Dottorato di Ricerca in Linguistica Generale,

Storica, Applicata, Computazionale e delle Lingue Moderne

L-LIN/01

Tesi di dottorato:

Copulae and Classifiers in the Arabic Noun Phrase

Candidato:

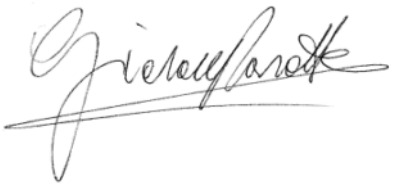
Francesco Grande

Tutori:

Prof. Giovanna Marotta

Prof. Antonella Ghersetti

Presidente: Prof. Giovanna Marotta



Triennio 2008-2010

A Veronica

Ritiene che le categorie siano ancora attuali?
Se interpretate in maniera metafisica, oggi sono anacronistiche.
Reinterpretate in maniera grammaticale, invece, continuano a esprimere le basi dell'analisi logica: sostantivi, aggettivi (quantitativi e qualitativi), relazioni, avverbi (di luogo e di tempo), verbi ausiliari (essere e avere) e forme verbali (attiva e passiva).

Piergiorgio Odifreddi, *Intervista ad Aristotele*

Copulae and Classifiers in the Arabic Noun Phrase

Francesco Grande

Acknowledgements

I would like to thank my advisor Giovanna Marotta for supporting from the onset my idea of investigating the syntax-phonology interface and giving me helpful advice concerning it; and my advisor Antonella Ghersetti for guiding me toward my focus on the Arabistic line of reasoning and her suggestion to investigate Arabic linguistic facts from the challenging point of view of the relationship between Old Arabic and Modern Colloquial Varieties.
I would also like to thank Jamal Ouhalla, for accompanying me throughout the entire work in transforming syntactic insights into scientific arguments and into a minimalist account as well.
Many thanks to Andrea Moro for encouraging me to develop some hypotheses on the relationship between syntactic theory and the Saussurean linearity and for reading Chapters 5, 6 and 8 of the dissertation. His feedback helped me to significantly improve the quality of my work.
I would like to express my gratitude to Jan Jaap de Ruiter, who responded with enthusiasm to the idea of reading my dissertation when we talked about it.
I am grateful to Bettina Migge for her attentive reading and invaluable comments on Chapters 1 and 2 and Na'ama Pat-el for reading some crucial sections of Chapter 6 and engaging in a high-level and stimulating discussion about them.
I am indebted to all the scholars I have had the fortune to exchange ideas with during my PhD studies: Giuliana Giusti, Tobias Scheer, Paolo Acquaviva, Marcel den Dikken, Giuliano Lancioni, Kristen Brustad, Giuseppe Longobardi, Paola Crisma, David Lightfoot, Alessandra Giorgi, Ur Shlonsky, Benjamin Fagard, Massimiliano Spotti and in particular Guglielmo Cinque and Jean Lowenstamm for introducing me to theoretical syntax and phonology respectively.
Thanks, as well, to my fellow PhD candidates for all the lively discussion about linguistic issues: Gabriele Rigon, Irene Tinti, Emanuela Sanfelici, Gabriele Lenzi, Marc Traebert, Giuseppina Silvestri, Martina Johnson, Carlo Arciero.
Special thanks to my Arabist colleagues in Milan - Paolo Luigi Branca, Roberto Bellani, Myrna Chayo - as well as to Wolfgang Sahlfeld for their continuing support of my professional life.
I would also like to thank my inspirational extended family and my friends who helped me to achieve this goal: Mary, Dido and Michi.
Last but not least, thanks to my wife, Veronica, who made this possible.

Abstract

This thesis proposes a reconsideration of the Arabic adnominal markers and of its modification structure, on the ground that the current analytical trend, which likens them to their Indo-european counterparts, runs into several problems. It develops an unified account of the adnominal markers, and argues that all of them are to be understood as copulae recursively embedded into the Arabic modification structure. Modifiers too are interpreted here in an unified way, all of them being rethought of as a non-restrictive relative clause, at least originally.
Among the consequences of this approach are reinterpreting nunation as a Numeral Classifier and the prenominal article as a medial article. As for case-endings, the present dissertation revives the Semitistic hypothesis that underlying to them is a morpheme *w* and proposes to identify it with the Conjunction *w* 'and' widespread in Arabic and Semitic, since this kind of conjunction crosslinguistically (e.g. in English, Somali) mediates the predicational relationship between the NP / Subject and the non-restrictive relative clause / Predicate. Moreover, it is argued here that Arabic case is actually a semantic opposition Subject vs. Non-subject that develops out of a copular opposition between lack vs. presence of Predicate Inversion.
The copula analysis of the Arabic modification structure has also some interesting implications for linguistic theory in general, because it provides a definition of word as well as a better understanding of the fusive, agglutinative and isolating morphological types.
This thesis capitalizes on the findings of some recent threads of research. On diachronic level, it adopts Owen's theory about Arabic dialects and Classical Arabic, according to which the former are older than the latter, as well as Garbini's and Durand's view that Arabic is a mixed language made up of an Amorite and a Pre-semitic parastrate. In synchrony, it follows Moro's theory of Dynamic Antisymmetry, which explains the displacement phenomenon typical of natural languages as a consequence of another key-property of them, namely Saussurean linearity.

Contents

1 Introduction **5**

1.1 General remarks . . . 5

1.1.1 Research topic . . . 5

1.1.2 Research rationale . . . 6

1.1.3 General assumptions and theoretical orientation . . . 7

1.1.4 Main claims and purpose . . . 14

1.2 Organization of the dissertation . . . 15

1.2.1 Chapter 2 . . . 15

1.2.2 Chapter 3 . . . 16

1.2.3 Chapter 4 . . . 17

1.2.4 Chapter 5 . . . 18

1.2.5 Chapter 6 . . . 20

1.2.6 Chapter 7 . . . 21

1.2.7 Chapter 8 . . . 23

I Arabic in its own terms **25**

2 The multi-faceted nature of Arabic language **26**

2.1 Preliminary remarks . . . 26

2.2 The Arabic language: the state-of-the-art . . . 27

2.2.1 Synchrony: the diglossia . . . 27

2.2.1.1 High Variety and Low Variety . . . 28

2.2.1.2 Classification of the Low Varieties . . . 29

2.2.1.3 Najdi (and Hijazi) Arabic . . . 30

2.2.2 Diachrony: the stages of the language under investigation . . . 32

2.2.2.1 Arabic and the Semitic family . . . 33

2.2.2.2 *Stammbaumtheorie* and *Wellentheorie* . . . 34

2.2.2.3 The 'logical matrix' and Owens's criticism . . . 43

2.2.2.4 Old Arabic . . . 47

2.2.2.5 Classical Arabic . . . 49

2.2.2.6 Old Arabic and modern Colloquial Arabic . . . 50

2.3 The Arabic linguistic tradition: an overview . . . 51

2.3.1 Object, aims and problems . . . 53

2.3.2 The transmission of the corpus: *muṭṭarid* and *šādd* . . . 57

2.3.3 The systematization of the corpus: *'aṣl* and *far'* . . . 59
2.3.4 An integrated methodology . . . 62
2.3.5 The primary sources of this work . . . 64
2.4 Summary . . . 66

3 The Arabic Noun Phrase: (rethinking) the received view 75
3.1 The Arabic Noun Phrase: the state-of-the-art . . . 75
3.2 The received view . . . 76
3.2.1 Case, number, gender, definiteness . . . 76
3.2.2 The modification-structure . . . 80
3.2.2.1 Adjectives and appositions . . . 81
3.2.2.2 Construct State and Free State . . . 108
3.2.2.3 Relative Clause . . . 115
3.3 Some open issues . . . 126
3.3.1 Case . . . 127
3.3.2 Gender and number . . . 129
3.3.3 The article . . . 139
3.3.4 Tanwīn . . . 141
3.3.5 Adjective . . . 143
3.3.6 Construct State and Free State . . . 144
3.3.7 Relative Clause . . . 147
3.4 Summary . . . 147

II A broader perspective 150

4 The minimalist framework: (Dynamic) Antisymmetry 151
4.1 Introduction . . . 151
4.2 Syntax, lexicon and the item-and-arrangement model . . . 152
4.3 Form and meaning . . . 154
4.4 More on syntax: the generative and transformational components . . . 156
4.5 The Minimalist Program: the biological foundations of language . . . 159
4.5.1 Syntax as a human faculty and recursion . . . 161
4.5.2 The Performance systems and displacement . . . 165
4.5.3 Phrase structure: Antisymmetry and Cartography . . . 170
4.5.4 Movement: Dynamic Antisymmetry . . . 186
4.6 Summary . . . 192

5 Predication in the NP: a crosslinguistic survey 194
5.1 A syntactic and crosslinguistic theory of predication . . . 194
5.2 Functional items and the theory of predication . . . 196
5.2.1 The Relator . . . 197
5.2.2 The Linker . . . 208
5.2.3 Predication and non-restrictiveness . . . 211
5.2.4 Predication and comparison . . . 219

5.2.5 The Reinstated Word Order . . . 223
5.3 Towards an unified syntactic theory of predication . . . 227
5.3.1 Relators and Base-generation . . . 227
5.3.2 Heads and Points of Symmetry . . . 230
5.3.3 More on Heads and Points of Symmetry . . . 234
5.3.4 Heads and copulae . . . 249
5.3.5 Triggers of movement . . . 250
5.3.6 Appendix . . . 253
5.4 Summary . . . 255

III Copulae and Classifiers in the Arabic NP 260

6 Phrase structure 261
6.1 Aims and introduction . . . 261
6.2 Reducing the typology of the Arabic RC . . . 262
6.3 The Amorite Relative Clause Type . . . 264
6.3.1 Diagnostics . . . 265
6.3.2 The minimalist analysis . . . 268
6.3.3 Further diagnostics and developments . . . 274
6.4 The Pre-semitic Relative Clause Type . . . 278
6.4.1 Diagnostics and developments . . . 279
6.4.2 The minimalist analysis . . . 280
6.4.3 Deriving the prenominal article . . . 284
6.4.4 The minimalist analysis . . . 289
6.5 Generalizing the copular analysis . . . 290
6.6 Classifiers in the Arabic NP . . . 291
6.6.1 Diagnostics and minimalist analysis . . . 296
6.6.2 The Possessive Classifier . . . 299
6.6.3 Deriving the non-pronominal Construct State . . . 306
6.6.4 The minimalist analysis . . . 314
6.7 Summary . . . 319

7 Towards a reconsideration of Arabic Case 323
7.1 The copula analysis of Case . . . 323
7.2 Arabic Case and subjecthood . . . 324
7.3 Arabic case and Hamitic . . . 325
7.4 Arabic case and prepositions . . . 327
7.5 Arabic Case and dialectal variation . . . 331
7.6 The fine-grained structure of the copulae *u, a, i* . . . 338
7.7 Deriving Arabic Case via Predicate Inversion . . . 345
7.8 Language Contact Hypothesis and dialectal variation . . . 347
7.9 Refining the Classifier Hypothesis . . . 349
7.9.1 Deriving dual . . . 349
7.9.2 Towards an unified account of Arabic plurals . . . 352

7.9.3 The bound -*t* as a Numeral Classifier . . . 355
7.9.4 Numeral Classifiers: an overall view . . . 359
7.10 Summary . . . 361

8 Transformations **364**
8.1 Phrase structure: an overview . . . 364
8.2 The universal hierarchy of functional projections and the Arabic NP . . . 367
8.3 The POS DP - N: from Spec, NP to Spec, AbsP . . . 369
8.4 The POS Abs - NP: Remnant movement of N to Spec, ConjP . . . 371
8.5 The POS Conj - DP: representational and derivational neutralization . . . 374
8.5.1 Representation: Abstract Incorporation in situ . . . 377
8.5.2 Derivation: From Spec, AbsP to Spec, DefP . . . 377
8.6 The POS Def - NP: representational and derivational neutralization . . . 379
8.6.1 Representation: Abstract Incorporation in situ . . . 380
8.6.2 Derivation: From Spec, ConjP to Spec, FocP . . . 383
8.7 The POS D - NP: Abstract Incorporation . . . 384
8.8 Residual issues . . . 386
8.9 Relativized Minimality in the pronominal Construct State . . . 390
8.10 Theoretical implications . . . 395
8.10.1 An inversion of perspective . . . 395
8.10.2 Parameter-setting as POS neutralization . . . 396
8.10.3 A note on den Dikken's Linker . . . 400
8.10.4 The syntax of predication as a 'chain reaction' of POS . . . 402
8.10.5 Deriving optional movement in secondary predication . . . 402
8.11 Summary . . . 405

9 Conclusions **407**
9.1 Main findings and results . . . 407
9.1.1 Arabic studies . . . 407
9.1.2 Minimalist Program . . . 412
9.1.3 Linguistic theory . . . 413
9.2 Synopsis . . . 414
9.2.1 Copulae in Arabic: a taxonomy . . . 414
9.2.2 Copulae in Arabic: the 'Russian-doll' effect . . . 415
9.2.3 Classifiers in Arabic . . . 415
9.2.4 The 'chain reaction' of POS . . . 415

References **416**

Chapter 1

Introduction

1.1 General remarks

1.1.1 Research topic

This thesis investigates the Arabic Noun Phrase in two main respects. first, it is a survey on the adnominal markers associated with the Arabic noun. Second, it deals with the Arabic modification, that is with the relationship that holds between a noun and its modifier, which can be compared to that holding between a subject and its predicate (den Dikken, 2006).
These two relationships will be also referred to here as SECONDARY PREDICATION and PRIMARY PREDICATION, respectively, along the lines of den Dikken (2006).
The former topic is represented by the markers of:

- number
- gender
- (in)definiteness
- case

as well as by the so-called

- SINGULATIVE

which will be described in greater detail in Section 1.1.3.
The latter topic revolves around the presence of a copular element intervening between the noun and one of the following modifiers:

- an adjective
- a relative clause
- a genitive phrase

These constructions are all attested and exemplified in Djidjelli Arabic, a variety spoken in Algeria, where the copular element is the allomorphic base *d-*:

(1) Djidjelli Arabic (Marçais, 1956, p. 505), (Pennacchietti, 1968, p. 42)

n-nās ĕddĕ-kbār
the-people GM-bigs

'elderly people'

(2) Djidjelli Arabic (Marçais, 1956, p. 418), (Pennacchietti, 1968, p. 41)

l-ḥāl eddi-ḫūyi
the-state GM-brother-me

'my brother's state'

(3) Djidjelli Arabic (Marçais, 1956, p. 493)

l-rāžəl əddi ža
the-man GM came

'the man who came'

The (socio-)linguistic scenario in which these phenomena occur is extremely rich, both at diachronic and synchronic level (Versteegh, 1997a; Suleiman, 2006; Bassiouney, 2009): as a first approximation, Arabic can be thought of here as a complex of varieties rather than as a compact entity. Ch. 2 covers this topic extensively.

1.1.2 Research rationale

The two topics under scrutiny have been given a different amount of attention in the literature. While the Arabic adnominal markers have been the subject of extensive work by both Arabists and Semitists (see Section 3.3), the predicational nature of the Arabic modification and its morphosyntactic reflex, ie the realization of a copular element in the structure, has been recognized only recently.
It should be added that the first systematic study in the field (Pennacchietti, 1968) has not been totally overtaken by later work (Ferrando 2000, Owens 2006, Ayoub 1991, Ayoub 2006) which have a narrower theoretical scope and a more limited empirical coverage.
Yet, in spite of the attention they have received, the Arabic adnominal markers remain obscure in many respects and warrant further investigation. It is exactly in this connection that the 'predicational' approach to the Arabic modification-structure turns out to be appealing, since it provides new interpretive tools in the analysis of the adnominal markers.
A case in point is the adnominal marker *-(i)n* that in the Bedouin varieties typically intervenes between a noun and an adjective. This structure has an indefinite meaning and therefore *-(i)n* is traditionally taken to be an exponent of indefiniteness. Nevertheless, this view can be hardly deemed correct, since indefiniteness is usually conveyed by a bare noun.
As Owens (2006) interestingly suggests, the most straightforward way to overcome this impasse is interpreting the adnominal marker *-(i)n* as a reflex of the predication-relationship holding between the noun and the adjective, ie as a copular element[1] (data from Najdi Arabic, a Bedouin variety):

(4) Najdi Arabic (Ingham, 1994, p. 48)

bēt
house

[1] Adopting a diachronic perspective, this function of *-(i)n* is a later development: see Chapter 6.

'a house'

(5) Najdi Arabic (Ingham, 1994, p. 54)

bēt-in ṭuwīl
house-GM tall

'a tall house'

The scenario that emerges form the discussion so far is twofold:

(I) Combining a predication-oriented analysis of the Arabic modification-structure with the investigation of the adnominal markers is a promising line of inquiry, which is worth developing from its embryonic state to a more fully-fledged form.

(II) Such an approach leads to the expectation that in the best case the Arabic adnominal markers are copular elements, or, alternatively, that they are somehow related to copular elements.

In particular, (II) relies upon the well-known Saussurean axiom that linguistic entities are parts of a system. This idea is pursued to its limits here: Arabic adnominal markers are taken to be not separate entities but different facets of the same phenomenon. An alternative approach, which concentrates on a single adnominal marker, is conceivable, but as a matter of fact all the analyses carried out so far in this spirit are not totally successful, as a review of the literature proposed in Chapter 3 will make clearer.
To sum up, (I), (II) are the two key-ideas that will be pursued and implemented here: in particular, (I) is the working hypothesis of this work, and (II) its aim.

1.1.3 General assumptions and theoretical orientation

This dissertation revives Pennacchietti (1968)'s analysis of the Arabic (and Semitic) Noun Phrase, which focuses on the role that syntax plays in this domain. There are (at least) two aspects to this syntactic approach.
First, particles that are traditionally taken to be articles, relative pronouns etc... are deprived of meaning and simply regarded as copular elements - elements with no semantic import occurring in a modification-structure between the noun and its modifier[2].
Pennacchietti dubs them NON-AUTONOMOUS-PRONOUNS and describes them as elements that:

> "require the presence of a complementary element, either a complement of specification or a relative clause. [Thus they are] void of a true semantic meaning, and simply serve as nuclei of analytic expressions of nominal character. "
> (Pennacchietti, 1968, p. 65, translation by Buccellati 1972, p. 296)

Second, the presence or absence of a given copular element is analyzed in comparative terms, by performing a feature-by-feature comparison between two or more Semitic languages. Thus, Pennacchietti (1968, p. 63) draws a dividing line between Arabic, where a copular element is allowed to occur between the noun and the adjective (cp. (1)) and other Semitic languages,

[2]What is actually modified is a Noun Phrase, since a noun associated with its modifier via a copular element can be considered as a whole and associated in its turn with another modifier, via a new copular element, and so on. To keep the structure transparent, however, the simplified scenario shown in the main text will be adopted.

where it is not (see Chapter 6 for details).
The approach thus characterized enjoys a renewed interest in Semitistic and Arabistic literature, although Pennacchietti's work is not always mentioned and evidence is gathered independently (Owens, 2006; Ferrando, 2000; Ayoub, 1991, 2006): in this respect, the present dissertation converges with a current trend in literature.
Nevertheless, it is felt here that the comparative perspective adopted by Pennacchietti is characterized by two non-trivial theoretical assumptions and that the researcher espousing his view should be aware of them.
On the one hand, the diachronic scenario Pennacchietti (1968) sketches out does not totally match the standard view. On the other hand, Pennacchietti's classification of the syntactic constructions displaying the copular elements leaves some linguistic facts out of the picture.
As for the diachronic scenario, what follows is a case in point. Pennacchietti (1968, pp. 4ff.) observes that the presence of copular elements inside the Arabic Noun Phrase is an old phenomenon in Semitic, already attested in an ancient language like Akkadian (3000 BC, approximately: Garbini & Durand 1994, p. 29), where the base *š-* performs the same copular function of (Djidjelli) Arabic *d-* (compare (6) with (1,2,3)).
Furthermore, he documents various manifestations of this phenomenon in some modern Colloquial varieties of Arabic and in the Hamitic family. For example, in these languages there occurs a construction, the so-called FREE STATE, in which the copular element *d-* (in Arabic) or *ō* (in Somali) intervenes between a noun and the genitival phrase modifying it (cp. (2), here repeated as (8) and (10), respectively).
Notice, however, that the Free State is not found in the Classical variety of Arabic (Pennacchietti, 1968, pp. 73, 83), where the noun seems to be directly associated with its genitival phrase, with no intervening copular element (and no preposition). This construction is usually called CONSTRUCT STATE, and is documented in Akkadian as well (cp. (7)):

(6) Akkadian (Caplice & Snell, 2002, p. 20)

šarr-u-m dann-u-m ša māt-i-m
king-NOM-GM mighty-NOM-GM GM land-GEN-GM

'mighty king of the land'

(7) Akkadian (Owens, 2006, p. 85)

beel biit-i-m
master house-GEN-GM

'master of the house'

(8) Djidjelli Arabic (Marçais, 1956, p. 418), (Pennacchietti, 1968, p. 41)

l-ḥāl eddi-ḫūyi
the-state GM-brother-me

'my brother's state'

(9) Classical Arabic (Wright, 1896, vol. II, p. 199)

ḥarr-u l-šams-i
the.heat-NOM the-sun-GEN

'the heat of the sun'

(10) Somali (Pennacchietti, 1968, p. 62)

níman-ka kalé ō béled-ka
men-the others GM village-the

'the other men of the village'

Of particular relevance here is that Arabists and Semitists usually interpret the Construct State (9) as older than the Free State (8), (8) allegedly being an analytic construction that developed out of the synthetic one (9) (see Owens 2006, pp. 111-113 and references therein). More generally, they regard the Colloquial varieties, among them the Bedouin ones, as derived from the Classical variety (see e.g. Moscati et al. (1980)).

Nevertheless, on a broader comparison, nothing supports their claim, because in Akkadian, one of the oldest manifestations of Semitic, not only the Construct state (7), but also the Free state (6) occurs. This is tantamount to saying that precisely in the same way as the Classical variety of Arabic shares with Akkadian old features like, say, some lexical items, the Colloquial varieties (and, remarkably, Somali as well) share with it the old feature of (*li*-less) Free State[3] (see also Owens 2006, pp. 111-113 for similar arguments).

Notice that pretheoretically the feature 'Free State' stands as an old isogloss regardless of Pennacchietti's interpretation. Recent work on the Free State (Ouhalla 2010, Bardeas 2009, and cp. also Chapter 6) strongly suggests that Pennacchietti's copular analysis of *d*-like elements is generally to be rejected in favor of a preposition-analysis, but this has no bearing on the fact that the Free State is an archaic construction.

Therefore, it arguably follows from Pennacchietti's account that:

(I) in Arabic the Colloquial varieties are at least as old as the Classical one, if not even older

(II) old Semitic features are still preserved in modern Hamitic languages

But this is exactly what is currently claimed in less 'orthodox' threads of Semitistic and Arabistic studies. As it will be seen in the Section 2.2, in the Arabistic literature Owens (2006) is mainly devoted to provide evidence in favor of (I), and is aware of (II), especially as far as adnominal markers are concerned (Owens, 2006, pp. 82-83). On the other hand, in the Semitistic literature, Garbini and Durand agree with both (I) and (II) (see Garbini & Durand, 1994, pp. 66, 173).

In particular, Owens's work dovetails with Garbini and Durand's in one interesting respect, namely that Arabic arises through language contact, involving a (relatively) recent Semitic superstrate, and an old Semitic, if not even Hamitic, substrate (see Owens 2006, p. 101, fn. 22, Garbini & Durand 1994, pp. 144-6 and *passim* and cp. also Anghelescu 2004, Part I, Ch. 4).

It seems therefore safe to maintain that Pennacchietti's comparative perspective implies a diachronic scenario along the lines just described, rather than along the lines of a more standard view, roughly exemplified by Moscati's position.

The second aspect of Pennacchietti's comparative perspective that needs some discussion is

[3]The Free State at issue is characterized by the lack of the preposition *li-* and should be kept distinct from the Free State exhibiting it, which is documented (Pennacchietti, 1968, pp. 53-54) only from Phoenician onwards (XIV c. BC).

the way he applies to syntax one of the most typical ingredients of comparativism, namely the taxonomical methodology (Lehmann, 1993, Ch. 2).
Buccellati (1972, p. 297), in fact, observes that Pennacchietti's account although providing an accurate listing of all the construnctions in which the copular elements are found, is not able to capture the TRANSFORMATIONS, ie the systematic relations holding among them and thus seems to be flawed in this crucial respect.
The relations at issue, usually called transformations in the American structuralism (see Section 4.4 for further discussion), are basically the native speaker's feeling that there exists a deep connection between two (or more) given syntactic constructions, and that this connection is such that the former construction is intuitively more 'basic' than the latter(s), the latter(s) being 'somehow' derived from the former (see e.g. Lyons 1968, p. 247ff., Chomsky 1995, p. 23). Thus, in English, active sentences are more basic than passive sentences, declarative sentences more basic than interrogative sentences etc...
A clarification at this point is in order. Although the phenomenon of transformation is a major concern of the Chomskian linguistic theory, it is arguably a pretheoretical datum, already observed in the traditional grammar as well as in the American structuralism preceding the Chomskian theory (Lepschy, 1966, pp. 179, 186-7). Therefore the phenomenon per se should be kept distinct from its most famous theoretical account, worked out in the Chomskian tradition. As Lyons puts it:

> "The term 'transformational' has unfortunately engendered a good deal of unnecessary controversy and confusion in the recent literature of linguistics. If we use the term in a general and rather informal sense, rather than in the particular sense in which it is defined in anyone theory, we can say, quite reasonably, that the 'deeper connexions' between sentences which 'cut across the surface grammar' (cf. the quotation from Hockett given above) are transformational relationships: this is a perfectly legitimate use of the term 'transformational'.
> Many of these transformational relationships between sentences are well-recognized in traditional grammar; but it is only recently that linguists have made any progress in accounting for them in an explicitly generative framework. Any grammar that claims to assign to each sentence that it generates both a deep-structure and a surface-structure analysis and systematically to relate the two analyses is a transformational grammar (whether it uses the label or not)".
> (Lyons, 1968, p. 248)

A particularly fitting example is Buccellati (1996)'s *A Structural Grammar of Babylonian*, which makes use of transformations within a non-Chomskian framework.
As far as the Arabic Noun Phrase is concerned, native speakers[4] are aware of a (deep) connection that ties a Construct State like (11) to a Free state like (12). Although they do not single out a copular element in the Classical Arabic Free state (cp. (5), (8)) and do not discuss the different

[4]When data from Classical Arabic and Old Arabic are used (see Ch. 2 for a definition of these terms), the native speakers referred to here are to be understood as the informants selected by the Early Arab grammarians (about 750 AD) according to their elicitation criteria (see e.g. Alhawary 2003). It is worth noticing in this connection that the Late Arab Grammarians repeat and take extracts from the Early Arab Grammarians' data, especially Sībawayhi (d. 177/798) (Fleisch, 1961, vol. I, p. 46, Bohas et al., 1990, p. 31, Owens, 2006, p. 7) to the effect that the variety of Arabic they describe can be hardly subsequent to 750 AD (Owens, 2006, p. 7).

degree of definiteness encoded in (11) and (12), they consider the Construct State as derived from the Free State via deletion of the adnominal marker *-n* and of the preposition *li-* 'to' or *min* 'from' (Bohas et al., 1990, p. 63). Accordingly, (11) is in modern terms a transformation of (12) (Owens, 1988)[5]:

(11) Classical Arabic (Bohas et al., 1990, p. 63)

ġulām-u Zayd-i-n
page-boy-NOM Zayd-GEN-GM

'Zaid's page-boy'

(12) Classical Arabic (Bohas et al., 1990, p. 63)[6]

ġulām-u-n li-Zayd-i-n
page-boy-NOM-GM to-Zayd-GEN-GM

'a page-boy of Zayd'

(13) (12) → TRANSFORMATION → (11)

(12) = N-*n* *li*-NP

TRANSFORMATION = Delete *-n*, *li-*

(11) N NP

Pennacchietti, however, does not take this piece of information (summarized in (13)) into account and interprets (11) and (12) as two different constructions.
Nevertheless, as a comparison between (11) and (12) shows, resorting to a transformational perspective is desirable not only for empirical reasons, because of its consistency with the native speakers' judgments, but also from a theoretical standpoint, since it provides a more economical account: Free State and Construct State turn out to be the same construction.
What is therefore needed is a theoretical framework that takes care of transformations: among the possible choices, the so-called Minimalist Program, a generative-transformational model, will be adopted here. The reason of this choice lies in the fact that the Minimalist Program is compatible with the two main claims of Pennacchietti's work discussed above, namely the comparative method applied to the domain of (morpho-)syntax, and the presence of copular elements inside the (Arabic) Noun Phrase.
In fact, as it will be discussed in the Chapter 4, the former aspect is accommodated within an influential research thread of Minimalist Program, the so-called Cartography, the latter into a recent development of the Minimalist Program, namely den Dikken (2006)'s theory of Relators and Linkers. It will be also discussed that these two branches of the Minimalist Program are compatible, because they are based upon a common theoretical background: Kayne (1994)'s Antisymmetry.

[5]The following is an example artificially constructed by Arab Grammarians for the sake of commodity in order to illustrate and discuss the more common properties of the Construct State. As such, it is meant to provide a descriptive generalization of the just discussed instances of Construct State like (9) etc...See Guillaume (2006) for details.

[6]From a diachronic vantage point, one may cast some doubts on identifying the deleted particle with the relatively recent preposition *li-* (cp. fn. 3) and may wish to seek a more archaic particle. This is discussed in Chapter 6.

In addition, the Minimalist Program, has the advantage of resorting to the version of comparative method developed by Greenberg (1963) and known as TYPOLOGY. This latter has as its object (morpho-)syntactic constructions ('types') by themselves, which do not necessarily belong to the same language-family. Such an approach makes it possible to extend the domain of inquiry and the empirical coverage to non-Semitic instances of Noun Phrase, provided they share with the Arabic (and Semitic) Noun Phrase similar morphosyntactic features.

Interestingly, this expectation is met in the case of the so-called NUMERAL-CLASSIFIER LANGUAGES, which are mainly found in East and Southeast Asia. They can be roughly described as languages where expression of number is not realized through suffixation, as in the Indo-european languages, but via insertion of a complex made up of a quantifying expression (e.g. a numeral) and the so-called singulative morpheme. This latter is a form denoting the individuation of a single unit from a collectivity, exemplified by the English word *head* in the expression *two heads of cattle*.

This results in a tripartite structure, usually called NUMERAL CLASSIFIER CONSTRUCTION, which is illustrated in (15).

As for Arabic, Bauer (1912) and Greenberg (1990b) liken the ending *-at, -ah* to a singulative morpheme (see Ch. 7 for more details):

(14) Old Arabic (*Šarḥ al-Kāfiyah*, vol. III, *al-Jam'*, p. 325)
naml: 'ants (COLLECTIVE)'→
naml-ah: 'a single ant (M / F, SINGULATIVE)'

and, more generally, they compare constructions involving *-at, -ah* as well as other morphemic material with the Numeral Classifier Construction (although the matching is sometimes partial: contrast (15)with (17)):

(15) Mandarin Chinese (Greenberg, 1990b, p. 188)

san bĕn shu
three CLF book

'three books'

(16) Omani Arabic (Greenberg, 1990b, p. 178)

'ashri:n ra:s finda:l
twenty CLF potato

'twenty potatoes'

(17) Omani Arabic (Greenberg, 1990b, p. 179)

thala:th baqr-a:t
three cow-CLF

'three cows'

Furthermore, numeral-classifier languages match with Arabic in that their Noun Phrase inserts a copular element between a noun and its modifier:

(18) Thai (den Dikken, 2006, p. 230)

khon thîi kéng
person GM smart

'the/a smart person'

In the typological literature, the former parallelism has been long discussed by Greenberg (1990b) (cp. also Aikhenvald 2000, pp. 120-121), the latter has been observed by Gil (2005, p. 247), which lists among the languages displaying a copular element between the noun and its modifier not only the numeral-classifier languages, but also Jewish Arbel Aramaic, a Semitic language.

The parallelism is even stronger if one agrees with Greenberg (1990b) in interpreting the singulative as an element that obligatorily mediates the relationship between the noun and its numeral modifier - which is precisely the function of a copular element. The American scholar, in fact, points out that in the numeral-classifier languages there co-occur

> "the aversion of collectives to direct construction with a numeral and the intervention of an individuated noun, the classifier, as one of the devices to avoid this direct confrontation".
> (Greenberg, 1990b, p. 184-185)

It ensues that the comparative-typological analysis carried out in this work, while focusing on the presence of copular elements in the Arabic Noun Phrase, cannot leave aside the issue of the Numeral-Classifier Construction. In so doing, this thesis is in line with recent trends in theoretical linguistics (Zabbal, 2002; Fassi Fehri & Vinet, 2004, 2007; Ouwayda, 2009), which explore the expression of number in Arabic in the light of the Numeral-Classifier Construction typical of Chinese, Japanese etc...

Returning to the advantages of the minimalist framework, it should be mentioned the possibility of working out an explanation of the (morpho-)syntactic facts which takes into consideration how (morpho-)syntax proper[7] interacts with non-syntactic factors, be they pragmatic or phonological.

For example, one may wonder whether in the Classical variety of Arabic the loss of the Free State featuring the *d*-element discussed in Pennacchietti (1968, Ch. 3) has to do with the extremely unstable nature of the phoneme *d* observed by Greenberg (1950).

To summarize, this dissertation revives Pennacchietti's approach to the Arabic Noun Phrase, which focuses on the presence of copular elements and adopts a comparative perspective. Such an approach is further developed into a generative-transformational framework, the so-called Minimalist Program. As a consequence, the comparative method is less constrained in terms of diachrony and genetical relationships, ie it is typological. This impacts on the object of the analysis, which will include, when relevant, phenomena that traditionally are regarded as not or only loosely related with Arabic, like the Numeral-Classifier Construction.

[7] This dissertation is situated in the tradition usually taken to begin in the American Structuralism's item-and-arrangement model, where no distinction is made between morphology and syntax: see Thornton (2005) for details. Thornton (2005, pp. 88-9) provides textual evidence that this unified treatment of morphology and syntax probably traces back even earlier, to Saussure. Be that as it may, since the generative theory takes as its starting point the item-and-arrangement model, the same architectural settings hold, at least until the development of the minimalist model, which either preserves the undifferentiated morphosyntactic scenario, or departs from it, according to its several branches (Thornton, 2005; Baker, 2002; Leu, 2008). In particular, the version of Minimalist Program adopted here, being substantially antisymmetric and cartographic, is conservative in that it relies upon the assumption that morphology is a part of syntax (Leu, 2008). These issues are discussed at length in Chapter 4.

Also, the Minimalist Program makes it possible to work on a broader (morpho-)syntactic scenario, which is also concerned with the interplay between (morpho-)syntax and discourse-related / phonological factors.
Nevertheless, when diachrony is taken into account the scenario adopted is consistent with Pennacchietti (1968)'s work and subsequent developments, according to which the Arabic Colloquial varieties are (at least) as old as the Classical one, and Arabic on the whole includes elements of an ancient Semitic or Hamitic substrate.

1.1.4 Main claims and purpose

In this work all the Arabic modifiers are taken to be (at least originally) non-restrictive relative clauses, except the pronoun of the Construct State. This proposal basically exploits Medieval Arab grammarians' view that in Arabic the modifier is a noun, ie an apposition, a construction traditionally taken to be very close to a non-restrictive relative clause (cp. the traditional synonim 'appositive relative clause').
Furthermore, the adnominal markers of the Arabic Noun Phrase are accounted for either as copular elements or as related elements: among these latters, the classifiers (see Section 1.1.3). Accordingly:

- The so-called (PRENOMINAL OR ADJECTIVAL) ARTICLE *l-* is a copula mediating the relationship between an (understood) *demonstrative-like* element and a noun. This holds true also for the *l-* introducing the modifier. The Arabic modifier in fact is an appositive relative clause and as such features an anaphoric *demonstrative-like* element and a noun. More precisely, the demonstrative-like element at issue is a covert OPERATOR.

- The so-called TANWĪN, ie the exponent *-n*, is an (animate) Classifier. Since the Classifier has a nominal nature, it is embedded inside a non-restrictive relative clause as much as the other Arabic modifiers. It later evolves into a General Classifier and then into a copula in complementary distribution with *l-* introducing the modifier.

- Arabic Case is at least originally a (composite) copula (cp. Nyberg 1920) mediating the relationship between a noun and its modifier. Its ability to express different semantic roles in Classical Arabic develops out of a binary opposition (still documented in Najdi Arabic) between the copulae *a, i*, capable to signal Predicate Inversion, and the copula *u*, not capable to do so.

- The so-called SINGULATIVE, *-(a)t* is an (inanimate) Classifier. It later turns into a feminine marker because the General Classifier *-n* takes over its original function (cp. Bauer 1912 and Greenberg 1990b).

- The relative marker *lla-d̲ī* and its allomorphs are better understood as a copula mediating the relationship between a noun and its modifier (cp. Pennacchietti 1968). It survives as a relic and has been generally replaced by Case.

- The semi-lexical noun that in the Colloquial Arabic varieties precedes the genitival complement of the Free State is regarded here as a (Possessive) Classifier. It is also taken here to be the intermediate stage in a (clinal) process of grammaticalization, which according

to Ouhalla (2010) has the Noun as its source and the Preposition as its target. Because of its nominal nature, the Classifier is embedded inside a non-restrictive relative clause.

- The covert particle which according to Arab Grammarians precedes the genitival complement of the Construct State is reinterpreted as a (Possessive) Classifier. It is also documented in cognate languages, where it has an overt realization, namely *ḏū*. In Classical Arabic, it can be replaced by the General Classifier *-n* (in the so-called TAMYĪZ-construction). Both *ḏū* and *-n* are subsequently dropped for phonological or semantic reasons. On this view, the Construct State is no more than a subcase of Free State (cp. Ouhalla 2010).

- This does not hold for the Construct State featuring a pronoun (cp. Ouhalla 2010). This latter, in fact, is a genitival complement directly annexed to the noun. This is due to the fact that semantically the entire construction typically signals inalienable possession.

All these claims will be motivated diachronically and typologically and then couched within a minimalist framework.
The present account makes extensive use of two phenomena, namely the coindexing between two adjacent word-level entities, and a phonological constraint, as depicted below:

(19) ...X_i Y_i...

(20) *$\underline{d}$; l

; = ordering irrelevant

1.2 Organization of the dissertation

1.2.1 Chapter 2

The Arabic language is better understood in synchrony as made up of Classical Arabic and the modern Colloquial Varieties. This Chapter introduces this state of affairs also from a diachronic perspective, relying upon recent research by Owens (2006). This latter plausibly demonstrates that Classical Arabic is an evolving entity, developing out of Old Arabic, while the modern Colloquial Varieties behave as a whole as an 'atemporal' entity (Owens, 2006, p. 114) insofar as some key-features are concerned - among them, lack of (morphological) Case.
In so doing, Owens's theory highlights the strong relationship between the modern Colloquial Varieties and Old Arabic and holds that at their core the former ones are as old as the latter. His claim is explored here with particular regard to a Bedouin Colloquial Variety, Najdi Arabic, as described e.g. in Ingham (1994), and is given a practical methodological implementation: analysis of the Arabic data of the present dissertation concentrates on the Colloquial Varieties and Old Arabic, assigning Classical Arabic a more modest role.
Chapter 2 follows Owens (2006) also in defining the nature, the chronological limits and the philological sources of Old Arabic and shows that such a definition is totally in line with characterization of the same linguistic entity provided in Semitic studies by Garbini & Durand (1994). Bringing these two theories in comparison has the advantage of accommodating Owens's diachronic theory on *Arabic* into the broader diachronic scenario of Garbini & Durand (1994), which describes the position of Arabic within *Semitic*. In greater detail, they regard Old Arabic

as a mixed language resulting from language contact between a Semitic - more precisely, Amorite - parastrate and a pre-Semitic parastrate, a linguistic system very close to Hamitic (Language Contact Hypothesis).
The practical implication of this move, introduced in this Chapter and developed in the next ones, is that the features typical of these parastrates seem to have played at least originally a role sensibly different from that designated by their traditional name. A case in point is the (putative) Case, probably to be rethought of as a copula (Nyberg, 1920), an analytical trend referred to here as Copula Hypothesis.
Finally, Chapter 2 provides in-depth discussion about the source of Old Arabic data, mainly drawn from Arab Grammarians' and Lexicographers' works, and defends their validity, stressing the point that they are framed within a mature scientific / linguistic paradigm.

1.2.2 Chapter 3

This chapter offers a critical review of the main data and hypotheses concerning the Arabic Noun Phrase (NP), as discussed in the received view and in less traditional threads of research. A closer and integrated investigation of already known data results in rejecting some commonly entertained hypotheses about the modification-structure in favor of two hypotheses, namely the Specificity Hypothesis (cp. Brustad 2000) and the Relative Clause Hypothesis (put forward for the Construct State in Bravmann 1961, and revived for the adjective in a recent study by Kremers 2003). Interestingly, such hypotheses are not mutually exclusive and provide an unified account of the Arabic modifiers, which can be *all* thought of as *specific* and *relative clauses*. In compliance with the results of Chapter 2, the present Chapter shows that this state of affairs holds not only for the modern Colloquial Varieties but also for Old Arabic, thus extending Brustad (2000).
After dismissing several interpretive trends of the mainstream Arabist and Semitist studies, Chapter 3 turns to consider and to revamp some alternative hypotheses worked out in literature. Of particular relevance among them is the Copula Hypothesis whose main proponent is Pennacchietti (1968), because of its ability to encompass not only (Classical Arabic,) Old Arabic and the modern Colloquial Varieties (South-Arabian included) but also Hamitic: that is, the same linguistic scenario discussed in Chapter 2. In particular, the Copula Hypothesis, if combined with Owens (2006) and Garbini & Durand (1994) interprets Case as a *copula* borrowed via a phenomenon of language contact from the *pre-Semitic parastrate* described in Garbini & Durand (1994), the scenario of borrowing being already hinted at in Owens (2006). As for number and gender, they are accounted for along the lines of the Classifier Hypothesis (see e.g. Bauer 1912, Greenberg 1990b), claiming, among others, that the SG marker *-at* is to be interpreted as a Classifier.
Finally, *l-, -n* are plausibly not to be regarded as markers of definiteness / indefiniteness. Instead, *-n* seems to act either as a copula (Copula Hypothesis: Owens 2006) or as a sign of nominality (Demarcative Hypothesis: see e.g. Ayoub 1991). On the other hand, *l-* is arguably an element originally intervening between a noun and its adjectival modifier, along the lines of Wackernagel's law (Copula Hypothesis: Testen 1998, Pat-el 2009). Last but not least Arabic language in general seems to belong to the agglutinating morphological type (Agglutination Hypothesis: Petracek 1981 and Garbini & Durand 1994 are but a sample).

The picture that emerges from this chapter is that a motivated account of the Arabic NP should take into account the Copula Hypothesis as well as its relationship with the Relative Clause Hypothesis, the Classifier Hypothesis and the Agglutination Hypothesis. Moreover, it is pointed out in this Chapter that the Copula Hypothesis, albeit grounded on a comparative, semantic etc... basis, is still vague in a crucial respect, in that it lacks *syntactic* evidence (e.g. diagnostics).

1.2.3 Chapter 4

Chapter 4 introduces the syntactic framework of the present dissertation: (Dynamic) Antisymmetry. This is a branch of the Minimalist Program (MP), the most recent incarnation of the Chomksian linguistic paradigm.

To begin with, the Chapter illustrates the basic architecture of grammar assumed in MP: one that starts with Lexicon, proceeds with Syntax and ends up 'bifurcating' into two components: semantic (C-I) and phonological (A-P). MP founds these architectural settings of grammar in biological and evolutionary terms: Syntax is a key-ingredient of human language only subsequent to some evolutionary process that exapted Syntax away from its original adaptive function (ie recursion: Chomsky et al. 2002, 2005). Then, the Chapter focuses on the evolutionary shift, driven by communicative needs, which turned Syntax and A-P from separate to interacting modules. This minimalist assumption, defended in Chomsky et al. (2002, 2005) has it that in primates other than humans etc... Syntax and A-P are incommunicado, while in humans they 'understand each other' because some cognitive strategies have been designed during the evolutionary process in order to ensure exchange of information between these so different systems. The same assumption, if translated in linguistic terms, means that Syntax and A-P are incommunicado in primates other than humans etc... because atoms of language (approximately, 'words') at A-P are ordered one after the other, given that their pronunciation puts them on the time axis (Saussurean linearity), but this is clearly not the case for Syntax. Here, in fact, atoms of language are *not* pronounced and therefore *not* located on the time axis - hence *not* ordered one after the other.

It also means that Syntax and A-P do interact in humans because the just mentioned cognitive strategies are to be understood in a linguistic sense as devices able *to bridge the gap* between lack of Saussurean linearity in Syntax (inability for atoms of language to be located on the time axis) and its presence at A-P (their ability to do so).

At this point, Chapter 4 explains which are the cognitive strategies at issue according to an influential branch of MP, the so-called antisymmetric program. Under this approach, they are basically two and consist of providing (human) Syntax with the same Saussurean linearity already found at A-P.

The first strategy provides Syntax with Saussurean linearity of A-P (thus putting Syntax 'in touch with' A-P) by forcing the atoms of language to combine in a special way, one that 'mimics' Saussurean linearity at A-P. Such a special way of combining them is what Bloomfield (1933) referred to as endocentricity / hierarchy. That is, Bloomfieldian endocentricity / hierarchy observable in Syntax is *not* a primitive property, but something modelled after Saussurean linearity at A-P. Technically, this special combination of atoms is either the so-called complex Spec - Head or the so-called X-bar format, two syntactic structures in complementary distribution:

the former is assigned to a substantive item, the latter to a functional one (Cinque, 2006).
The just described strategy was first posited by Kayne (1994) and is designated as Antisymmetry. Further developments include Cartography (Cinque & Rizzi, 2008, and references therein). The second strategy, referred to as Dynamic Antisymmetry by its proponent, Moro (2000), provides Syntax with Saussurean linearity of A-P by forcing the atoms of language already combined in an undesired way to re-combine in the legitimate X-bar format. Moro (2000) contends that the undesired combination at issue, which he calls Point of Symmetry (POS), is made up of two *unordered* atoms of language A and B. It is precisely such lack of ordering in the POS of Syntax that according to Moro prevents A-P from 'understanding' it, and then A-P can be enabled to 'understand' Syntax if lack of ordering of A, B is eliminated, which takes place leaving just one atom in situ, by *separating* A from B via displacement. This 'repair-strategy' is therefore what is usually designated as displacement (or movement) and is indicated in this work as *derivational neutralization of a POS*. A particular instance of POS is the configuration X - XP (Head - phrase) where XP has clitic-like properties.
Under this approach, the trigger of displacement in natural languages is the need to ensure exchange of information between Syntax and A-P, and Bloomfieldian endocentricity / hierarchy obtains not only by (a special) combination of two atoms but also through their displacement - hence the designation 'Dynamic Antisymmetry'. To put it shortly, this Chapter lays out a (dynamic) antisymmetric framework, where the phrase structure is thought of as a complex Spec-Head (substantive items) or as an X-bar format (functional items), and movement is triggered by the need to (derivationally) neutralize a syntactic structure which is neither of them, ie the POS.

1.2.4 Chapter 5

Chapter 5 explores and improves a *syntactic* version of the Copula Hypothesis as developed in the antisymmetric program for languages other than Arabic, setting the scene for its application to Arabic in the next Chapters. In this respect, the Chapter proposes a substantial revision of den Dikken (2006)'s antisymmetric theory of predication, especially relative to its manifestation within the NP. Since den Dikken's theory has a large empirical coverage (some Semitic, Indo-European, and Numeral Classifier Languages), Chapter 5 also lays out a cross-linguistic context, which makes it possible to cast the Arabic data set forth in Chapters 2, 3 within a typological and non-language-specific perspective.
According to den Dikken (2006), this consists of a three-fold typology of syntactic structures of predication in the NP ('secondary predication'):

- Canonical Predication: Thai *[khon] [kéng]*, lit. 'person-smart' ie 'smart person' (where the head noun conveys *old* information)

- Inverted Predication: Burmese [ca-naw weh] [thii] [sa-ouq], lit. 'I bought-copula-book ' ie 'the book I bought', itself to be further divided into attributive and comparative Inverted Predication

- Reinstated Word Order, a kind of hybrid type which displays the word order of Canonical Predication but also *the copula occurring in Predicate Inversion*: Thai *[khon] [thîi]*

[kéng], lit. 'person-copula of Pred. Inv.-smart', ie 'smart person'but with the head noun conveying *new* information

The revision of den Dikken's theory undertaken in this Chapter aims at fixing some of its shortcomings. The first, already observed in Tanase Dogaru (2008) is that he does not always provide an accurate description of the syntactic structure of the predication occuring in the NP. A second related problem is that sometimes he posits in the syntactic structure elements whose presence is not justified on semantic grounds. To get around such problems, this Chapter evidences the claim that crosslinguistically three are the kinds of syntactic structure underlying to the NP, all of them instances of Canonical Predication (see immediately above) featuring one and the same Predicate / modifier: a phase (CP / DP) headed by pro.
Two other drawbacks of den Dikken's account highlighted in this Chapter are that according to him movement applies for at least two different reasons (pro-licensing and feature-checking), which threatens the internal coherency of his theory, and that he is sometimes forced to posit syntactic structures violating the antisymmetric framework supposed to derive them. Chapter 5 proposes that these problems evaporate as soons as Moro's *dynamic* antisymmetric account replaces den Dikken's antisymmetric account in describing the syntax of predication of the NP, showing that Moro's POS *alone* replaces *both* of den Dikken's pro-licensing and feature-checking.
Furthermore, Chapter 5 discusses the linguistic phenomenon of Word Formation, given that it plays a key-role in den Dikken (2006)'s syntactic theory of predication. In particular, den Dikken characterizes it along the lines of Baker (1988)'s and Julien (2002)'s Abstract Incorporation, where it boils down to fusion of:

- *two* adjacent atoms of language

- that leaves just *one* atom in situ (a word-like unit)

These two properties of Word Formation make it a phenomenon highly reminiscent of Moro's POS. The present chapter develops this insight and argues that displacement and Word Formation are two facets of one and the same phenomenon, namely the need of neutralizing an illegitimate binary POS leaving just one atom out of it. In this light, displacement and Abstract Incorporation only differ in the way they leave just one atom in situ: the former is *derivational* (*separating* A from B), the latter *representational* (*fusing* A and B). Thus, Chapter 5 puts forward that the so far mysterious phenomenon of Word Formation (actually, representational neutralization of a POS) is to be explained as a strategy (along with X-bar format and displacement) to provide Syntax with Saussurean linearity so that A-P 'understands' it.
This is an extension of Moro's Dynamic Antisymmetry that, it is argued in this Chapter, has far-reaching implications, since it provides not only a definition of word but also a better understanding of the fusive, agglutinating and isolating morphological types, all of them being basically different manifestations of a POS. Accordingly morphology (= Word Formation) is no longer seen as a pathology of language (Aronoff 1998 but cp. also Chomsky 1995), but as a *natural* consequence of the Sintax - A-P interplay.

1.2.5 Chapter 6

Chapter 6 lends additional syntactic arguments (diagnostics) in favor of the findings of the diachronic research on Arabic discussed in Chapters 2, 3, in particular the relationship between Modern Colloquial Varieties and Old Arabic à la Owens, as well as the Language Contact Hypothesis, the Copula Hypothesis, the Classifier Hypothesis, the Relative Clause Hypothesis. Then, it interprets the resulting syntactic scenario within the dynamic antisymmetric theory of predication developed in Chapters 4, 5.

The Chapter takes as its departure point a reduced typology of the Arabic Relative Clauses (RCs), based on the Language Contact Hypothesis. This move makes it possible to single out *just two* RC types: a pure Amorite RC type and a pure Pre-semitic RC type. All the other RC types one way or the other are mixed types deriving from these two RCs.

It is claimed in this Chapter that the Amorite RC type is characterized by development of the demonstrative base *(lla)ḏī* from a relative pronoun (documented in South Arabian) into a Conjunction copula (cp. *and* in *John, who/ and he is the smartest of us all*). This shift is documented in Old Arabic and Nedroma Arabic and points to the non-restrictive nature, at least originally, of this RC type. A two-fold diagnostic tool, based on Arab Grammarians' work, is worked out in order to empirically ground this claim.

Chapter 6 argues that the Pre-semitic RC type too is originally non-restrictive, as shown by another diagnostic tool worked out on the basis of the behavior of non-restrictive RCs in Somali. Moreover, it is argued here, the strict solidarity between the Somali RC and the Pre-semitic one, as found in *Old Arabic*, allows to identify the putative Nominative of this latter - actually, a morpheme *u / w* of unclear function so far - with the conjunction *u / w* 'and' widespread in the *Modern Colloquial Varieties* and in Semitic (cp. also its Somali equivalent *oo*). It ensues that the putative Nominative *u / w* found in the Old Arabic Pre-Semitic RC is a Conjunction copula as much as its Amorite counterpart *(lla)ḏī*. The minimalist analysis of these RC types developed here takes the Conjunction copula to be a dedicated functional projection Conj, adopting the antisymmetric copula analysis of the Somali RC by Frascarelli & Puglielli (2006).

Chapter 6 then turns to discussing a later Pre-semitic RC type, one that exhibits an overt RC marker *l-* (medial *-l-*), still documented in both Old Arabic and in some Modern Colloquial Varieties. It builds diagnostics that leads to assign the medial *-l-* the categorial status of C (complementizer) and moreover it proposes that the prenominal *l-* is to be rethought of as an instance of medial *-l-*, providing an unified account of this morpheme. The argument is built as follows.

Diachronic evidence from Garbini & Durand (1994) and Putzu & Ramat (2001) that the article-construction *l-* + Noun is chronologically preceded by the demonstrative construction *(hā)ḏā(lika)* + *l-* + Noun is discussed. Then, it is shown that the real article in this latter is the base *ḏā*, *not l-*, for two reasons. First, following Brustad (2000) in the modern Colloquial Varieties the base *ḏā* develops from a demonstrative into a (quasi-)article and textual research is made in this Chapter to evidence the same state of affairs for Old Arabic. Second, in Moroccan Arabic etc..., (medial) *-l-* can be combined to the indefinite article *waḥ(e)d* (Edzard, 2006), which rules out its function of (definite) article. Finally, it is shown on the basis of Greenberg (1950) that *ḏ* and *l* give rise to a non-euphonic cluster.

From this standpoint, the article construction *l-* + Noun is reinterpreted as no more than the

demonstrative construction *d̲ā* + *l* Noun where *d̲ā* has evolved into an article and has been dropped in order to avoid the non-euphonic cluster *d̲-l*, so that it has a covert realization.
Implicit in this unified analysis of *l-* as medial *l-* is a copula analysis, since *l-* *intervenes* either between a Noun and its RC, to be interpreted as a Subject and its Predicate following den Dikken (2006) or between a Demonstrative and its Noun, to be interpreted in the same way following Campbell (1996) (and Arab Grammarians). This unified analysis of *l-* is then discussed relative to its consistency with a minimalist/ antisymmetric framework, where elements intervening between demonstratives and nouns are taken to be determiners (D), as per Giusti (2002) and it is shown that the *phonological* and *copular* identity of C *l-* and D *l-* is a reflex of their common nature of phase (on which see e.g. Chomsky 2008).
Last but not least Chapter 6 concentrates on a particular instance of Pre-semitic RC type, ie the Arabic nouns modified by *-n* (*tanwīn*) in the SG and SOUND PL (see end of Section 1.2.6 on BROKEN PL), and develops syntactic diagnostics to evidence that they are Numeral Classifier Constructions, where the noun is followed by a RC made of a Numeral Classifier *-n*, in its turn followed by a (possibly covert) quantifier .
In a similar vein, Arab Grammarians's hypothesis that the Free State and the non-pronominal Construct State are one and the same construction is combined with the Classifier Hypothesis. Under this perspective, the construction at issue is a Possessive Classifier Construction, characterized by the presence of a Possessive Classifier intervening between the Possessee and the Possessor, realized either overtly (Free State) or covertly (non-pronominal Construct State).
The Possessive Classifier at issue is identified with the particle *d̲ū*, which in Old Arabic still means 'Possessee of' rather than 'Possessor of'. The covert realization of *d̲ū* is explained here looking at its form, when definite (another instance of the non-euphonic cluster *d̲-l*) or at its meaning, when indefinite (*-n*, having evolved from Numeral into General Classifier, replaces the Possessive Classifier *d̲ū*, undergoes semantic weakening and is lost, a drift attested in Numeral Classifier Languages).
The several values ascribed to *-n* are straightforwardly accounted for if the Classifier Hypothesis is adopted, since from this vantage point they are the result of a clinal grammaticalization process, starting with the Classifier (Ayoub's Demarcative Hypothesis) and ending up with a C (Owens's Copula Hypothesis: on C as copula see immediately above).
The present Chapter also offers a dynamic antisymmetric account of the Amorite and Pre-semitic RC types, which is able, among others, to derive their postnominal position and the evolution of the Possessive Classifier once again in terms of neutralization of a POS. It also paves the way to an account of Arabic Case in terms of Predicate Inversion, an idea implemented in Chs. 7 and 8.

1.2.6 Chapter 7

This Chapter revives Arab Grammarians' view that the vowels *a, i*, if morphemic, have the *phonological representation awi, uwi* in order to substantiate (and improve) the Agglutination Hypothesis espoused by Petracek (1965), Garbini & Durand (1994) and other Semitists to describe Arabic Case. According to these scholars, *a, i*, when acting as 'case-endings' have:

- an abstract representation *decomposable into distinct morphs* (see e.g. Garbini & Durand 1994, p. 100)

- a semantic value that does not necessarily correspond to Indo-european Case (see e.g. Garbini & Durand 1994, p. 101).

As for the former property, Chapter 7 demonstrates that the morphemic *a, i* occur precisely in the same phonological environment that Arab Grammarians take to be responsible for deriving the phonetic representation *a, i* from the phonological representation *awi, uwi*. As for the latter property, Chapter 7 revives Nyberg (1920)'s copula analysis of Case and uses the syntactic diagnostics elaborated by Kern (2010) to show that (*awi, uwi* phonetically realized as) *a, i* are copulae, and more precisely copulae originally signalling Predicate Inversion, as still documented in a relic construction (Ingham 1994's Transposed NP Strucuture) found in Nadji Arabic, a conservative modern Bedouin Variety.
The Chapter also deals with the syntax of the Hamitic / Pre-semitic Absolutive, the behavior of *hamzah* in Old Arabic, as well as with the alternation of the Semitic demonstrative base *ya- / 'i* in the aim of bringing evidence in favor of identifying the morphemic material *awi, uwi* underlying *a, i* as an agglutinative complex, made up of:

- (the morph *a*, a relic of Absolutive marking the Predicate / Modifier)
- the morph *w*, ie the Conjunction copula (see Section 1.2.5)
- the morph *'i*, a former demonstrative base turned into a copula of Predicate Inversion very much as the Thai particle *thîi* (see Section 1.2.4 and den Dikken & Singhapreecha 2004)

The same arguments corroborate the claim that the phonological representation of the putative case-ending *u* is better understood as *aw*, which minimally yet crucially differ from that of *a, i* in that the latter ones include the copula of Predicate Inversion *'i*, the former does not:

(21) *u* = (a) + w ... - 'i ⇔ Predicate Inversion: NO
a, i = (a) + w ... + 'i ⇔ Predicate Inversion: YES

As such, the Agglutination Hypothesis, if combined with the just mentioned syntactic and phonological arguments, reveals that in Old Arabic (/ Classical Arabic) the putative case-endings *u, a, i*:

- have an internal and fine-grained structure
- are organized into a *binary* opposition *u* vs. *a, i* based on the lack vs. presence of a copula of Predicate Inversion (*'i*) in their phonological representation

Chapter 7 then proceeds to consider the same issue from the perspective of Arab Grammarians and of the Modern Colloquial Varieties. The former ones claim that *u, a, i* are involved into another binary opposition, ie presence (*u*) vs. lack (*a,i*) of the ability to act as Subject. In the latter ones, *u* is opposed to *a,i* in terms of dialectal variation, *u* being found in the Hijazi branch, *a,i* in the Bedouin one (Owens, 2006).
This multi-level parallelism characterizing the binary opposition *u* vs *a, i* is taken here to explain the nature, the birth and the development of Arabic case-system.
In origin Arabic case-system is just a dialectal opposition between the Hijazi and the Bedouin branches (observed in Old Arabic and still documented in the modern peninsular Colloquial Varieties) in terms of lack vs. presence of the copula of Predicate Inversion *'i* in the phonological

representation of their composite copulae *u* vs. *a,i* (cp. (21)).
The copular value of *u, a, i,* entails that they intervene between a Noun/ Subject and a Modifier/ Predicate (underlyingly RC, as discussed in the previous Chapters): Noun + *u, a, i* + Modifier. It ensues that this syntactic structure, when marked for *u* (lack of copula of Predicate Inversion *'i*) is an instance of Canonical Predication, whereas its counterpart marked for *a, i* (presence of copula of Predicate Inversion *'i*) is an instance of Reinstated Word Order (Section 1.2.4).
Such a state of affairs is highly reminiscent of Thai, where *khon kéng* (person-smart: Canonical Predication) is opposed to *khon thîi kéng* (person-copula of Pred. Inv.-smart: Reinstated Word Order) in terms of:

- lack vs. presence of the copula of Predicate Inversion *thîi*
- informational profile: old vs. new information encoded in the head noun

When these two branches fuse into Classical Arabic (see Owens 2006 and references therein), their endings *u* vs. *a,i* compete for the role of Subject, and the Hijazi *u* prevails over the Bedouin *a, i* because the Hijazi *u* is associated with an old informational profile of the head noun, felt as a more suitable candidate for the role of Subject than the new informational profile of the head noun associated with the the Bedouin *a, i* (for old information as typical property of the Subject, see Comrie 1981). Hence the Bedouin *a, i* polarize into non-Subject markers.
Finally, Chapter 7 explores the positive side-effects that a copula analysis of Arabic Case along these lines has on the Classifier Hypothesis, since it demonstrates that the copula analysis of Arabic Case thus characterized makes it possible to realign nouns marked for DU, bound *-t* and, to a certain extent, BROKEN PL to the same Numeral Classifier Construction typical of SG and SOUND PL dealt with in Chapter 6. The dynamic antisymmetric implementation of the copula analysis of Case developed in this Chapter is postponed to Chapter 8.

1.2.7 Chapter 8

This Chapter provides a dynamic antisymmetric account of the Arabic NP based on the findings of the previous chapters, where they have been motivated on *independent* semantic and phonological grounds.
These findings, if couched within an antisymmetric framework, bring out a hierarchy of functional projections highly reminiscent of that posited in previous antisymmetric / cartographic accounts of the Arabic NP (e.g. Cinque 2003, Shlonsky 2004, Ihsane 2003), at least insofar as the left and the right peripheries are concerned.
Nonetheless, the dynamic antisymmetric account proposed in this Chapter departs from the standard antisymmetric / cartographic approach in that the former totally dispenses with the Checking Theory of movement adopted in the latter. Instead, neutralization of a POS is invoked as the real trigger of movement.
In consequence of this, the entire derivational history of the Arabic NP turns out to be a 'chain reaction' of POS:

(22) Base Generation → POS: DP - N →
Merger of AbsP (= GenP) → POS: Abs - NP →
Merger of ConjP (= WP) → POS: Conj - DP →

Merger of DefP (= FP) → POS: Def - DP →
Merger of FocP (= ZP)

More generally, the dynamic antisymmetric account elaborated in this Chapter focuses on the contrast between *representational* and *derivational* neutralization of a POS which, it is argued, is plausibly able to recast in phonological terms and to *replace* the YES / NO parameter setting typical of the Principles and Parameter model, based on semantic assumptions (interplay between interpretable and uninterpretable features).
The same contrast between *representational* and *derivational* neutralization of a POS also derives under a minimalist approach the dialectal opposition between the Hijazi and the Bedouin branch in terms of lack vs. presence of copula of Predicate Inversion *ʾi* (and hence, indirectly, the opposition Subject vs. non-Subject arising from it in the Classical Arabic 'case-system').
Finally, such a contrast provides a principled explanation of why the Inverted Predication structure sometimes remains as such and sometimes turns into the Reinstated Word Order (an issue left unexplained in den Dikken 2006).

Part I

Arabic in its own terms

Chapter 2

The multi-faceted nature of Arabic language

2.1 Preliminary remarks

The topic of the present work necessarily calls for a definition of what is exactly meant by Arabic Noun Phrase.
Nevertheless, the appropriate characterization of this linguistic object in its turn depends on how the Arabic language itself is defined: although this may seem an obvious methodological step, it is far from being so in the case of Arabic (AL-'ARABIYYAH), which in both the academic practice and the popular thinking is usually a cover term designating many varieties at synchronic and diachronic level (Suleiman, 2006).
A further problem in defining the Arabic Noun Phrase lies in the technical terminology, since, intuitively, the traditional terms of Western (= Indo-european) grammar were not originally devised to denote exhaustively the Arabic (= Semitic) linguistic entities (forms, words, constructions and sentences / utterances).
In addition to that, Arabic is one of the few languages in the world (along with Sanskrit) for which a developed linguistic theory exists since the early stages of language, to the effect that the Arabic linguistic entities already have a well established terminology (Owens, 2006, pp. 34-35), which is still in use today in the literature in original or translated form[1]. It goes without saying that a correct understanding of this traditional terminology implies some knowledge of the universe of scientific discourse it refers to.

[1]This dissertation is in line with a recent epistemological trend in the Arabistic studies, pursued by Owens (2006), Hallaq (2005), Leaman (2009), which considers the traditional Arabic Islamic theology, law, linguistics etc... as fully fledged sciences since the beginnings of this society, rather than products of naive speculative thought as traditionally assumed in the past (cp. also, *mutatis mutandis*, Lévi-Strauss 1962):

> "The problems associated with "beginnings" have for long stemmed more from unproven assumptions than from any real historical evidence. Hence, the classic Orientalist creed that the Arabia of the Prophet was a culturally impoverished region, and that when the Arabs built their sophisticated cities, empires and legal systems, they could not have drawn on their own vacuous cultural resources."
> (Hallaq, 2005, pp. 3-4)

For these reasons, this chapter will tackle the issue of the Arabic language and of the Arabic Linguistic Tradition, while the issue of the Arabic Noun Phrase per se is taken up in the next chapter.

2.2 The Arabic language: the state-of-the-art

2.2.1 Synchrony: the diglossia

If one were to ask native speakers which is the major dividing line they would draw inside Arabic, probably the answer one would get would not differ from that of linguists: here and now, Arabic is made up of two varieties - the 'pure' one, the so-called *fuṣḥà* or CLASSICAL ARABIC, and the so-called ARABIC DIALECTS or MODERN COLLOQUIAL VARIETIES (Suleiman, 2006).
Linguists generally call this phenomenon DIGLOSSIA, and describe it as a situation

> "where one linguistic variety has a higher status than another (or others), and in which linguistic functions are partitioned between the two in complementary fashion".
> (Owens, 2001, p. 423)

This is admittedly a sketchy definition of diglossia, based on its first theorization by Marçais (1930) / Ferguson (1959b), which puts the focus on the opposition holding between the two varieties at issue in terms of social prestige / status rather than on their interaction (see also Ferguson 1991 for a reformulation of Ferguson 1959b).
The import of the latter has been emphasized in later sociolinguistic work (see Owens 2001, Larcher 2007, Boussofara-Omar 2006, and Bassiouney 2009 for references and details), where it has been characterized as an intermediate space of varieties separating the two varieties in complementary distribution, to be better understood as two poles encompassing this space. For example, in Egypt a native speaker that wants to convey the meaning *I will say* will have at his disposal not only the two strings *ḥa'ūl* (Egyptian Colloquial Arabic) and *sa'aqūlu* (Classical Arabic), but also some related mixed forms, as reported below and illustrated in Fig. 2.1:

(23) Classical Arabic, Egyptian Colloquial Arabic (Visser, 2005, p. 15)

sa-'a-qūl-u,	sa-'a-qūl,	ḥa-'a-'ūl,	ḥa-'ūl
FUT-I-say-GM	FUT-I-say	FUT-I-say	FUT-I.say

'I will say'

From a sociolinguistic standpoint, the intermediate varieties have been given either a non-variationist or a variationist interpretation, the Arabic language being regarded, respectively, as 'une pluralité d'arabes' or as 'un arabe pluriel': diglossia has thus become a PLURIGLOSSIA or a DIGLOSSIC CONTINUUM (Larcher, 2007, p. 254).
This work, being mainly concerned with the contrast holding between Classical Arabic and the Colloquial varieties, is neutral with respect to the nature of the intermediate varieties and abstracts away from them. In other words, the attention here is on the poles of the Fig. 2.1, rather than on the space between them.
Consequently, the working definition of diglossia given above will be adopted, which opposes

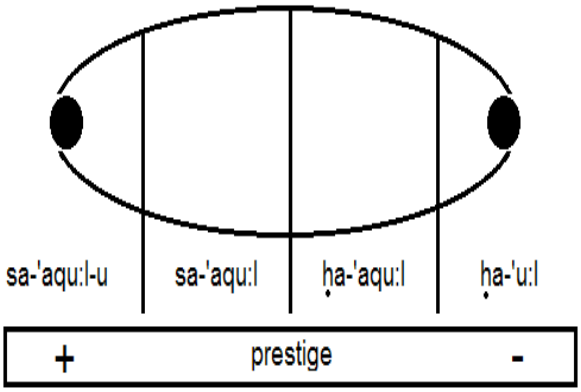


Figure 2.1: Arabic Diglossia

a variety of high prestige (or HIGH VARIETY) to a variety of low prestige (LOW VARIETY). However, when this distinction turns out not to be relevant, the comprehensive term 'Arabic', which translates the original *al-'Arabiyyah* will be used instead.

2.2.1.1 High Variety and Low Variety

In a given linguistic area, the functional domain of the High Variety includes formal occasions, literary and written production, religious moments and rites. In these contexts, purity and codification of the language seem to be privileged against spontaneity of speech, to the effect that the variety used is a fossilized language and people resorting to it are not native speakers. It ensues that in the same area native speakers make use of the Low Variety in all the remaining contexts: at home, when talking with friends, at the marketplace etc...
As far as Arabic is concerned, the linguistic area in which a given Low Variety is spoken is often taken to coincide with a given Arabic country or region (Palva, 2006) - e.g. Egyptian Colloquial Arabic can be conceived to a certain extent as a variety spoken in Egypt. On this basis, Ferguson (1959b) in his classical study *Diglossia* seems to set a one-to-many relationship between one single High Variety and several (national / regional) Low Varieties:

> "DIGLOSSIA is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety... "
> (Ferguson, 1959b, p. 336)

This taxonomical criterion, however, does not do justice of many isoglosses that cut across regional or national boundaries: in particular, Ferguson (1959a) himself, in another classical study (*The Arabic Koine*) argues for the existence of fourteen features that allow to group all the Low Varieties into a single linguistic entity, and oppose this latter as a whole to the High Variety.
Noticeably, these features are very old, since they go back to the 'first centuries of the Muslim era' (Ferguson, 1959a, p. 617), ie about 610-750 AD.

The High Variety, vice versa, is arguably more fragmented than at first sight, since its role is nowadays played not only by Classical Arabic but also by its simplified offshoot, known as MODERN STANDARD ARABIC (MSA), as observed, among many others, by Versteegh (1997a, Ch. 11) and Bassiouney (2009, Ch. 1).
The overall picture that emerges from Ferguson (1959a)'s *The Arabic Koine* is a many-to-one relationship, which is exactly the opposite of what is argued in Ferguson (1959b)'s *Diglossia*, as schematized in Table 2.1. Notice in passing that this scenario has non-trivial diachronic

Table 2.1: Ferguson's view on High and Low Varieties

Ferguson's paper	High Variety	Low Variety
Diglossia	one	many
Arabic Koine	many	one

implications, since it opposes an High Variey which undergoes diachronic change to a Low Variety (idealizedly) immune to it, as the stable nature of its features plausibly suggests.
This aspect will be taken up in Section 2.2.2.3.

2.2.1.2 Classification of the Low Varieties

Leaving aside the just discussed Fergusonian 'super-group' of modern Colloquial Varieties[2], two other major criteria of grouping of Low Varieties have been recognized in the literature: the former distinguishes between Eastern and Western varieties, the latter between Bedouin and Sedentary varieties (Fischer & Jastrow, 1980; Abboud-Haggar, 2006; Palva, 2006).
The former criterion has a geographical nature, since the Eastern group includes all the varieties spoken in what Medieval Arabic Geographers considered to be the East (MASHREK, MAŠRIQ) of the Arab world, ie the area corresponding to the present Egypt, (Arab enclaves in) Israel, Palestinian National Authority, Jordan, Lebanon, Syria, Hatay region of Turkey, Cyprus, Iraq, Saudi Arabia, Kuwait, Bahrein, Qatar, United Emirates, Oman, Yemen.
The Western group covers all the varieties spoken in the so-called MAGHREB or MAĠRIB, ie present-day Libya, Tunisia, Algeria, Morocco, Saharawi Democratic Republic (the 'West' of Medieval Arabic Geographers).
The most famous distinctive isogloss between the two groups is the inflection of the first person singular and plural:

(24) Eastern varieties 1SG 'a(-ktib) 1PL ni(-ktib)
Western varieties 1SG ni(-ktib) 1PL ni(-ktib/ kitb-)u
'I write, we write'

At syntactic level, there are two important distinctive isoglosses that concern the Noun Phrase, reported by Palva (2006, p. 605) and Edzard (2006) and illustrated in Table 2.2[3]: Apart from the two just mentioned isoglosses, the distinction Eastern / Western is not crucial for this work, but it will nevertheless be employed for the sake of clarity, in order to locate one or more varieties in the right geographical context.

[2]See Section 2.2.2.3 for further discussion.

[3]Under the approach developed here the construction: Noun + *-in* reported in the table under the label 'Indefiniteness' encodes this piece of information not in (the Classifier) *-in*, but in an understood Numeral 'one' following it, a phenomenon typical of Numeral-Classifier Constructions (see Greenberg 1990a and Chapter 6).

Table 2.2: Syntax of the Noun Phrase: distinctive isoglosses #1

Variety	Eastern	Western
Free State	low frequency	high frequency
Indefiniteness	bare Noun, Noun + *-in*	wāḥ(ə)d l-

More relevant here is the distinction between Bedouin and Sedentary varieties, which is sociologically-based rather than geographical. It has to be emphasized with Palva (2006, p. 605) that these designations reflect the history of the nomadic communities or of the settled populations (*diachrony*) in the Arab-speaking world, irrespective of their present-day situation (*synchrony*). For example, in the Libyan capital Tripoli, the majority of the population speak a Bedouinized former Sedentary variety (Palva, 2006, p. 610).
The resulting scenario has a good deal of complexity and falls far beyond the scope of this work, which is concerned only with the Bedouin varieties spoken in the Arabian Peninsula. This latter, in fact, might be plausibly regarded as the common place of origin of all the Bedouin varieties, the present-day geographical distributions being the consequence of migrations out of it, as per Rosenhouse (2006, p. 260) and Owens (2006, pp. 2, 3, 161).
Abstracting away from these socio-historical factors, the Bedouin varieties stand out as a coherent and homogeneous group because of some distinctive isoglosses which seem to be morphosyntactic rather than phonological. As Palva puts it[4]:

> "In a classification exclusively based on linguistic contrasts, scarcely any single criterion besides the reflex of **q* distinguishing between the Bedouin-type and the sedentary-type dialects can be found. However, there are a number of prominent typological features, some of which are shared by all sedentary dialects, yet without constituting a contrast with all dialects of Bedouin type, and vice versa, and there are features constituting significant partial contrasts between the two groups".
> (Palva, 2006, p. 605)

The morphosyntactic isoglosses at issue are illustrated in Table 2.3[5], whose notation follows Palva (2006, p. 605) in that (A) indicates the kind of contrast based on features which are 'shared by virtually all the dialects belonging to the group', while (P) indicates the kind of contrast based on features which are shared by 'a substantial part of the respective group'.
Among these isoglosses, of utmost importance for this study are the presence of the interdental fricatives, the expression of indefiniteness and the apparent aversion for the Free State as documented in the Bedouin varieties[6], and especially in that spoken in the Arabian Peninsula and known as Najdi Arabic, which is described in the next Section.

2.2.1.3 Najdi (and Hijazi) Arabic

For the purposes of the current research, the Arabian Peninsula can be divided in four linguistic areas, as per Abboud-Haggar (2003, pp. 104-105, 118-119) (see also Rosenhouse 2006 for a slightly different view):

[4]The starred form in the quoted text indicates a reconstructed phoneme, that is, simplifying, a morpheme which is not directly attested: see Section 2.2.2.2 for the notion of RECONSTRUCTION.
[5]See fn. 3.
[6]See Chapter 6.

Table 2.3: Syntax of the Noun Phrase: distinctive isoglosses #2

Variety	Bedouin	Sedentary
Interdental fricatives	retained (A)	> postdental stops (P)
Indefiniteness	generalized Noun + *-in* (P)	bare Noun, except in formulaic expressions (A)
Gender distinction in plural	retained (P)	not retained in finite verbs and personal pronouns (P)
Verb modifiers in the Imperfect	NO (P)	YES (A)
Productive internal passives	YES (P)	NO (A)
Productive Form IV	YES (P)	NO (P)
Free State	very low frequency	high frequency

(I) North / North-Eastern / Central / Eastern sector: it has as its central core the Desert of Najd, where the Najdi Arabic variety is spoken. The Bedouin varieties spoken in the remaining part of this sector all stem from Najdi Arabic (see Ingham 1982a and Ingham 1997 for details). They can be further divided as follows:

  (a) Bedouin varieties of the Eastern subsector (Kuwait, Qatar, Bahrein, United Arab Emirates, Oman)

  (b) Bedouin varieties of the North-Easternmost subsector (from Southern Iraq to the southern side of Zagros mountains)

  (c) Bedouin varieties of the North-Eastern subsector (simplifying, Syro-Mesopotamian Desert)

(II) North-Western sector (Desert of Negev, of Sinai, Southern Jordan, eastern coastal fringe of the Gulf of Aqaba)

(III) Western sector (regions of Hijaz, Asir, Tihama till the Yemeni coastal fringe facing the Red Sea)

(IV) South / South-Western sector (Yemen and the region of Hadramawt)

This results in the map reproduced in Fig. 2.2[7], which can be simplified and made more handable by merging the North-Western, Western and South / South-Western sectors into a larger group, as Ingham (1994, pp. 1-3, 8-9) proposes on the basis of linguistic and historical considerations (see also Sections 2.2.2.4, 2.2.2.6).

There obtains a major dividing line that distinguishes between Najdi Arabic plus cognate languages (= Bedouin Northern, Central, Eastern varieties) and Hijazi Arabic, which includes Yemeni Arabic (= Western and Southern varieties).

This opposition is traditionally referred to as Eastern vs. Western *(Peninsular)* varieties, as per Ingham (1994, *Ibidem*), Prochazka (1988, Ch. 1) and is not to be confused with the more general opposition between Eastern vs. Western varieties discussed in Section 2.2.1.2:

[7] Abboud-Haggar (2003, pp. 103-104) provides also a comprehensive review of the body of work about these linguistic areas and the interested reader is referred to her volume for detailed references. See also Fischer & Jastrow (1980).

(25) Ar. Peninsula: Eastern var.: Najdi Ar. and cognate varieties

Ar. Peninsula: Western var.: Hijazi Ar., Yemeni Ar.

It should be added for the sake of completeness that in Yemen and in the region of Hadramawt other varieties of Modern Colloquial Arabic are spoken, the so-called South-Arabian dialects, which will be only cursorily dealt with here (see Sections 2.2.2.2, 2.2.2.3 for more details).
We are now in a good position to turn to Najdi Arabic[8].
This latter can be split into three subgroups, following Ingham (1994, pp. 4-5) and Ingham (2006, p. 326) (see also Fig. 2.3):

- Central Najdi, spoken by the tribes living in the Desert of Najd, in which concentrations of wells are enough to permit permanent settlement and agriculture, as well as by the 'Anizah tribes of the Syrian desert
- Northern Najdi, the speech of the people living in the mountainous region of Jabal Shammar ('Shammar's Mountain') and of the Shammar tribes of the Northern Najd and of the Jazirah region, located above the Syro-Mesopotamian Desert
- Mixed Northern Central Najdi, spoken in the Al-Qasim region and by the Dhafir tribe
- Southern Najdi, spoken by the Najran and Ghatan tribes in the Southern Najd as well as by the Al Murrah and Ajman tribes in the (South-)Eastern Najd

In spite of the just described situation of dialectal variation, all these speeches

> "can be classed as Najdi in a general sense in sharing certain definable features setting them off from other dialects of the surrounding area".
> (Ingham, 1994, p. 5)

Moreover, they 'do not differ from each other essentially in syntax' (Ingham, 1994, p. 8).
On these grounds, the present work considers Najdi Arabic as a coherent variety and the term is meant to denote all the speeches listed above, unless otherwise stated.

2.2.2 Diachrony: the stages of the language under investigation

The history of Arabic and its exact position within the Semitic languages are still a matter of debate, because Arabic exhibits common features with more than one branch of Semitic (see Versteegh (1997a, Ch. 2) and the next Section).
In addition to that, an inquiry into this subject seems to be made even more complicated by the material conditions in which the Arabic language developed and by some ideological factors. On one side, the corpus of written documentation preceding the redaction of the Koran (about 650 AD) is relatively scanty, which makes it difficult to draw robust conclusions about the

[8]The most updated works in the field are Ingham (1982a), Ingham (1994), Ingham (1997), this latter being actually a selection of eight articles published by the author over the last thirty years on various Peninsular varieties. In these works Ingham provides also bibliography of the previous literature on the subject and an essential bibliography is also found in Ingham (2006). Among the studies Ingham suggests, particularly outstanding is Prochazka (1988).

diachrony of Arabic. According to Garbini & Durand (1994) this is mainly due to the nomadic nature of the ancient (pre-)Arab society, which by definition relies upon orality, not writing, in the production and in the transmission of knowledge, art, literature.
On the other side, (both Western and Arab) scholars appear to superimpose their sociocultural system of representation on the object of study much more than usually happens in other disciplines, as Owens (2006) points out (see Section 2.2.2.3 for further discussion).
On such a fragmented landscape, the diachronic interpretation of the Arabic language given here, which reproduces (and tries to develop) Owens (2006)'s theory combining it with Garbini & Durand (1994) is only partially compatible with other trends of research and can also turn out to be at odds with them.

2.2.2.1 Arabic and the Semitic family

The Arabic language belongs to the Semitic family, which is traditionally taken to have three branches, as per Moscati et al. (1980, p. 4), Versteegh (1997a, Ch. 2):

- East Semitic (Mesopotamia): Akkadian (and its descendants)
- North-West Semitic (Syro-Lebanese Region): Hebrew, Phoenician, Aramaic (as well as less documented languages like Amorite and Ugaritic)
- South-West Semitic (Arabia and Ethiopia): Arabic, South Arabian, Ethiopian

In this taxonomy, Arabic is regarded as a part of the South-West Semitic, as depicted in Fig. 2.4 (recall from Section 2.2.1 that the comprehensive term Arabic indicates both the Classical and the Colloquial varieties, but in this case, it includes also Epigraphic North Arabic, an early variety of Arabic described in Section 2.2.2.3).
This is what can be called the 'traditional' or 'standard' view, and has as one of its most authoritative representatives Moscati et al. (1980), which develop Brockelmann (1908)'s work (see also Garbini & Durand 1994 and Versteegh 1997a for more details).
Moscati et al. (1980) explicitly state that a grouping of the Semitic languages along these lines is 'usually based on their geographical distribution' and think that the latter 'corresponds tolerably well (though not without certain exceptions) to the distribution of gross linguistic features'.
It has been pointed out (see Versteegh (1997a, Ch. 2) for more detailed discussion and references) that these exceptions are not negligible in the case of Arabic, which, on closer examination, shares important features not only with South-West Semitic languages, as traditionally assumed, but also with North-Semitic ones.
Among them, the so-called BROKEN PLURAL, ie a kind of apophonic plural, and the SUFFIX CONJUGATION, which is basically a subcase of the pattern:

(26) 'Adjective' + Noun

In the Suffix-Conjugation, a pronoun replaces the noun but while in Akkadian the entire construction means something like '*I am ... / you are ... / he is ...*' and so on, in Arabic it evolves into the perfect tense: '*I ...-ed / you ...-ed*' etc... This state of affairs is illustrated in Table 2.4, setting forth data from Versteegh (1997a, *Ibidem*).

Another case in point is the fact that Arabic exhibits at the same time both innovative and conservative features. All the North-Semitic features found in Arabic, in fact, stand out as an

Table 2.4: The ambiguous position of Arabic #1

Branch	South-West	North-West	Arabic
Phonology			
f - *b* opposition	YES$_x$	NO	YES$_x$
Morphology			
Broken Plurals	YES$_x$	NO	YES$_x$
Pronominal stems *-t, h-*	NO	YES$_y$	YES$_y$
Syntax			
Prenominal article	NO	YES$_y$	YES$_y$
Suffix-conj.	NO	YES$_y$	YES$_y$

innovation, while Case (or, in generative terms, morphological Case) represents on the traditional view a good example of a conservative feature, since it is shared exclusively by Classical Arabic and Akkadian.
This situation is summarized in the Table 2.5[9].
These considerations have brought about a crisis in the traditional conception of the Semitic

Table 2.5: The ambiguous position of Arabic #2

Branch	East (= Akkadian)	North-West	Classical Arabic
Morphology			
Pronominal stems *-t, h-*	NO	YES$_y$	YES$_y$
Syntax			
Prenominal article	NO	YES$_y$	YES$_y$
Suffix-conj.	NO	YES$_y$	YES$_y$
Case	YES$_z$	NO	YES$_z$

family, and several alternatives have been devised to justify the 'anomalies' summarized in the Tables 2.4, 2.5, as it will be seen in the next Section.

2.2.2.2 *Stammbaumtheorie* and *Wellentheorie*

The central tenet of (Brockelmann's /) Moscati's model is the STAMMBAUMTHEORIE, originally worked out in the field of Indo-european historical linguistics in the XIX century by the so-called German NEOGRAMMARIANS (Schleicher, Bopp etc... see Lehmann (1993, Ch. 1) for details). This theory holds that a given language B (DAUGHTER LANGUAGE) is created from a given language A (MOTHER LANGUAGE) preceding it chronologically, by a process of gradual modification in time (DIACHRONIC CHANGE).
This genetical relationship is based on some features that B shares with A, which can be detected in spite of the modification that B has undergone, if a rule is found to explain the modification.
Roughly simplifying, a word like English *foot* can be taken to descend from an hypothesized

[9]This work, in the spirit of Owens (2006, pp. 83-84), is not concerned with instances of declension that in Semitic appear as not fully-fledged or problematic, namely Ugaritic, Eblaitic and Epigraphic South Arabian (see Sections 2.2.2.2, 2.2.2.3 for more details on these languages).

word *pod-*, originally found in the Indo-european language, and still documented e.g. in the Ancient Greek word *pod(a)*, via the phonological rule $p \rightarrow f$.
As the example shows, the mother language is often no longer attested and can only be inferred (RECONSTRUCTED) on two empirical grounds.
First, the feature under scrutiny, say, *p* of *pod* is found in a relevant sample of ancient languages, its modified counterpart (e.g. *f*) standing as an isolated and chronologically later case.
Second, the posited rule has a very broad application domain in the language where it occurs, that is it is REGULAR: for example, the rule $p \rightarrow f$ is part and parcel of a robust and generalized phonological process known as Germanic Sound Shift or (Rask's /) Grimm's law.
Under this approach, features of the language B deviant from the rule are accounted for as 'intruders' borrowed from another language that, not belonging to the language B originally, cannot covered by its rules (Lehmann, 1993, p. 287). In Latin, for example, *s* of Indo-european becomes *r* if found between two vowels (contrast *florem* 'flower' with English *blossom*, 'idem') but this rule does not apply in a word like *caseus*, the equivalent of (Indo-european and) English *cheese*, because *caseus* is a word that Latin borrowed from the Sabine language (Palmer, 1954, pp. 37, 69).
Clearly, the just sketched process of feature borrowing, usually called LANGUAGE CONTACT, is seen in the *Stammbaumtheorie* as playing a negligible role in the diachronic change turning the language A into B, a change that primarily rests on B's internal regular rules.
The negligible role of language contact within the *Stammbaumtheorie* impacts also on the characterization of the linguistic features. While features shared by the two languages under comparison (= the mother and the daughter language) form a coherent and detectable linguistic system (SUPERSTRATE), all the remaining features, referred to as SUBSTRATE, only partially represent the linguistic system they are borrowed from, to the effect that *prima facie* they appear as scattered phenomena and are not easily detected by native speakers. Such a state of affairs has led scholars to liken features of the substrate to the residues in chemistry (Lehmann, 1993, p. 26)[10].
In an enlarged perspective, this implies that the shared features (= similarities) between two languages, say the language B and the language C, are typically due to the fact that they are daughter languages inheriting these features from a common mother tongue, e.g. the Indo-european ancestor, already possessed of these latter. Simply put, feature-sharing is feature-retention.
Thus, the above mentioned lexical similarity between English *foot* and Ancient Greek *pod(a)* goes back to Indo-european *pod* and somehow retains it.
This genetical relationship can be diagrammed as a genealogical tree, having the mother language at the top and the most recent daughter languages at the bottom: this is exemplified by the Figure 2.4.
On the other hand, cases in which two languages share a common feature via language contact, like the non-mutation of *s* documented in both Latin and Sabine, are *in principle* ruled out and taken to be exceptional, so that they find no room in the genealogical tree diagram.

[10]Since this kind of research mainly concentrates on the diachronic change and on comparing the features shared by the mother and the daughter language, historical linguistics is often dubbed diachronic or comparative linguistics. Such an equivalence, however, is not totally correct, because this discipline *de facto* involves other techniques, e.g. the so-called INTERNAL RECONSTRUCTION (Lehmann, 1993, Ch. 2) and, which is of more interest here, the DIFFUSIONIST APPROACH (Garbini & Durand, 1994, Ch. 4), as it will become clearer shortly.

It has been seen in the previous Section that the *Stammbaumtheorie*, if applied to the Semitic languages, runs into several problems (cp. also Tables 2.4, 2.5). Following Garbini & Durand (1994, pp. 131-133) this seems to generally hold not only for Brockelmann's / Moscati's model, but also for other approaches based on the same conceptual setting, like Wright (1890), de Lacy O'Leary (1923), Hetzron (1976, 1997), Lipinski (1997) which therefore will not be further reviewed here.

According to Garbini & Durand (1994), these problems evaporate as soon as one rethinks and enhances the role that language contact plays in the diachronic change, integrating the *Stammbaumtheorie* with the so-called WELLENTHEORIE or DIFFUSIONIST THEORY[11].

This is an approach originally formulated in the XIX century by the German scholars Schmidt and Schuchardt (Lehmann, 1993, Ch. 1) as a reaction to the neogrammarians' contention that similarities are an *old* and conservative phenomenon and diachronical change is overwhelmingly based on regular rules (COMPARATIVE METHOD: see fn. 10).

The *wellentheorie* basically holds that a given language C is exposed to the influence of a given language B that radiates from a centre of political or socio-cultural power (e.g. a metropolis, a tribal clan, a court) and spreads from it in concentric waves: accordingly, the centre at issue is designated as CENTRE OF IRRADIATION.

In this case, feature-sharing by B and C turn out to be a *recent* phenomenon (an innovation), stemming from a process of language contact triggered by B on C, by virtue of which C becomes similar (CONVERGES) with B for some features.

In this connection, the partisans of Schmidt's and Schuchardt's approach, the Italian NEOLINGUISTS (as they styled themselves in opposition to neo-grammarians) warn that the features shared by two (or more) Indo-european languages scaled in time are not necessarily retained from a common mother language. According to Neolinguists, the features at issue can also turn out to be an innovation, ie the result of the just described process of convergence (Bartoli, 1925, p. 98). The gist of their position is exemplified by the English word *flower*, which far from being an old feature directly inherited from Indo-european and shared with its Latin counterpart *florem* by virtue of a genetical relationship, is a new feature, (indirectly) borrowed from Latin (through the mediation of French).

This phenomenon is a subcase of convergence which will be called here PSEUDO-FEATURE RETENTION.

Returning to the general case, it is noteworthy that the process of convergence, crucially relying on a political entity, is enclosed within the borders of this latter, ie is restricted to its AREA OF INFLUENCE, and decreases in force as it moves to the borders getting away from the centre, like a stone thrown into a pond of water. This entails that a peripheral language D will converge with the language B influencing it only for a limited amount of innovative features, while preserving archaic features in significant quantity.

This point has been stressed particularly by Italian Neolinguists (Bartoli, 1925, pp. 6-10) which dub it LATERAL AREA PRINCIPLE (NORMA DELLE AREE LATERALI), and derive it from the geographical study of Romance dialects. For example, an innovative feature like the Late Latin word *caballus* 'horse' extends over the centre of irradiation (Italian, language B: *cavallo*) and its neighboring regions (French, language C: *cheval*), while offshoots of the more ancient (Classical Latin) word *equus* 'horse' are found in some varieties spoken in the peripheral area

[11] A further improvement of this approach is Trudgill (1974)'s GRAVITY MODEL.

of the Pyrenees (French-Spanish border), which in this respect resist the Late Latin innovation (languages D)[12].

The *Wellentheorie* has in common with the *Stammbaumtheorie* the view that the features shared by the two languages under comparison (= the centre of irradiation B and the peripheral language D) form a coherent and detectable linguistic system. But the former theory crucially differs from the latter in that it characterizes also all the remaining features of the language D as a coherent and detectable linguistic system (*qua* fully fledged language preexisting irradiation). The two bundles of features, being on the same footing, are no longer defined as superstrate and substrate, but as two PARASTRATE(S) (or, with a synomym, ADSTRATE(S)).

Such a state of affairs leads Palmer (1954) to claim that the peripheral context in which Bartoli (1925)'s Lateral Area Principle applies is no more than a sociolinguistic situation of *bilingualism*:

> "It is such maps which have led linguists to lay down the principle that marginal areas tend to be archaic. But we must bear in mind that in most countries where such geographical studies of language have been carried out we find a congeries of local communities organized in a national state where the dialects are exposed to the influence of an all-pervading standard language radiating from a centre of culture and administration. What it is essential to grasp is that the impingement of the standard language on a dialect speaker is a phenomenon of bilingualism - the dialect speaker at least understands the standard language. If his community is isolated, he has less contact with speakers of the standard language and so is less exposed to innovations proceeding from that source".
> (Palmer, 1954, p. 27)

On these grounds, the peripheral language is to be identified with a *mixed language*, which, by definition, is characterized exactly along Bartoli (1925)'s and Palmer (1954)'s lines: it 'emerged in situations of community bilingualism' and 'arise[s] through continuous socio-ethnic separateness resisting pressure to assimilate' (Matras & Bakker, 2003, pp. 1, 14).

As for Semitic, the choice of introducing elements of the *Wellentheorie* into the the *Stammbaumtheorie* to solve the interpretive problems summarized in the Tables 2.4, 2.5 is justified by a well-known historical fact, namely the presence of nomadic dynamics in the regions of Mesopotamia, Syro-Lebanese region, Arabian Peninsula and Ethiopia. Nomadism, in fact, provides an ideal context for language contact (Garbini & Durand, 1994, pp. 131-135) - notice in passing that this scenario is also recognized in Moscati et al. (1980, pp. 3-4).

In particular, archaelogical and epigraphical evidence plausibly demonstrate that around 3000 BC Semitic people are located in the Syrian plateau (indicated as II in the Fig. 2.5) and that nomadism (and hence language contact among Semitic languages) originally takes place in this region (Garbini & Durand, 1994, *Ibidem* and p. 27).

In this light, the Bedouin / Arab nomadism documented in the Arabian Peninsula and discussed in the Sections 2.2.1.2, 2.2.1.3 is the late (and most famous) reflex of a life style inherited from ancient times and enhanced by camel taming (see below).

Two other areas of subsequent occupation that will be dealt with here are the bordering regions

[12]Similar offshoots are also found in the French Massif Central and in the Alps, areas detached from the centre of irradiation not in terms of geographical distance, but in terms of geographical morphology (mountain regions). They are often referred to as SPRACHINSEL(S) and covered by Bartoli (1925, pp. 3-6)'s ISOLATED AREA PRINCIPLE (NORMA DELL'AREA PIÙ ISOLATA). Thanks to Irene Tinti for helpful discussion on these issues.

of Mesopotamia (on the east and south east) and of Canaan (on the west and south west).
This latter term covers the coastal region spanning from Gaza to Tell Sukas / Latakia, with its immediate hinterland (Garbini & Durand, 1994, p. 36): it correspons to present-day Palestinian National Authority, Israel, Lebanon (Fig. 2.6), and to the portion of Syria indicated as I in the Fig. 2.5.

Accordingly, the Syrian plateau as well as its neighborhoods set the main scene for the linguistic history of Semitic people. The chronological limits are from 3000 BC to 1200 BC, the conventional date in which ancestors of Bedouins / Arabs began their settlement of the Arabian Peninsula (Garbini & Durand, 1994, p. 142).
Following Garbini & Durand (1994, p. 137), since the III millennium BC a major centre of irradiation in this macroregion has been a portion of the Syrian plateau, namely the northern area located between Mesopotamia and present-day Syria and occupied by Amorites. This is a Semitic ethnic group living in (semi-)nomadic tribes.
Hence, the Syrian plateau and its neighborhoods are seen by Garbini & Durand (1994, p. 139) as the area of influence of this centre[13].
The language spoken by Amorites in 3000 BC is of the utmost importance in Garbini & Durand (1994)'s proposal, since it one way or the other is an important 'ingredient' (*qua* substrate or parastrate) of *all* the North-West and South-West Semitic languages emerged from 2000 BC onwards.
On the other hand, East Semitic is independent from Amorite and this fact is not accidental, because its earlier manifestations, ie Akkadian and plausibly Eblaite (3000 BC) constitute a centre of irradiation on their own (Garbini & Durand, 1994, pp. 148-152), which by definition resist convergence.
The two centres of irradiation at issue differ in an interesting respect. While Akkadian and Eblaite represent the oldest linguistic typology within the Semitic family[14], the Amorite speech is more prone than them to diachronic change and evolution (Garbini & Durand, 1994, pp. 34, 137), probably because of its nomadic and hence non-written / non-codified nature.
The same factor results in the lack of original written documentation, so that Amorite language is known only indirectly through Akkadian etc... sources in its essential characters (Garbini & Durand, 1994, 33-34). Among them, of particular relevance here are the innovative suffix conjugation and the innovative pronominal / demonstrative system (*d̲-, h-* etc...), as already seen in Section 2.2.2.1 and in the Tables 2.4, 2.5 as well as discussed in Garbini & Durand (1994, pp. 107, 114, 136-137).

[13]Cp. Garbini & Durand (1994, p. 136): 'il seminomadismo amorreo, che vediamo nel III millennio a.C. particolarmente vivace nella zona intermedia tra le città della Mesopotamia e quelle della Siria'; Garbini & Durand (1994, p. 139): 'un'area linguistica, localizzabile nella Siria settentrionale non urbana nella quale...sono nate e si sono progressivamente sviluppate delle innovazioni che le popolazioni nomadi, in seno alle quali queste si manifestavano, hanno di volta in volta trasferito nelle zone di sedentarizzazione'; Garbini & Durand (1994, p. 144): 'il gruppo innovatore amorreo nella regione sub-urbana della Siria'.

[14]Cp. Garbini & Durand (1994, p. 136): 'L'accadico e l'eblaita rappresentano pertanto una tipologia linguistica arcaica'.
Notice also that by the term 'proto-Semitic' (Garbini & Durand, 1994, p. 144) mean exactly this archaic and *historically documented* linguistic typology, not Neo-grammarians' reconstructed mother-tongue or *Ursprache*. As Garbini & Durand (1994, p. 136) put it: 'Le lingue semitiche non sono dunque manifestazioni parallele nate dalla diversificazione di una più antica realtà linguistica (il «protosemitico»)... quelle [= Akkadian and Eblaite] scritte nei centri urbani, alle quali spetta, storicamente parlando, il nome di «proto-semitiche»' .

In the same period in the Syrian plateau there are (indirectly) attested three other non-written languages, which Garbini & Durand (1994) dub Cananean, South Palestinian, Pre-Akkadian (Cananeo, Sud-palestinese, Pre-accadico, respectively).
These languages, on the whole, resemble in one respect. They are all instances of Semitic chronologically prior to Akkadian and Eblaite, dating back to 4000 BC or even before and characterized by features also found in the Hamitic languages. Among them (Garbini & Durand, 1994, pp. 143) the phonological opposition *f*-*b*, which is better understood as a former opposition between approximants $\phi - \beta$, and the broken (= apophonic) plurals (cp. also Section 2.2.2.1 and Tables 2.4, 2.5). Accordingly, Garbini and Durand call them 'pre-Semitic languages'.
At this point one may wonder which is the exact relationship holding between Semitic and Hamitic under the diffusionist approach. Garbini & Durand (1994, pp. 171-174) bring arguments in favor of the hypothesis that Hamitic languages arise as a consequence of language contact between pre-Semitic languages and African languages, so that a given feature shared by both a Hamitic language X and a Semitic language Y, is no more than the same pre-Semitic (= very archaic Semitic) feature got in the past by the Hamitic language X via language contact, and retained over the time by the Semitic language Y.
Nevertheless, they leave open the possibility, advocated in the traditional *Stammbaumtheorie* (Moscati et al., 1980, pp. 16-17), that pre-Semitic and Hamitic descend from a common ancestor (HAMITO-SEMITIC or AFROASIATIC). Interestingly, they also point out (Garbini & Durand, 1994, pp. 173) that both the hypotheses yield the same result: namely, that Hamitic languages preserve pre-Semitic relics[15]. This remark will play a key-role in this work.
Returning to the pre-Semitic languages, South-Palestinian and Pre-Akkadian are similar relative to their location, because both of them are found at the *periphery* of the Amorite area of influence: the former in the Sinai region (present-day Egyptian-Israeli border), the latter in the Western coastal region of the present-day Persian Gulf. On the other hand, Cananean was spoken in Canaan, as the name itself suggests (Garbini & Durand, 1994, pp. 37, 76, 143-148, 171-174)[16].
With all this in place, the *Wellentheorie* predicts that:

(I) each of North-West and South-West Semitic languages is no more than a wave of Amorite language that has impacted on one of the pre-Semitic languages encountered on its path

(II) retention of the pre-Semitic features is a function of the Amorite wave's distance in space and in time from its centre of irradiation - more precisely, an inverse proportion. The closer

[15] 'Ora, tanto nell'ottica diffusionista quanto in quella genealogica, è sin dal periodo proto-semitico (inteso nel senso spiegato pocanzi) che si attuano i processi evolutivi sulla base dei quali in Africa si formeranno i diversi gruppi camitici. La conclusione da trarre è che le lingue camitiche, siano esse rami staccatisi anticamente dal tronco afroasiatico comune ed evolutesi autonomamente o lingue africane semitizzate, racchiudono falde di fossili linguistici semitici'.

[16] Cp. Garbini & Durand (1994, p. 144): 'Ma lontano dalle città, e senza alcun rapporto, se non forse sporadico, con queste [= cities of the Syrian plateau], continuavano a vivere comunità agricole a livello neolitico e gruppi di cacciatori ancora a livello epipaleolitico [= speakers of South-palestinian and pre-Akkadian]; genti il cui patrimonio linguistico non doveva differire molto da quello in possesso delle popolazioni che sullo scorcio del IV millennio a.C. si trovarono, forse loro malgrado, coinvolte nel processo di urbanizzazione [= speakers of Cananean]...parlate che, a vari livelli cronologici, dovevano rappresentare diverse forme di pre-semitico (o, in termini comparativistici tradizionali, di camito-semitico)'.

the Amorite wave, the weaker the pre-Semitic feature-retention (substrate), the more distant the Amorite wave, the stronger the pre-Semitic feature-retention (parastrate).

Garbini & Durand (1994, pp. 140-141) in their work explicitly state (I), setting the following equivalences:

- Ugaritic = Amorite + Cananean (or perhaps a different pre-Semitic language): present-day Latakia (Syria), 2000-1400 BC
- South Arabian = Amorite + pre-Akkadian: Mesopotamia, 2000-1400 BC
- Cananaic (Hebrew, Phoenician) = Amorite + Cananean: present-day Southern Lebanon and Israel, 1200 BC
- Arabic = Amorite + South Palestinian, 1200 BC

This list does not include Aramaic, a language documented in written form from 1000 BC onwards, since it is a pure instance of Amorite and more precisely the endpoint of the just mentioned process of innovation typical of this latter, a process that has not interrupted from 2000 to 1000 BC (Garbini & Durand, 1994, pp. 138, 149)[17].
Nor is Ethiopian included in the list, because it is regarded in its essence as a further development of the 'mixture' Amorite plus pre-Akkadian (Garbini & Durand, 1994, pp. 149-150).
Finally, Hebrew and Phoenician are classified as two varieties of the same language (Cananaic / Cananaico, to be kept distinct from its pre-Semitic 'ingredient' Cananean / Cananeo). In spite of the clearly different political and religious identity, in fact, they share the same linguistic background (Garbini & Durand, 1994, pp. 36-38).
As for (II), although Garbini & Durand (1994) do not grasp it, they nevertheless recognize the more and more increasing resistance of pre-Semitic to converge with Amorite. They in fact qualify pre-Semitic as a weak substrate in the case of Ugaritic, but as a strong substrate in the case of Cananaic and even as a parastrate when it is the 'ingredient' of a language located in a peripheral area: Arabic. South Arabian deserves a separate treatment, because the wave irradiating from the Amorite centre overlaps with that of the Akkadian centre, as shown immediately below.
In consequence of this, Ugaritic and Cananaic can be likened to two daughter languages of Amorite (Garbini & Durand, 1994, p. 149), while Arabic cannot, being rather a (slightly) mixed language - 'una lingua leggermente mista' (Garbini & Durand, 1994, *Ibidem*).
The overall scenario is illustrated (in an idealized way) in the Fig. 2.7 and in the Fig. 2.8.
This different reaction (= convergence) of the pre-Semitic languages can be measured relative to the above mentioned features *f-b* and broken plural, which survive in the Arabic parastrate (and in South Arabian), but not in the Ugaritic and Cananaic substrate. In this light, the traditional 'South-West Semitic' features of Arabic turn out to be features of the pre-Semitic parastrate (Garbini & Durand, 1994, p. 143).

[17]Cp. Garbini & Durand (1994, p. 149): 'lingua fortemente innovatrice, l'amorreo non poteva cristallizzarsi nella forma che aveva nel 2000 a.C., sì che, sia per evoluzione interna sia per il fatto che le parlate delle tribù non sedentarizzate differivano ovviamente da quelle delle tribù fissatesi in Siria e Palestina, dopo un millennio l'amorreo appare trasformato in aramaico. E' assai difficile dire se l'amorreo debba essere considerato come una lingua «figlia» dell'amorreo o non piuttosto come la stessa «madre» amorrea trasformata, come dopo un bagno nella miracolosa fonte della giovinezza.'

Another feature of Arabic that, according to Garbini & Durand (1994, p. 145), belongs to pre-Semitic is what is traditionally regarded as 'North-West Semitic' , namely the prenominal article *l-*. The term 'article', however, is used for the sake of commodity, given that, following Garbini & Durand (1994, pp. 102-103) the morpheme *l-* developed the function of article only *subsequent to* the the pre-Semitic period.

The typological and comparative considerations recently made by Testen (1998) and Pat-el (2009) (see Section 3.3.3 for details) prompt to extend the pre-Semitic (/ Hamitic) hypothesis also to the prenominal article found in Cananaic .

Notice also that South Arabian and Ethiopic, which do not inherit from pre-Semitic the expected prenominal article, are a motivated exception, because, as Garbini & Durand (1994, p. 146) observe[18], these languages are also exposed to the Akkadian centre of irradiation and in Akkadian, one may add, the prenominal article does not occur.

As already seen above, the remaining traditional 'North-West Semitic' features of Arabic, ie the pronominal stems *-t, h-* and the suffix-conjugation are reinterpreted under the diffusionist approach as features of the Amorite centre of irradiation.

Once again, the exceptional behavior of South Arabian and Ethiopic in this respect (neither the stem *-t* nor the suffix-conjugation occur) is motivated under the same approach as a consequence of interference by the Akkadian centre of irradiation. Their features can be therefore explained either as a convergence with Akkadian, or, as Garbini & Durand (1994, pp. 106, 147, 150) suggest, as a manifestation of pre-Semitic, which in this case has not been impacted by the Amorite wave (although non-trivial Amorite influence is detectable in other sectors of the grammar, e.g. the consonantal system)[19].

To sum up, similarities between Arabic and South Arabian are either due to language contact with Amorite or inherited from their pre-Semitic 'ancestors' (where they are also related with the Hamitic family).

The picture that emerges from the discussion so far is that the problems raised by the *Stammbaumtheorie* are circumvented if the interpretive tool of language contact is adopted. This seems to hold valid also for a feature which has been left undiscussed so far, namely Case.

An account of this feature in terms of *Wellentheorie* has not been not explored by Garbini & Durand (1994), but has been suggested by Owens (2006), which hypothesizes a situation of language contact between Arabic and Akkadian.

The main problem in this account lies in the fact that only very rarely is Case borrowed from another language (Matras, 2009)[20]. Yet, Owens's idea can be retained if one reinterprets Case

[18] As they put it: 'Il predominio assoluto della cultura scritta cuneiforme, sumerica e accadica, non ha lasciato spazio a quelle varietà linguistiche di cui è più ricca la regione siro-palestinese' . For example, while in almost all Semitic languages the meaning of 'to write' is conveyed by the root *k t b*, in South Arabian there occurs the root *š ṭ r*, which is only found in Akkadian (Garbini & Durand, 1994, p. 146).

[19] Cp. Garbini & Durand (1994, p. 150): 'Quando, molto più tardi, saranno documentate queste lingue « meridionali» (sudarabico, geez, ed etiopico meridionale) appariranno evidenti, specie nelle prime due, le tracce di una presenza amorrea (confermata, come si è già detto, dalla più antica onomastica sudarabica), ma queste sono tali da consentire, al massimo, una definizione di lingua «amorreizzata» [...]. La nostra totale ignoranza della situazione linguistica a sud della Mesopotamia nel II millennio a.C. non ci consente di valutare il ruolo svolto dall'amorreo in questa zona: possiamo dire soltanto che esso non fu secondario.[...], mentre è sicuro il contatto produttivo tra le parlate locali (poco fa si è accennato a una specie di pre-accadico[...]) e l'amorreo'.

[20] Thanks to Bettina Migge for pointing me out this aspect of grammatical borrowing and for helpful discussion on this point.

as a connective signalling the relationship between a noun and its modifier and intervening between them, as suggested in Chapter 1 (see also the next Chapters), because the connective is a grammatical category that may naturally undergo borrowing (Matras, 2009).
Given that a connective with these distributional and structural properties is documented in Hamitic (e.g. Somali) and given that under the diffusionist approach feature-sharing between Hamitic and Semitic points to a pre-Semitic origin, the actual source of borrowing of Arabic 'Case' is not Akkadian 'Case' but a pre-Semitic connective found in its South-palestinian parastrate.
The Akkadian Case, in this light, is explained exactly in the same way: Akkadian, in fact, because of its archaicity still shares some pre-Semitic features with the peripheral language referred to as pre-Akkadian (Garbini & Durand, 1994, p. 147)[21]. The most striking example in this regard is the Akkadian prefix-conjugation, which is paralleled not only by Ethiopic, but also, in Hamitic, by Ancient Egyptian and Lybico-Berber (Garbini & Durand, 1994, p. 120). Notice that this similarity is recognized also in more traditional accounts, such as Moscati et al. (1980, p. 133).
To recap, the ambiguous position of Arabic has been deeply reconsidered, the features previously ascribed to North-West or South-West Semitic being replaced by Amorite or Pre-Semitic features. The bundle of features at issue is summarized in Table 2.6, to be contrasted with Tables 2.4, 2.5 (notice that the Table 2.6 abstracts away from Akkadian and South Arabian / South-West Semitic, because the presentation of their features would involve adding to the picture a third variable, namely the Akkadian wave, which is of less interest here).
Last but not least, it should be noticed that what has been referred to as Amorite and pre-

Table 2.6: The ambiguous position of Arabic: revised

Branch	Pre-Semitic	Amorite	Arabic	Cananaic ('North-West')
Phonology				
f - *b* opposition	YES_x	NO_y	YES_x	NO_y
Morphology				
Broken Plurals	YES_x	NO_y	YES_x	NO_y
Pronominal stems *-t, h-*	NO	YES_y	YES_y	YES_y
Syntax				
'Prenominal article'	YES_x	NO	YES_x	YES_x
Suffix-conj.	NO	YES_y	YES_y	YES_y
'Case'	YES_x	NO_y	YES_x	NO_y

Semitic languages in this Section for expository purposes, are better seen as a set of speeches, as per Garbini & Durand (1994, pp. 144, 149). Once again this is plausibly due to the nomadic style of life and to lack of writing.
Therefore, if one considers that Arabic as we know it since its earlier manifestations in the Arabian Peninsula in IV AD (see below) is subject to the same material and socio-cultural con-

[21]They define in fact pre-Akkadian as 'una fase linguistica parallela a quella dalla quale nei centri urbani si formò l'accadico, una specie di pre-accadico,' adding that 'L'esistenza, in queste lingue [= Ethiopian and South Arabian], di due forme verbali a prefissi, anche se con funzioni diverse da quelle dell'accadico, costituisce un argomento notevole a favore di questa ipotesi'.

ditions, it comes with no surprise that it is not a compact language but a set of speeches. Such a state of affairs was recorded by Arab Grammarians and Lexicographers (cp. Sections 2.2.2.4, 2.2.2.6) and cannot be easily dismissed.
Hence, the linguistic situation of Arabic is perfectly natural for a diachronic approach along the lines of Garbini & Durand (1994), but seems to raise a further problem for the Semitic *Stammbaumtheorie*, as it will become clear in the next Section.

2.2.2.3 The 'logical matrix' and Owens's criticism

Before dealing with the situation of linguistic variation within the Arabian Peninsula and its accounts in the modern linguistic theory, something else should be said to bridge the gap between the proto-history of Arabic and South Arabian in the Syrian plateau, outlined in the previous Section, and their historical vicissitudes, which took place in the Arabian Peninsula.
Arabic and South Arabian have originally a different location and a different history in the Syrian plateau, so that their migration towards the Arabian Peninsula is the result of a multifaceted migration process, began in 1200 BC for the speakers of Arabic (see previous Section), and in 2000 BC for the speakers of South Arabian (Garbini & Durand, 1994, p. 146).
The process is complete in 900 BC, when the first epigraphical attestations of their languages appear and (heavily simplifying) the former group of migrants have occupied all the Arabian peninsula except Yemen and Hadramawt, which is inhabited by the latter one.
The variety of Arabic recorded in the epigraphic material, called North Arabic, differs from later stages of Arabic in that it displays the prenominal article *(h)n-* instead of *l-* and its last attestations date to 400-500 AD. In these centuries, North Arabic is supplanted by a stage of Arabic generally (but not correctly) designated as 'Classical (North) Arabic' in both traditional Moscati et al. (1980, p. 14) and less traditional works (Garbini & Durand, 1994, pp. 64-65).
About in the same period, 'Classical Arabic' supplants also epigraphic South Arabian (as well as a language closely related to it and known as Hymiaritic[22]).
At this point, the picture proposed by the *Stammbaumtheorie* gets blurred.
On one side, it argues that 'Classical Arabic' is a compact language, characterized by the presence of Case, a feature shared with Akkadian but not with the modern Colloquial varieties (contrast the Table 2.5, where the feature Case is ascribed to '*Classical* Arabic', with the Table 2.4, where all the remaining features are ascribed to Arabic *as a whole*). This behavior is made sense of assuming that Case of 'Classical Arabic' is an old feature, since it is found also in Akkadian, and that the modern Colloquial varieties ('dialects') are descendants of 'Classical Arabic' which have lost it (Moscati et al., 1980, pp. 14, 96).
More generally, all the Semitic languages other than 'Classical Arabic' and Akkadian, being caseless, would have dropped case-endings.
On the other side, the *Stammbaumtheorie* reports two facts at odds with this scenario, namely[23]:

- dialectal variation is documented not only *subsequent* but also *prior* to Classical Arabic (Moscati et al., 1980, p. 14)

[22]Other poorly documented languages related with 'Classical Arabic' are the language of the so-called Protosinaitic Inscriptions and Nabatean, which will be not dealt with here. In particular, in line with Garbini & Durand (1994, p. 51) and Mascitelli (2006, pp. 66-67), Nabatean is regarded as an 'arabized' Aramaic, rather than as (a variety of) Arabic.
[23]Thanks to Antonella Ghersetti for pointing me out these two important aspects.

- modern (and caseless) Colloquial varieties are not necessarily descendants of Classical Arabic, since they still preserve archaic elements (Moscati et al., 1980, pp. 5, 14)

The partisans of the standard view take great pains to harmonize these two data with the theory (see Section 2.2.2.6 and Table 2.7 for more details and examples).
The main challenge that archaicity of dialectal variation and of caseless Colloquial varieties present is that they project the present-day situation back - at least - to the beginnings of the history of 'Classical Arabic' in the Arabian Peninsula (IV c. AD), so that caseless 'Colloquial varieties' might plausibly have coexisted with 'Classical Arabic' (Owens, 2006, p. 74).
It goes without saying that the resulting scenario makes the argument of archaicity of Case (to be better understood as a connective) not cogent. After all, as seen in the previous Section, feature-sharing between Akkadian and 'Classical Arabic' is not necessarily to be taken as retention of an old feature but, according to Owens, might plausibly be the product of language contact. Thus, feature-retention is only a deceiving effect of the fact that 'Classical Arabic' has been in contact with the same linguistic feature of Akkadian.
This is the phenomenon described by the Italian Neolinguists and indicated in the Section 2.2.2.2 as Pseudo-feature retention:

(27) $\begin{array}{llll} \text{English} = \text{Anglo-Saxon} & + \text{French}_{Latin:flower} & \leftrightarrow \text{Indo-eur.}_{Latin:flower} \\ \text{Arabic} = \text{Amorite} & + \text{South-Pal.}_{Pre-Semitic:Case} & \leftrightarrow \text{Akk.}_{Pre-Semitic:Case} \end{array}$

But this is tantamount to saying that Case of 'Classical Arabic' cannot be longer taken to be an archaic feature, to be opposed to an allegedly 'younger' lack of Case of the other Semitic languages (modern Colloquial Arabic included). In Owens (2006)'s terms, there is no 'Old-New dichotomy' between Case and lack of Case, and any attempt to establish it is an untested postulate (or, in Owens's terms, a LOGICAL MATRIX).
Unluckily, it is precisely this stipulative dichotomy that, according to Owens, has characterized Arabist studies over *all* of its history:

> "The problem with Arabic, as will be seen, is that the historical relation among varieties of Arabic was basically fixed by the mid-nineteenth century, and thereafter impervious, if not oblivious, to reinterpretation, up to the present day".
> (Owens, 2006, p. 35)

Given that the just discussed archaicity of dialectal variation and of caseless varieties is a problem not easy to circumvent, the issue of the (alleged) diachronic development from 'Classical Arabic' to the modern Colloquial varieties from IV c. AD onwards is still a matter of debate - not to speak of the diachronic development from the (alleged) Semitic *Ursprache* to 'Classical Arabic' as discussed at length in the previous Section.
With the notable exception of Moscati et al. (1980, p. 14), which signficantly refrain from entering into the debate, the main points of divergence among scholars seem to mainly revolve around three still unsettled questions:

- What triggered the loss of case endings?
- Which are the stages of language Classical Arabic passed through before developing into the modern Colloquial Varieties?
- Which is the link between the modern Colloquial varieties and the old ones?

It is clear that the first question is orthogonal to (Owens's) diffusionist approach, where absence vs. presence of Case boils down to absence vs. presence of borrowing of this feature via language contact.

Nonetheless, for the sake of completeness, it is worth mentioning Fück (1950)'s positions on this issue[24], since the German scholar was the first to offer an explanation of Arabic Case in terms of language contact (although within the Old-New dichotomy framework). According to him, loss of case endings is the result of massive grammatical simplification introduced by non-native speakers of Arabic when they were conquered by non-Arabs and began learning Classical Arabic.

Fück's theory had two spin-offs: Ferguson (1959a)'s Koine Hypothesis, according to which Classical Arabic underwent a simplification in the military camps about 700 AD, which heralded in the modern Colloquial Varieties, and Versteegh (1987, 2004)'s Pidginization Hypothesis, which reframes Fück's ideas in the light of the findings of the second language research (pidgin languages, second language learning etc...).

The second question as well has no theoretical status under (Owens's) diffusionist approach since no development of this kind takes place. However, the reader interested in the genetic approach may wish to consider Fleischer (1854)'s proposal, which splits Arabic into Old, Middle and New Arabic.

Blau (1981) and Fischer (1982, 2006b) give a different implementation to Fleischer's ideas - Fischer, in particular, revives in part classical work in the field by Nöldeke (1897), while Blau adds into the picture elements of language contact and second language research.

Fleischer (1854)'s work was also largely influential to Arabic dialectology: cp. Fischer & Jastrow (1980) and, more recently, Abboud-Haggar (2003).

Turning to the third question, it is a central concern for (Owens's) diffusionist approach, because it calls into the picture the two main problems (archaicity of dialectal variation and of caseless varieties) that cause the breakdown of all the theories based on the Old-New dichotomy, revealing this latter as a logical matrix, and test the validity of any approach that is meant to replace them.

Notice that, in a diachronic perspective, bringing into comparison the present-day Colloquial varieties with those described by the Arabic Linguistic Tradition from 800 AD onwards, and probably going back to 300 AD (see next Section) does not necessarily imply an analysis of the intermediate stages of language occurring between them (say, the nature of Colloquial varieties in the X or in the XVIII c. AD). Owens (2006), in fact, does not take into account this time span and concentrates on some features of the old Colloquial varieties that have been kept identical throughout the centuries till the present-day Colloquial varieties, like the so-called *'imālah*[25] (Ch. 4), some phenomena of epenthesis and of pause (Ch. 7), some pronominal forms (Ch. 8). Furthermore, the nature of the data make it difficult to compare modern South Arabian Colloquial Varieties with their old counterparts, only scarcely documented by the Arab Grammarians and Lexicographers (Rabin, 1951, Ch. 4), and the same difficulties are encountered when comparing modern Colloquial Varieties with Epigraphic South Arabian (Garbini & Durand, 1994,

[24] This work has been also translated in French by Claude Denizeau in 1955, under the title *Arabiya. Recherches sur l'histoire de la langue et du style arabe.*

[25] This term indicates in the Arabic Linguistic Tradition the change of $\bar{a}$ to an $\bar{e}$-like value: e.g. *ḥimi̲er* instead of *ḥimā̲r* 'donkey' in Classical Arabic and *si̲emiḥ* instead of *sā̲maḥ* 'he forgave' in modern Colloquial Lybian Arabic. See Owens (2006, Ch. 4) for further discussion.

pp. 62-63). Accordingly Owens (2006) narrows down the scope of his research, cursorily mentioning these varieties only two times (pp. 74, 275).
Finally, Owens's perspective disallows evolution of 'Classical Arabic' into the modern Colloquial Varieties, not evolution of 'Classical Arabic' in and by itself and in fact Owens (2006) sees in this latter a development passing through three stages. Roughly simplifying, they are a Koranic, a Medieval and a Modern stage, three notions that will be made more precise in the three next Sections.
The overall picture that emerges is that Owens (2006), while dismissing the theoretical apparatus of the Old-New dichotomy, revamps some of its contents (although this move is not explicitly admitted by Owens himself). Actually, he regards

- 'Classical Arabic' as an evolving entity, along the lines of Fleischer (1854) and Fischer (1982, 2006b)
- old Colloquial varieties as an 'atemporal' entity (Owens, 2006, p. 114), insofar as features like lack of Case, *'imālah*, epenthesis-phenomena and some pronominal forms are taken into account, and opposes these varieties as a whole to the multiform 'Classical Arabic', a scenario highly reminiscent of that proposed by Ferguson (1959a)'s Koine Hypothesis

Owens's, in fact, differs from Ferguson's only in that (Owens, 2006, p. 114) the group of Colloquial varieties does not descend from 'Classical Arabic' but is directly linked to proto-Semitic - being irrelevant here if this term is to be understood as the early historical manifestation of Semitic, à la Garbini & Durand (1994) or as a reconstructed mother language, à la Owens (2006)[26].
The difference between the logical matrix / Old-New dichotomy and Owens (2006)'s approach is illustrated in the Table 2.1) above, here repeated as (28):

(28) Logical matrix HIGH VARIETY: one LOW VARIETY: many
Owens's approach HIGH VARIETY: many LOW VARIETY: one

For the reasons mentioned above, Owens (2006)'s approach will be followed here and in particular its methodological choices:

- no analysis of the intermediate stages between old and modern Colloquial varieties
- little attention to old and modern South Arabian Colloquial varieties[27]
- a fine-grained view of 'Classical Arabic'

This has non-trivial implications for an appropriate definition of the notion of 'Arabic' in diachronic and historical terms (ie after its emergence in the Arabian Peninsula in IV c. AD), as it will be seen in the next Section.

[26]Notice also that in Ferguson (1959a), very much as in Owens (2006), the bulk of the discussion is devoted to Colloquial varieties other than South Arabian ones. For example Ferguson states in connection with the relative marker *'illi*: 'it is significant that throughout the non-Arabian dialects [= Colloquial varieties spoken outside the Arabian Peninsula] the only forms found are those which may be derived from a presumed **'illi*' (Ferguson, 1959a, p. 630).
This passage clearly shows that Ferguson is aware that his generalizations do not always apply to modern South Arabian Colloquial varieties - which in fact exhibit a relative marker *di* or the alike: see e.g. Pennacchietti (1968, pp. 44-47).

[27]Unless otherwise stated, these latter are not included in the term 'Colloquial varieties'.

Before concluding, however, a clarification is in order: Owens (2006)'s diachronic characterization of Arabic makes use of terms already existing in the Arabistic and Semitistic literature, but with a new meaning.
Thus these terms, when used throughout this work, are not to be understood in the familiar (e.g. Fleischer's, Fischer's, Ferguson's etc...) sense.

2.2.2.4 Old Arabic

This term designates a stage of language documented from 300 AD to 800 AD (Owens, 2006, p. 6). The *terminus a quo* is the so-called Nemara inscription, located southeast of Damascus in Syria and dated 328 AD, while the *terminus ad quem* is represented by the corpus of forms, words, constructions and sentences / utterances reported in three written sources[28]:

- the Koran (*al-Qur'ān*) as transmitted in the *Qirā'āt*, ie the Koranic readings (or: 'renditions') which albeit written down in the X c. AD by Ibn Mujāhid (d. 324 / 936), are associated with an eponymous reader that lived in the VIII c. AD.
- the grammatical treatise *al-Kitāb* ('the Book') by Sībawayhi (d. 177 / 793)
- the Koranic commentary *Ma'ānī l-Qur'ān* by al-Farrā' (d. 207 / 822)

The linguistic materials found in other texts of the same period, namely:

- Pre- and Early Islamic poetry
- the so-called *'ādāb* literature, ie courtly literature, *belles lettres*

are not considered by Owens (2006, pp. 7-8, 39) as attestations of Old Arabic because they bear marks of later manipulation and systematization, which assimilated them to Classical Arabic[29]. It is worth mentioning in this connection that Sībawayhi and al-Farrā' when reporting a given linguistic fact of Old Arabic in their treatises, may associate with it some lines of Pre- and Early Islamic poetry, but these latter seem to have no bearing on the authenticity of the linguistic facts under discussion:

> "The verses...used by the philologists themselves, are meant to serve as examples, not as evidence for the existence of the phenomenon they illustrate. For that reason in most cases it does not matter much whether they are genuine or not."
> (Rabin, 1951, p. 15)

Finally, it should be added that late scholarly sources other than *al-Qur'ān* of the *Qirā'āt*, *al-Kitāb*, *Ma'ānī l-Qur'ān* can be taken to be reliable sources of Old Arabic provided they merely repeat and take extracts from one of the three just mentioned works. This holds for both Late Grammarians (Owens, 2006, pp. 7, 88, 198-199) and Lexicographers (Owens, 2006, pp. 63-64) (see Section 2.3.5 for more details).
On the whole, the language described in these texts is mainly characterized by a considerable

[28] See Section 2.3.2 for the data elicitation and gathering practices resorted to in these sources.

[29] Owens (2006, pp. 7-8, 46-47) rules out the so-called Middle Arabic texts, dating back to VII c. AD, as sources of Old Arabic, because he regards Middle Arabic as a style, not as an historical stage of language. The reader should be aware, however, that Middle Arabic, as Owens (2006, pp. 6-7) admits, is a difficult matter to discern and therefore other alternatives are viable as well: see e.g. Blau (2002).

amount of internal variation, so that a given form, word, construction or sentence / utterance is commutable with other variants keeping its core-meaning unaltered, as exemplified by (30), (31) below.

As Owens (2006, pp. 11-12) remarks is extremely difficult - in spite of Rabin (1951)'s impressive work[30] - to take out from the just mentioned variation a complete phonological and morphosyntactic account of discrete dialects. Nonetheless, the available data make it possible to flesh out two branches of Old Arabic, the Tamimi and the Hijazi ones, as per Owens (2006, pp. 119-120, 136 and cp. also pp. 60, 98-99, 207-208, 263)[31].

These two branches still survive today, respectively, in the Eastern and Western Peninsular varieties already discussed in Section 2.2.1.3 and schematized in (25): this parallelism will be dealt with in the next Section (see in particular the Tables 2.7 and 2.8).

Such a state of affairs opposes Old Arabic to Classical Arabic, where the increasing standardization and codification at all descriptive level either excludes altogether the variants at issue or, being primarily interested in general principles and rules, treats them separately and as peripheral topics (Owens, 2006, p. 88).

As far as the Noun Phrase is concerned, the most telling example of this opposition is the construction featuring what is traditionally taken to be the Arabic Case:

(29) Noun + 'Case' + Modifier

Owens (2006, pp. 94-95), in fact, reports instances of 'Case' alternation in Old Arabic, which actually do not involve Case semantics at all, and according to Sībawayhi, are only a matter of different dialectal usage[32].

Thus, according to Owens, in Old Arabic sentences like (30), (31), documented among others in Sībawayhi's *al-Kitāb*, the two modification-structures featuring the 'Case'-vowels (*-ā-*, *-ū-*, respectively) are to be reinterpreted as caseless forms displaying an intervening vowel, whose value varies according to the dialectal branch. In fact, the construction (31) is dubbed by Arab Grammarians *mā hijāzī*, that is 'the (negation) *mā* as used in the Hijaz'[33]:

(30) Old Arabic, non-Hijazi (Tamimi?) branch (Owens, 2006, p. 94)

mā 'abd-u-llāh-i 'ax-ū-ka
not slave-NOM-God-GEN brother-NOM-you

'Abdallah is not your brother'

(31) Old Arabic, Hijazi branch (Owens, 2006, p. 94)

mā 'abd-u-llāh-i 'ax-ā-ka
not slave-NOM-God-GEN brother-ACC-you

[30]Since Rabin (1951) few has been added to the corpus of Old Arabic data and in fact one of the most up-to-date and clear overviews of Old Arabic (Al-Sharkawi, 2006) is largely based on this work.

[31]This distinction was already known to Arabic Linguistic Tradition, and was revived and corroborated in modern Arabist studies by Vollers (1906), then further developed by Rabin (1951, 1955) and Owens (2006). Arab Grammarians' view slightly differ from that of modern scholars in that the dichotomy Tamimi vs. Hijazi is based more on an ethnographic (tribal) criterion than geographical one. Nonetheless, Rabin (1951, Ch. 1) shows that an in-depth investigation of the former provides useful information about the latter and allows to identify the Tamimi (Bedouin) group (at least) with the Najdi area, whereas the Hijazi group corresponds to a good extent to the Hijazi area.

[32]This hypothesis is taken to its extreme consequences in this dissertation: see Chapter 6.

[33]A third variant, attested in the *Qirā'āt*, is the zero vowel, ie a generalized pausal form: see Owens (2006, Ch. 4).

'Abdallah is not your brother'

Notice that this explanation holds independently of whether the vowel at issue is epenthetic, or morphemic, two possibilities put forward by Owens (see e.g. Section 2.2.2.2).
Later on, in codifying Classical Arabic, the Arab Grammarians will tend to look for the general principles and rules governing the distribution of 'Case'-endings, taking great pains to assign them their appropriate distributional contexts either on a formal or on a semantic basis (Guillaume, 1998; Bohas et al., 1990). Consequently, constructions like (31) will be deemed exceptional by them (while in Chapter 6 the hypothesis will be entertained that this 'paradigm shift' has to do with a process of resemantization of the 'case'-endings, from discourse-related markers to semantic role markers).

2.2.2.5 Classical Arabic

Based on the foregoing definition of Old Arabic, Classical Arabic turns out to be 'the endpoint of a development within the complex of varieties of Old Arabic' (Owens, 2006, p. 5). That is, it inherits its features from both the Tamimi and the Hijazi branches and can be characterized as a mixture of them, as per Owens (2006), which revives Rabin (1951, Ch. 1) and Vollers (1906), if not even Arab Grammarians' view.
Following Owens (2006, pp. 114-116), this development is basically the transition from a stage in which nouns do not obligatorily bear 'Case' (whatever its nature) to a stage in which they do.
Furthermore, this stage of language undergoes codification. At least in this respect, Owens (2006)'s characterization of Classical Arabic is in line with more traditional positions (see Fischer 2006b and references therein), which ascribe this process to political and religious reasons (see Section 2.3 below).
Insisting on the aspect of codification has the advantage of providing the chronological limits of the linguistic stage 'Classical Arabic' which is conventionally identified with Arab Grammarians' period of codifying activity, ie from 800 AD to 1500 AD, as per Bohas et al. (1990, Ch. 1) and Fischer (2006b).
This chronological delimitation leaves post-Classical sources largely untouched here, as by and large they perpetuate the Classical scholastic mode until the modern times. Following Fischer (2006b, p. 189), in fact, the last representative of this post-Classical tradition is the Lebanese Faris al-Shidyaq (d. 1887), who flourished in the first half of 1800.
In the same period, Modern Standard Arabic emerges in the context of the movement of socio-cultural (and technological) renewal known as NAHḌAH (Versteegh, 1997a, Ch. 11), but it is basically a simplified version of Classical Arabic (see, among many others, Versteegh (1997a, *Ibidem*) and Owens (2006, p. 5)), which is of low interest for the the present purposes.
In short, Classical Arabic is no more than a 'portion' of Old Arabic which later on gets endowed with obligatory 'Case' and standardization, as illustrated in Fig. 2.9.
Beyond Case, another feature that distinguishes Old Arabic from Classical Arabic is the syntax of verbal agreement (Owens, 2006, p. 111-112). While the former allows full agreement between the verb and its subject in the word-order VSO, the latter does not in the same context:

(32) Old / Classical Arabic(Rabin, 1951, p. 168)

'akala-nī l-barāġīṯ
ate-me the-lice

'The lice ate me up'

(33) Old Arabic (Rabin, 1951, p. 168)

'akal-ū-nī l-barāġīṯ
ate-PL-me the-lice

'The lice ate me up'

(34) Classical Arabic (Rabin, 1951, p. 168)

*'akal-ū-nī l-barāġīṯ
ate-PL-me the-lice

'The lice ate me up'

(33) is the most famous example of the syntactic feature 'full agreement' that distinguishes Old from Classical Arabic and therefore this latter is also indicated in the literature as the (VSO-)type *'akal-ū-nī l-barāġīṯ*.

2.2.2.6 Old Arabic and modern Colloquial Arabic

On diachronic level, characterizing Old Arabic as caseless suggests a direct and strict link between this latter and the modern Arabic Colloquial varieties, which behave in the same way. This link

> "creates a geometric figure out of what in comparative Semitics has too often been defined as a one-dimensional structure beginning with Akkadian and ending with Classical Arabic [...]. Adding the modern Arabic dialects [= Colloquial varieties] creates a geometric structure with at least two dimensions in the sense that developments and / or archaisms from proto-Semitic may move directly from the proto-language to the modern dialects, bypassing Classical Arabic completely."
> (Owens, 2006, p. 114)

Fig. 2.9 illustrates this state of affairs.

Owens (2006, *Ibidem*), capitalizing on work by Zaborski (1995) and Retsoe (1995) also points out that, apart from Case, there at least three other features shared by Old Arabic and the modern Colloquials, which reinforce the link between them.

First, some pronominal forms, which can be exemplified for the purposes of this dissertation by the lexical items *'āna*, 'I', (V +)*-k(a)*, '(V +) you (Subject)', reported by Rabin (1951, pp. 51, 151) and Al-Sharkawi (2006).

Second, following Blanc (1970) the so-called PSEUDO-DUAL, ie a nominal form which either designates two entities (like a proper dual) or a number that range from two to two dozen (unlike than a proper dual) (see Section 3.2.1 for details and examples).

Finally, some agreement phenomena, like that exemplified as (33) in the previous Section, and here designated as SPECIFICITY-DRIVEN AGREEMENT: see Section 3.2.2 as well as Hashemi (2006b, and references therein) for a more accurate description.

The picture that emerges from the discussion so far can be summarized in the Table 2.7.

Table 2.7: Old Arabic and Modern Peninsular Varieties: common features

	Old Arabic	Modern Colloquial Varieties	Classical Arabic
'Case'	NO	NO	YES
Pronouns 'ānа, (V +) -k(a)...	YES	YES	NO
Pseudo-dual	YES	YES	NO
Specificity-driven Agreement	YES	YES	NO

The parallelism is even stronger if, among modern Colloquial varieties, the Peninsular ones are considered. They exhibit not only the same distribution of the just mentioned features into two main dialectal groups, but also additional features in common with Old Arabic. The features at stake are the elision of the high vowel in an open, non-final unstressed syllable, as well as the change $a > i$, or HIGH VOWEL DELETION and LOW VOWEL RAISING, following Abboud (1979)'s terminology.

In the literature, their manifestation has been reported for Old Arabic by Rabin (1951, pp. 97, 101) and Owens (2006, p. 60) independently from the manifestation of the same phenomena in the modern Colloquial Peninsular Varieties, described by Ingham (1994, pp. 27, 33), Prochazka (1988, pp. 19, 27-28) and, more extensively, in Abboud (1979).

In greater detail, High Vowel Deletion and Low Vowel Raising are realized in the modern Eastern Peninsular branch as well as in the ancient Tamimi branch, and oppose these latter to the Western Peninsular branch and to the Hijazi branch, respectively, as shown in Table 2.8.

Data on Old Arabic are drawn from Rabin (1951, p. 101), Owens (2006, p. 60), data on Eastern Peninsular Arabic from Prochazka (1988, pp. 27-28)[34].

To sum up, Owens (2006)'s claim that modern Colloquial Arabic are at least as old as Old Arabic rests upon the features listed in the Table 2.7.

It has been furthermore shown here that this holds valid in particular for the Peninsular varieties: plausibly, the distinctive features laid out in the Table 2.8 constitute evidence enough to maintain that the modern groupings Eastern vs. Western Peninsular Varieties correspond to the old groupings Tamimi vs. Hijazi, respectively.

2.3 The Arabic linguistic tradition: an overview

It has been seen in the previous Sections that the transition that took place in 800 AD from a stage of linguistic variation (Old Arabic) to a stage of codification (Classical Arabic) is strictly

[34]The application domain of Low Vowel Raising is the open syllable, nonetheless in the (old) Tamimi branch it seems to unexpectedly apply also in the context of a close syllable, e.g. *ḥijj*, 'pilgrimage' vs. the Hijazi form *ḥajj*. But application of Low Vowel Raising in this case is loose, because, as Rabin (1951, p. 101) remarks, in the Tamimi branch there occur regular forms like *'axḏ*, 'taking' that sharply contrast with *'ixḏ*, attested in other varieties of Old Arabic. In the just mentioned examples, one cannot exclude the interaction of other factors, like a guttural or velar environment, as hinted at in Prochazka (1988, p. 19).

Similar considerations hold for (modern) Western Peninsular forms like *ričib*, 'he rode ', manifesting the Low Vowel Raising that would be normally expected in their Eastern Peninsular counterparts. This anomalous behavior can be explained in terms of vowel-harmony, as per Prochazka (1988, *Ibidem*).

Noticeably, (exceptional) application of Low Vowel Raising is loose also in this case, because in the Western Peninsular branch forms like the just mentioned *ričib* (6 out of 13 documented variants) coexist with forms like *račib* (7 out of 13 documented variants), where, as expected, Low Vowel Raising does not apply.

Table 2.8: Old and Modern Peninsular Varieties: distinctive isoglosses

Sector	Stage	Branch	LOW VOWEL RAISING	HIGH VOWEL DELETION	Gloss
			NO	NO	
West	Old (Arabic)	Hijazi	na͟hy	‘a͟lima	rain-pool, he knew
	Modern (Coll. Var.)	Western Pen. Var.	ka͟tab	ka͟tabat	he, she wrote
			YES	YES	
East	Old (Arabic)	Tamimi	ni͟hy	‘a͟lma	rain-pool, he knew
	Modern (Coll. Var.)	Eastern Pen. Var.	ki͟tab	k͟tibat	he, she wrote

entangled with the development of the Arabic linguistic tradition.

This comes with no surprise, given that the Arab Grammarians were the ‘material executors’ of the operation of LANGUAGE CODIFICATION that the political (and religious) establishment planned in this century as a reaction to the breakdown of its hegemony (Fischer, 2006b, pp. 397, 402).

At that time, in fact, non-Arab sedentary people (*mawālin*) represented the new rising class taking over the power and they had to be ‘arabized’. Simply put, they were required to be taught and master Arabic, the official language and the language of administration. Hence, Arabic language experts were called to flesh out the grammar and define the lexicon of such a *langue d’empire*, in order to promote linguistic (and cultural) homologation.

But this is not the whole story. The establishment involved grammarians and lexicographers also in a policy of LANGUAGE PRESERVATION which chronologically precedes that of language codification and coincides with a deep sociocultural transformation passed through by the Arabic-Islamic society after Muḥammad’s death (11 / 632). The Bedouin elites begin in fact a territorial expansion that lead them to sedentarize and this results in a progressive lost of connection with the nomadic traditional language (= Old Arabic) and culture. Perhaps, the most striking example is Koran, which was considered under threat of corruption.

The establishment reacts with assigning grammarians and lexicographers the task of transmitting the linguistic heritage (Guillaume, 2006, pp. 117-178), also by what today would be defined as special field research, ie by traveling extensively throughout Central Arabia in order to receive instruction from the Bedouin themselves (Seidensticker, 2006, p. 35). On the other hand, codification plays little or no role in this spate of philological activity, as evidenced by the fact that its terminology is largely taxonomic: e.g. *ḥarakah*, ‘vowel’, *ṣifah*, ‘modifier’, *waqf*, ‘pause’ (Carter, 2006, p. 183).

Interestingly, these researchers on the field report that the variety of Old Arabic spoken by their Bedouin informants, which they define KALĀM AL-‘ARAB is understandable and akin to their Old Arabic (in change?) (Fischer, 2006b, pp. 399-400). This state of affairs which cannot be taken to hold for the later scholars working on codification after 800 AD. Whatever the sociolinguistic situation of the varieties of Arabic they spoke, which is not under discussion here, it seems safe to maintain that the just mentioned sedentarization process made such varieties sensibly distant from and hence not mutually understandable with *kalām al-‘arab*.

Thus, the codification process of 800 AD represents a fundamental dividing line in time not only

for the history of Arabic (see Section 2.2.2.5), but also for the history of its linguistic tradition, be it grammatical or lexicographical.
For the purposes of this dissertation, the grammarians and lexicographers active before the end of the eight century, like 'Abū l-'Aswad al-Du'alī (d. 69 / 688), the founder of the Arabic grammar, and Sībawayhi (d. 177 / 793), are referred to here as Early Grammarians and Lexicographers, while those flourishing after this date will be called Late Grammarians and Lexicographers.
Before concluding, it is worth noting that the (so to speak) nationalist connotation of the linguistic research generally precluded any interest in languages other than Old Arabic, although Arab Grammarians and Lexicographers themselves were sometimes native speakers of Persian (e.g. Sībawayhi) or Greek (e.g. 'Ibn Jinnī (d. 392 / 1002)).

2.3.1 Object, aims and problems

The just discussed change of sociocultural and sociolinguistic conditions - from sedentarization / language preservation to non-Arab escalation / language codification - seems to have had a significant impact on the Arabic Linguistic Tradition, producing a paradigm shift in it.
It is generally agreed (Bohas et al., 1990; Guillaume, 2006; Carter, 2006), in fact, that before 800 AD the discipline does not set forth its theoretical principles explicitly and systematically, drawing its legitimacy less from them than from the personal authority and experience on the field of its leading practitioners (see the end of Section 2.3.2). Scholars are primarily concerned with the transmission of the corpus of linguistic data as they are found in the usage (SAMĀ') and in the communicative context (Bohas et al., 1990, Ch. 2), so that language seems to be seen as a social product (Carter, 2006, p. 185), in a way highly reminiscent of Saussurean structuralism.
This explains why in grammar the basic unit of analysis is the utterance (KALĀM) associated with the speaker's intention (MURĀD) (Bohas et al., 1990, Ch. 2).
Bohas et al. (1990), Guillaume (2006), Carter (2006) also point out that that on the other hand after 800 AD Late Grammarians and Lexicographers keep transmitting the received corpus, but *de facto*

(I) refuse to increase this latter on the basis of the personal authority and experience on the field

(II) restrict their attention to a close set of standard data and are mainly preoccupied with organizing them within a general and abstract framework, which should be able to bring out their patterns of regularity (QIYĀS).

Accordingly, in grammar the basic unit of analysis is the sentence (JUMLAH) and its structural constituency, regardless of its context, to the effect that a JUMLAH is further characterized either as 'ISMIYYAH, ie as having a

(35) Topic + Comment

constituency or as FI'LIYYAH, ie as having a

(36) Verb + Subject

constituency, while as far as it is known the definitions *kalām 'ismī*, *kalām 'fi'lī* are never used in the technical literature (Bohas et al., 1990, p. 57).

In this period, therefore, language seems to be a sum of structures that can be decomposed without resorting to their semantics (Fleisch, 1961, vol. I, p. 46), an approach that can be likened to American structuralism[35].

The make up of the new scientific discourse makes quite evident the link between the change of scientific attitude and the sociocultural transformation of 800 AD, since Late scholars' tendency to organize the data according to universal theoretical principles (II), having much in common with logically-based foundational assumptions, is closely related with the renewed interest in Hellenic thought shown by the emergent Non-Arab class (*mawālin*). After Sībawayhi's death (d. 177 / 793), Greek philosophy and especially logic was circulating in the theological milieux (cp. the so-called Mu'tazilite or 'rationalist' school) and have already entered grammatical works (Carter, 2006, p. 185).

This process of Hellenization reached its full accomplishment around 930 AD in both religious and linguistic thought, as evidenced, respectively, by the works of the theologians al-Juba'ī father (d. 303 / 915) and son (d. 321 / 933) (Peters, 1976, pp. 17-19) as well as by 'Ibn al-Sarrāj's (d. 316 / 928) grammatical treatise *'Uṣūl al-Naḥw* (Bohas et al., 1990, pp. 10-11).

By the same token, Late scholars' tendency to close the set of sources (I), rejecting the personal authority of an expert in the field was influenced according to Carter (2006, p. 185) by the legal practice known as 'the closure of the door of endeavor' (*'insidād bāb al-'ijtihād*), fully achieved in 900 AD, which proposed similar guidelines for jurisprudence.

Remarkably, Hallaq (2005) has recently shown that this legal trend develops between 700 and 800 AD and is deeply rooted in the already discussed phenomenon of Bedouin sedentarization begun after Muḥammad's death (11 / 632). In the garrison towns of that period, in fact, law was still run by the (pre-Islamic and) Bedouin institution of the arbitrator (*ḥākim*), a man deemed to be in possession of experience and authority, who made large use of *'ijtihād (al-ra'y)*, ie of his discretional opinion and personal reasoning (literally, of his 'endeavor (of opinion)'), to solve the disputes among tribesmen that fell under his jurisdiction.

But under the gradual relocation of Arabs from the garrison towns to the cities, began in 710 AD, they less and less functioned in their capacity, their authority and personal reasoning being progressively limited in favor of a fixed set of written legal sources (among them, Koran).

This is explicitly stated around 800 AD by the jurist Muḥammad Ibn 'Idrīs al-Šāfi'ī (d. 204 / 820), which holds that *'ijtihād (al-ra'y)* is not applicable in the face of conclusive textual evidence (Hallaq, 1993, p. 592). This paves the way to the just mentioned 'closure of the door of endeavor' in law. Thus, in chronological terms, the flourishing and decline of the (personal authority of) arbitrators finds a striking parallel in the flourishing and decline of the (personal

[35] One may further speculate that Late Grammarians' and Lexicographers' approach is justified by and arises as a consequence of the same (poor) material conditions observed in the American structuralism. This is the fact that the language studied by scholars (either *kalām al-'Arab* or Native American) is by and large unintelligible and to be deciphered (see above). Given that in these circumstances scholars cannot easily access word meanings to understand and attain the language, the only alternative left for them is to make use of formal (= morphosyntactic) tools, like distributional analysis, immediate constituency analysis and so on.

Benveniste (1954) was the first to put forward this epistemological explanation for emergence of American structuralism (see also Lepschy 1966, pp. 95-96, 151-152 for further discussion). What is temptatively suggested here is extending Benveniste's considerations to the Late Arabic linguistic tradition.

authority of) Early Grammarians and Lexicographers (I).
For the sake of practical orientation, the former paradigm (Early Arabic Grammar) can be roughly compared to modern descriptive and / or field-linguistics, the latter (Late Arabic Grammar) to modern prescriptive and / or armchair-linguistics, although such generalizing labels can be sometimes misleading, since nowadays these approaches turn out to overlap (e.g. the generative paradigm is usually regarded as armchair-linguistics but not as prescriptive linguistics).
Arab Grammarians and Lexicographers themselves were aware of this paradigm shift, since they, according to the circumstances, compared themselves to either the rationalist theologians, the so-called *mutakallimūna* (Bohas et al., 1990, p. 27)) or the jurists (the *mutafaqqihūna*: cp. Carter 2006, p. 189).
Moreover, they neatly divided practitioners of the linguistic discipline into two main trends, the so-called 'school of Baṣrah' and 'school of Kūfah', the former focussing on the usage (*samā'*), the latter on the structural patterns (*qiyās*).
This distinction was formulated in particular by the winning paradigm, ie 'Baṣrans' / Late Grammarians, in a way such that they could deny their historical links with the intellectual milieu of their predecessors, the 'Kūfans' / Early Grammarians, and even rejected the latters' data and interpretations if disturbing the new paradigm (Bohas et al., 1990, pp. 6-8, 68-70).
But, as pointed out by Owens (1990, Ch. 8), Bohas et al. (1990, Ch. 1), Guillaume (2006) this is just the version told by the winners, which obscures the fact that the Arabic linguistic tradition as a whole was a self-consistent science, with a shared object, aims, problems and methods. It has already been mentioned, in fact, that *kalām al-'Arab* is the object of research of both Early and Late scholars, and that both of them devoted themselves in different ways to the transmission of this language - a task often referred to in their texts by the Arabic equivalent NAQL AL-LUĠAH (Bohas et al., 1990, pp. 17-20).
It should be added that recent work by Owens (1990, Ch. 8), Baalbaki (2004, Ch. 13), Owens (2006, Ch. 3), Baalbaki (2008) has convincingly shown that organizing the raw linguistic data into a coherent system is an aim together with their accurate transmission not only for the Late Arabic Linguistic Tradition, but also for their predecessors. Already al-Xalīl (175 / 791), Sībawayhi's teacher, shares with the Late Grammarians and Lexicographers the view that the usage (*samā'*) of *kalām al-'Arab* is characterized by order and harmony (*naẓm*) and that linguistic description should capture this state of affairs (Versteegh, 1997b, p. 55).
This attitude is echoed in the Arab Grammarians' and Lexicographers' equivalent of 'data systematization', 'data classification', ie 'AXḎ AL-LUĠAH (Bohas et al., 1990, pp. 20-22). This expression, taken literally, simply means 'taking of the linguistic data', plausibly because the language, on their view, is already organized in itself so that arrangement of the data automatically follows from its proper understanding and is not imposed externally by the scholar.
Rather, what Early and Late Arabic Linguistic Tradition differ in is the epistemological status assigned to *'axḏ al-luġah*. While for the latter it is a central theoretical concern, to be justified in abstract and logical terms, in the former it is not necessarily given such a theoretical status Carter (2006, p. 185), because it is regarded as an ancillary tool designed to help the authors in the presentation of a massive set of data, from the less to the more complex (Owens, 2006, Ch. 3).
It has been in fact noticed (Owens, 2006, *Ibidem*) that in grammar, Sībawayhi's exposition

proceeds from the simpler to the more complex syntactic structure, and by the same token that in lexicography (Seidensticker, 2006), the simpler phonotactic structure, that is the consonantal root, is taken as the starting point of a more and more complex series, in which the consonantal root is progressively augmented by some affixes. This organizational principle of the entries, which Seidensticker (2006) dubs 'ROOT - CLASSIFICATORY SYSTEM' stems from the first Arabic (semasiological) dictionary, al-Xalīl's *Kitāb al-'Ayn* (Seidensticker, 2006) and is reinterpreted here as the same and more general systematizing force at work in Sībawayhi's grammatical treatise[36].

This epistemological parallel is not surprising, at least for two reasons. From an historical viewpoint al-Xalīl was Sībawayhi's teacher and it is widely agreed that Sībawayhi's linguistic thinking is heavily indebted to al-Xalīl - see e.g. Haywood (1965, p. 22) and Carter (2004, pp. 29-33).

In addition to that, the simpler forms of Sībawayhi's morphosyntactic analysis (phrases or sentences) and the simpler forms of the Xalīlian phonotactic analysis (consonantal roots) are assigned the same term *'aṣl* (to be dealt with in Section 2.3.3).

To summarize, this Section has clarified the object (*kalām al-'Arab*), the aim (primarily, language preservation)[37] and the two central problems (transmission of the corpus or *naql al-luġah*, and systematization of the corpus or *'axḏ al-luġah*) of the science which is usually referred to as Arabic Linguistic Tradition. As for its methods, it is felt here that they are largely driven by the two just mentioned problems and hence they will be treated in the two following Sections as related either to the former or the latter.

[36]Western scholars (Haywood, 1965; Versteegh, 1997b; Seidensticker, 2006) usually do not consider the just described technique of proceeding from the less to the more complex ('root - classificatory system') as the backbone of al-Xalīl's dictionary, and rather take it to be the so-called PERMUTATIVE SYSTEM. This consists of grouping under the same entry all the roots sharing the same consonants, regarding them as permutations of a single 'super-root', a technique judged by Western scholars as cumbersome and making little sense.

Under the approach advocated here, however, the permutative system turns out to be a necessary precondition for the root - classificatory system. The basic idea is that the simpler form needed by the root - classificatory system as its starting point, ie the consonantal root, can be singled out from a sequence of word consonants only by permutating these latter.

This is tantamount to saying that al-Xalīl's permutative system is in its essence a (reliable) way to extract / to get the root consonants - a way that is cumbersome and not self-evident unless the perspective of the language games is adopted.

This phenomenon, in fact, is well documented in the modern Arabic Colloquial varieties, where it precisely consists of a permutation that affects only the root consonants, leaving other affixal material unaltered and is a perfectly natural pastime for the native speakers (see Walter 2003 for details and references).

In this light, al-Xalīl's permutative system turns out to be a language game, which he probably borrowed from some Bedouin ludic linguistic behavior and adapted for his purposes of systematization of linguistic facts.

Language game thus becomes a diagnostics (permutative system) to retrieve in the lexicon the simpler forms on phonosyntactic level (consonantal roots), which allows to take them as a departure point to list the remaining complex forms (consonantal root plus affixes), according to an increasing order of complexity (root - classificatory system). It would therefore ensue that the permutative system is largely instrumental to the root - classificatory one.

[37]Recall that the second aim, ie language codification is later in time, and has to do with Classical Arabic, not with Old Arabic / *kalām al-'Arab*.

2.3.2 The transmission of the corpus: *muṭṭarid* and *šādd*

Arab Grammarians and Lexicographers resort to a binary opposition to classify the transmitted data, namely the contrast between MUṬṬARID and ŠĀDD (Guillaume, 2006, p. 177).
These concepts, following Bohas et al. (1990, pp. 20-21) and Guillaume (2006) can be taken to be the equivalent of the Western concepts of 'regular' and 'irregular' respectively (see the beginning of Section 2.2.2.2).
The distinctive property of this opposition is the presence vs. the lack of what Arab Grammarians call QIYĀS, a term that, among the possible translations, will be here rendered into English as 'measurability'[38].
This is the ability for a new piece of linguistic data A, or *primum comparandum*, to be 'measured' (= compared) against an already known piece of linguistic data B, or *secundum comparatum*, in order to test the matching degree between them. If the matching is incomplete, the new piece of linguistic data is regarded as unable to be measured (or: as not possessed of *qiyās*), ie as irregular (*šādd*).
If, on the other hand, the matching is total, the new piece of linguistic data is regarded as 'measurable' (possessed of *qiyās*), ie as regular (*muṭṭarid*).
For example, the phrase *lam yakun*, 'he was not', if 'measured against (*yuqāsu 'alà*)' the phrase *lam yaqum*, 'he did not stand', is felt as *muṭṭarid* because both of them match relative to the pattern *lam yaCuC*. On the other hand, the phrase *lam yaku*, albeit semantically equivalent to *lam yakun* cannot be measured against *lam yaqum* because it does not match this latter relative to *lam yaCuC*, and then is said to be *šādd*.
A non-trivial by-product of the operation of 'measurement' is that it allows the observer to find out the property shared by A (say, *lam yakun*) and B (*lam yaqum*), a kind of standard unit of measurement or *tertium comparationis*. This property turns out to be a pattern of regularity on morphosyntactic and morpho-phonological level, e.g. *(lam) yaCuC*, which in morpho-phonology is technically denoted as WAZN, 'weight' or ṢĪĠAH, 'form'.
Therefore, a more appropriate translation of *qiyās* would be 'the ability to be measured on the basis of a pattern of regularity'.
Arabic Linguistic Tradition, however, did not content itself with observing the patterns of regularity at issue and already Early Grammarians, adopting a modern scientific attitude, relied upon *observation* of *regular* phenomena to *replicate* them. Actually, they tried to invent new words or to ipercorrect the existing irregularities via analogical extension of the linguistic regularities they had previously discovered: a technique called in Arabic Linguistic Tradition *'ihmāl* and (very approximately) translated here as OVERGENERATION (Bohas et al., 1990, pp. 149-150).
On these grounds, the notion of *qiyās* can be also likened in part to the Western concept of analogy, and opposed to the notion of linguistic usage (*samā'*) provided that one is aware of the fact that in the Arabic Linguistic Tradition 'analogy (vs. usage)' is only the applicative side effect[39] of a theoretical trend in which *qiyās* is generally taken to be *not* opposed to but

[38] This word is an infinitive form (MAṢDAR) from the verb *qāsa, yaqūsu*, 'to measure'. The infinitive form in Arabic conveys both an active and a passive meaning (more or less like the English infinitive *-ing* form). The passive interpretation proposed here is based on the fact that the finite counterpart of *qiyās* in the grammatical treatises is the expression *X (lā) yuqāsu 'alà Y*, 'X can(not) be founded on Y' (Bohas et al., 1990, p. 20).

[39] This account is in line with Bohas et al. (1990, p. 26)'s claim that the opposition *qiyās* / analogy vs.*samā'* / usage is 'a secondary development'. Versteegh (1997b, p. 35) adopts a stronger position, contending that *qiyās*

intrinsic in (the order (*naẓm*) of) *samā'* (see above).
It is worth observing in this connection that 'analogy (vs. usage)' was criticized explicitly by Sībawayhi (Versteegh, 1997b, p. 35), because he considered this kind of *qiyās* as an attempt to modify the *kalām al-'Arab* attested in the linguistic usage (*samā'*), something totally at odds with the aim of language preservation that founds the Arabic linguistic tradition. As such, this meaning of *qiyās* is negligible here.
Nonetheless, the debate around the applicative / analogical *qiyās* has the merit of raising the question of the reliability of data: to which extent did Arabic Linguistic Tradition (and its informants) manipulate and / or forge the transmitted corpus?
To answer this question, it is convenient to take into account textual research by Alhawary (2003), which shows that in their works Early Arab Grammarians and Lexicographers implement, report, and debate around a data collection procedure based on four factors, summarized in the following:

- Informants' background and locale:
  three criteria for informant selection were established:
  - Locale selection: informants were sought in their own habitat, especially in the middle of Arabia, to control for outside influences on their speech habits.
  - Native speakers' competence: informants born or raised in the desert and, in the best case, illiterate. People born in the desert but dwelling in the city were tested periodically, probably to investigate language attrition. At any rate, their information is considered weak and unreliable.
  - Reaction to grammatical errors and linguistic terminology: informants exposed to ungrammatical forms, words, constructions and sentences / utterances, if unable to detect the errors, were disqualified. So were people exposed to sentences / utterances containing linguistic terminology, which appeared to understand it.
  - Trustworthiness: Sībawayhi often vouches for his informants as trusted persons. As for the scholars transmitting the corpus, their chain of transmitters has to be traceable and every transmitter is required to be of recognized status.

- Informants' gender and age:
  the sample of informants ranges across gender and age in that both young and elderly, male and female native speakers are represented, although these latter sometimes go unreported or are not identifiable 'due to cultural-religious and / or personal reasons' (Alhawary, 2003, p. 5)

- Elicitation techniques:
  they can be characterized as follows:
  - informants are not given any information about the elicitation session
  - elicitation prompts are open-ended and targeting unrehearsed data
  - informants are in some cases constrained by time limitations

'represents a totally different concept' from analogy.

- Data recording:
  there are direct references and witnesses in the literature that elicitors did not rely on memory but rather committed the data to writing on the spot, using boards and sheets. The elicitors' reports include the use of hedges and detailed orthographic description (e.g. mood and case endings). An example in point is the so called 'commentary' (*ta'līqah*) by 'Alī ibn 'Abī Ṭālib (d. 40 / 660), the Prophet's cousin and son-in-law.

According to Alhawary (2003, p. 21) this demonstrates that, barring human errors and poetry, the corpus of data transmitted by the Early scholars was gathered 'from reliable sources using specific criteria' and that it was 'objectively elicited and promptly committed to writing'.
In this light, the traditional data collection procedure has no longer that defective status which makes modern scholars suspicious of the authenticity of the corpus known as *kalām al-'Arab* (see Alhawary (2003) for references). Accordingly, data taken to be reliable by Arabic Linguistic Tradition will be generally considered as such also in this work.
The only notable exception is poetry, since in its case a serious and scrupulous data collection procedure is not documented, which prevents to consider it as reliable information. This fact, however, seems to have no bearing on the authenticity of the linguistic facts discussed by the Arab Grammarians and Lexicographers (see Rabin's remark reported in Section 2.2.2.4).

2.3.3 The systematization of the corpus: *'aṣl* and *far'*

A major interpretive tool of Arab Grammarians' and Lexicographers' analysis - if not even the most important one - is the so-called 'AṢL. All the other tools one way or the other have to do with this key-notion, as it will become clear shortly.
The term *'aṣl* may be generally translated as 'origin', 'base' when referred to linguistic data (Baalbaki, 2006, p. 193)[40].
One of the most stringent definitions of this concept is provided by the Late Grammarian al-Zajjājī (d. 340 / 951) in his work *al-'Iḍāḥ fī 'ilal al-naḥw* (*apud* Versteegh 1995, p. 51[41]):

> "A thing can have a generally accepted status of *'aṣl* [*'aṣl mujtama' 'alay-hi*], and then, some of its elements can depart from it because of a weakness [*'illah*; it should be noted that this term also means 'cause' and 'explanation'] which affects them."

The element that departs from the *'aṣl* can be designated with the term FAR', which has the advantage of pinpointing the opposition between these two linguistic entities. It is however important to bear in mind that such an opposition in the Arabic Linguistic Tradition is only a *methodological fiction* introduced by the Late Grammarians (Baalbaki, 2006, p. 193), which should not obscure the ontological status of the *far'*. This is no more than an *'aṣl* plus 'something else' (the so-called *'illah*):

(37) Far' = 'Aṣl + $X_{_'illah}$

On this scenario, *identifying the 'illah* (in Arabic, TA'LĪL) is one of the major concerns of the Arabic Linguistic Tradition, because once that one has found this element and abstracted away

[40] Baalbaki reports a third translation of *'aṣl*, namely 'principle' but later (Baalbaki, 2006, p. 195) points out that it applies when *'aṣl* is used in an epistemological sense: cp. also the discussion below for further details.
[41] Translation and glosses are based on Guillaume 2006, p. 179, with minor modifications

from it, one is in a good position to reduce the *far‘* to the *’aṣl*. Thus, the *‘illah* could be likened in some respects to the disturbing factor of physics (e.g. attrition).
On Arab Grammarians’ view, these disturbing factors are generally due to ‘immediate sensory perception [*ḥiss*] that some sounds or sequences of sounds are ‘heavier’ or ‘lighter’ than others (i.e. that their pronunciation entails a greater of smaller expenditure of energy for the speaker...)’ (Guillaume, 2006, p. 178).
For example, the phonological change *awa* > *ā* is regarded by them as an *‘illah* that applies to abandon the ‘heavy’ (*ṯaqīl*) cluster *awa* in favor of the ‘light’ (*xafīf*) sound *ā*.
A non-trivial corollary of the operation of *ta‘līl* is that the *’aṣl* shows up vey clearly, which makes it possible for the researcher to focus on its properties.
In this connection, Late Arab Grammarians and in particular ’Ibn Jinnī (d. 392 / 1002), ’Ibn ‘Uṣfūr (d. 669 / 1271), al-’Astārābāḏī (d. 686 / 1287) furthermore observe that the *’aṣl*, when opposed to the *far‘* can have two different manifestations (Baalbaki, 2006, pp. 191-192 and references therein):

(I) it is the attested morphosyntactic entity which has undergone diachronic change and has been assigned as the origin from which a certain entity... has developed: e.g. the *’aṣl* of the phrase *lam yaku*, ‘he was not’, is *lam yakun*, a phrase documented in Old and Classical Arabic. What the two phrases differ in is the dropping of *-n*, which is therefore to be considered as the *‘illah*

(II) it is the supposed, but not necessarily attested, origin of a certain morphosyntactic entity, ie it is not necessarily meant to represent an historical stage of the language which has been later abandoned: e.g. the *’aṣl* of *qāma*, ‘he stood’ is *qawama*, a form not documented in Old and Classical Arabic, but posited on the basis of the general pattern *CaCaCa* denoting the perfect (e.g. *kataba* ‘he wrote’). What distinguishes *qāma* from its *’aṣl qawama* is the phonological change *awa* > *ā*, which is therefore to be considered as the *‘illah* (see above)

For expository purposes, the former type of *’aṣl* (Baalbaki (2006)’s ii.) will be referred to here as DIACHRONIC ’AṢL, the latter as SYNCHRONIC ’AṢL (Baalbaki (2006)’s iii.). Some clarifications at this point are in order.
First, the diachronic characterization of *’aṣl* summarized in (I) entails that (Late) Arab Grammarians, although generally lacking an historical perspective (= the comparative method), as traditionally assumed (Fleisch, 1961, vol. I, p. 46), ‘were certainly aware of the notion of change’ (Owens, 2006, pp. 75-76).
Textual research (Baalbaki, 2006, and references therein) has recently shown that this is actually the case and, in addition to that, this seems highly probable on epistemological grounds, since the very notion of linguistic change by definition stems from an approach to the linguistic data focusing on their *form* rather than on their meaning - which is precisely how Arab Grammarians conceived their discipline:

> “in language comparison priority must be given unconditionally to the form. If two forms correspond exactly or according to the rules, this compensates for some degree of discrepancy in the meaning. On the other hand, if two forms cannot be referred to a ‘common denominator’, this is not offset even by total agreement in

meaning. For reasons of form, *deus* (still *deiuos* in Latin) can no more be connected with [Ancient Greek] *theos* than with English *God*".
(Szemerényi, 1996, p. 16)

Second, the opposition *'aṣl* vs. *far'* on synchronic level illustrated in (II) can be identified with the modern concept of transformation after Owens (1988, pp. 220ff.) because of the non autonomous status of the *far'*, as depicted in (37).
In this regard, Guillaume (2006, p. 180) remarks that explaining a given linguistic fact under the lens of the opposition *'aṣl* vs. *far'* 'usually consists in relating it to another, supposedly more basic or prototypical one'. This statement that can be easily rephrased in transformational terms (see Section 1.1.3): the opposition *'aṣl* vs. *far'* captures the connection between two (or more) given linguistic entities, the connection being such that the former entity is intuitively more 'basic' than the latter (and is included in it).
This equivalence holds not only with respect to the *relationship* betwen the linguistic entities in and by itself, but also with respect to the linguistic *level* the linguistic entities belong to. The transformation, very much as the opposition *'aṣl* vs. *far'*, operates not only on (morpho-)syntactic level, affecting syntactic constructions, but also on (morpho-)phonological level, affecting phonological strings (Chomsky & Halle, 1968, pp. 358-364, 399, 427).
Accordingly, in this work the synchronic *'aṣl* (II) is taken to correspond to the underlying (morpho-)syntactic or (morpho-)phonological structure of modern generative-transformational grammar (or to its most recent incarnation). All this is said with the caveat that such an equivalence cannot be posited under certain circumstances (not discussed here: see Baalbaki 2006 and references therein for details) and that the synchronic *'aṣl* (II) has also been likened to the central category of the prototype theory (Guillaume, 2006).
Third, Baalbaki (2006, p. 191) finds in the Arab Grammarians' and Lexicographers' texts two other manifestations of the concept of *'aṣl*: it can also indicate the consonantal root and a linguistic entity which 'is perceived as the norm' (his i., iv.).
But it has been discussed at the end of Section 2.3.1 that the consonantal root, ie the phonotactic *'aṣl* is plausibly to be interpreted along the same lines of the morphosyntactic *'aṣl*, that is as the simplest / most basic phonotactic form, to the effect that Baalbaki (2006, p. 191)'s distinction between the two no longer holds. Accordingly, the consonantal root will fall either into the diachronic or the synchronic *'aṣl*[42].
As for the *'aṣl* that works as a normative referent, it clearly falls outside the description of the

[42]As a matter of fact, both the interpretations are found in the current literature. Sometimes, in fact, the Arabic root is identified with a basic form documented in diachrony (DIACHRONIC 'AṢL): e.g. the root of the perfect tense *'akala* 'he ate' is said to be (what has later developed into) its infinitive *'akl* 'eating'. This account is often dubbed in the literature STEM-APPROACH and is defended, among others, by Wright (1896, vol. I, p. 59), Garbini & Durand (1994, p. 112), Larcher (1995), Benmamoun (1999).
In the vast majority of cases the Arabic root is identified with a basic triconsonantal skeleton which is not attested in the usage, but can be extracted from several morphological forms in synchrony (SYNCHRONIC 'AṢL): this position is known in the literature as ROOT-APPROACH and is argued for by Cantineau (1950), McCarthy (1979), Prunet et al. (2000) and much other work.
Notice that in the present work Bohas (1997, 2000)'s groundbreaking proposal (the so-called THEORY OF ROOTS, ETYMONS, MATRICES) is subsumed under the root-approach label. Notice also that according to Bohas et al. (1990, Ch. 4) the Arabic linguistic tradition practised not only the root-approach, as traditionally assumed (Fleisch, 1961, vol. I, pp. 247-251), but also the stem-approach. Mahfoudhi (2007) provides a critical overview of both the approaches.

primary linguistic data driven by the need of preserving the *kalām al-'Arab*. Indeed, it calls into the picture elements of prescriptivism having to do with the need of codifying the *kalām al-'Arab* into Classical Arabic, an irrelevant matter here (see above).
It should be added that in the Late grammatical works, the term *'aṣl*, especially in the plural *'uṣūl*, also denotes the 'foundations' or 'principles' of the grammar. This is because classification of grammatical topics was made according to an (onto-)logical hierarchy which proceeds from the general, ie the *'aṣl* itself, to the particular, ie the *far'*, also called in this case *mas'alah*, '(subsidiary or secondary) question' (Baalbaki, 2006, p. 193).
To sum up, the discussion so far has led to characterize the opposition *'aṣl* vs. *far'* as either diachronic or synchronic. In the former case, the *'aṣl* is the old or relic form etc..., while the *far'* is the new form. In the latter case, the *'aṣl* represents the (syntactic or phonological) underlying structure (or the alike), the *far'* being the derived form etc... that is actually spelled out.

2.3.4 An integrated methodology

Upon closer scrutiny, the opposition regular (*muṭṭarid*) vs. irregular (*šāḏḏ*), which is empirically-grounded, seems to interact with the opposition *'aṣl* vs. *far'*, axiomatic in its nature. The careful reader will have in fact noticed that in the just discussed pair *lam yakun - lam yaku* the former opposition is interlocked with the latter, because *lam yakun* is at the same time *'aṣl* and regular, while in *lam yaku* the status of *far'* coexists with its being irregular.
In principle, the interaction of these two oppositions yields four logical possibilities, which can be doubled if one recalls that the opposition *'aṣl* vs. *far'* is found both on diachronic and synchronic level:

- in diachrony, a regular *'aṣl* is opposed to a regular *far'*
- in diachrony, an *ir*regular *'aṣl* is opposed to an *ir*regular far'
- in diachrony, a regular *'aṣl* is opposed to an *ir*regular *far'*
- in diachrony, an *ir*regular *'aṣl* is opposed to a regular *far'*
- in synchrony, a regular *'aṣl* is opposed to a regular *far'*
- in synchrony, an *ir*regular *'aṣl* is opposed to an *ir*regular *far'*
- in synchrony, a regular *'aṣl* is opposed to an *ir*regular *far'*
- in synchrony, an *ir*regular *'aṣl* is opposed to a regular *far'*

As a matter of fact, Fleisch (1961), Owens (1988), Bohas et al. (1990), Versteegh (1997b), Baalbaki (2006), Guillaume (2006) etc... report and recognize in the primary sources all these logical possibilities, except the fourth and fifth one, which, however, can be easily observed in the following passage of al-Xalīl's Introduction to his *Kitāb al-'Ayn* (*apud* Haywood 1965, pp. 29-30, 34, 36, his translation and glosses):

> "To make biliterals like "qad", "hal", and "law" into nouns, you must introduce the tashdīd and say (for example): "hādhihi lawwun mukawwanatun"[...], adding a wāw to a wāw [...]. Then you incorporate the two, putting a tashdīd, which is

the sign of incorporation and (indication of) the third letter (of the noun).[...] The Arabs frequently derive reduplicated forms from the stressed triliteral in which the second and third radical are identical[...]. Do you notice that they say: "ṣalla al-lijām" (the bridle crinkled)[...]. When they say that onomatopoeically, they stretch and stress the lām. But they lightened it in "ṣalṣala". The stress is a lengthening, while the reduplication is a repetition which has the effect of lightening. Such roots cannot be properly conjugated until they are either reduplicated or stressed (that is, the second radical is doubled), as I have described to you.[...] Know that the biliteral doubled word runs in two permutations, as radda and darra, shadda and dashsha. "

The relevant examples are provided in the Tables 2.9, 2.10[43], while the interplay between the two oppositions under discussion is summarized in the Fig. 2.10, which takes into account also some notions interlocked with such oppositions[44].
On the whole, Tables 2.9, 2.10 and the Fig. 2.10 summarize what will be henceforth referred to as 'AṢL-BASED MODEL.

The Old Arabic data discussed in the next Chapters, being mainly drawn from Arab Grammarians' work, may turn out to be framed within the *'aṣl*-based model: this makes it possible to obtain useful information about the diachronic phenomena and the mental processes in which such data are involved.
Notice also that, by extension, the same information may apply to the data taken from the modern Colloquial Varieties when these latter are identical to Old Arabic, as is often the case (see Section 2.2.2.6).

[43]The synoptic account given here flattens the historical landscape in which the notion of *'aṣl* developed. For example, as Baalbaki (2006, p. 192) observes, in Sībawayhi's work *'aṣl* is strongly associated with the notion of 'regular' (*muṭṭarid*) and opposed to that of 'irregular' (*šāḏḏ*).
Another case in point is the *'aṣl*-form *'istaḥwaḏa*, 'he overwhelmed ' (see Table 2.10), taken to be irregular in the Late Arabic Linguistic Tradition, but regular by the Early Lexicographer 'Abū Zayd al-'Anṣārī (d. 215 / 830): see Lane (1863, vol. II, p. 663) and *Lisān al-'Arab*, vol. II, p. 1041, s.v. *Ḥ W Ḏ*.
These details, however, do not affect the point made here and are not further discussed.
On the other hand, some other consideration should be borne in mind throughout the examination of the Table 2.10. They concern the just cited passage taken from al-Xalīl's work and particularly the *'aṣl* of the particles and the nouns mentioned in it.
Al-Xalīl, in fact, explicitly states that the particles *qad*, *hal*, and *law* and the nouns *yad-un*, 'hand', *dam-un*, 'blood', *fam-un*, 'mouth', in spite of the identical surface stem differ in their underlying phonological structure (*'aṣl*). On this level, the former are biconsonantal, the latter triconsonantal, given that their last consonant, being a weak consonant (a glide) is dropped in the spell-out (*yadw-un, damw-un, famw-un*).
In Haywood (1965, p. 30)'s translation: 'nouns do occur formed from two letters, whereas their complete form and their significance is triliteral as in "yad", "dam", and "fam". For the third letter has gone because of its weakness'.
On the other hand, in the Early Arabic Linguistic Tradition the quantifier *kam*, 'how many / much', is interpreted as a noun and assigned the *'aṣl kam-un* (see Table 2.10 for details): thus, this latter cannot be 'measured against (*yuqās 'alà*)' the *'aṣl yadw-un* etc..., because of the different stem.
The same holds for the *far' kam*, which differs from the *far' yad-un* etc... for its inability to be associated with the morphemic material *-un* in non-pausal forms. In other words, both *kam-un* and *kam* are irregular.

[44]The interplay among *'aṣl* and the other concepts of the Arabic Linguistic Tradition is also discussed at length in Suleiman (1999), although it is given an account different from that exemplified in the Fig. 2.10.

Table 2.9: The manifestations of *'aṣl* #1

'Aṣl		Description		Western equivalent	
Synchronic		posited by Arab Grammarians		Transformation	
	'AṢL	FAR‘	‘Illah	Source	*apud*
	REGULAR	REGULAR			
Sample	(qawama)	qāma	awa > ā	Ibn Jinnī etc...	Baalbaki (2006, p. 192) etc...
Gloss	he stood up	idem			
Cp.	ḏahaba	kāna			
Gloss	he went	he was			
	IRREGULAR	IRREGULAR			
Sample	(kam-un)	kam	No declension	al-'Axfaš etc... (d. 221 / 835)	Owens (2006, p. 76)
Gloss	how many?, how much?				
Cp.	-	-			
Gloss	-	-			
	REGULAR	IRREGULAR			
Sample	(quwuws-un)	qisiyy-un	wuw > uww uww > iyy	al-Mubarrad etc... (d. 285 / 898)	Bohas et al. (1990, p. 21)
Gloss	bows	idem			
Cp.	‘uluwm-un	-			
Gloss	sciences	-			
	IRREGULAR	REGULAR			
Sample	(ṣala)	ṣalla	Ca > CCa	al-Xalīl	(see above)
Gloss	it crinkled	idem			
Cp.	-	radda			
Gloss	-	he got back			

2.3.5 The primary sources of this work

The Old (and Classical) Arabic data set forth in this work are mainly drawn from the works mentioned below.
They are briefly contextualized along the lines of Guillaume (2000):

- *'Al-Kitāb*
  - English translation of the title: The Book
  - Author: SĪBAWAYHI, ‘Amr 'Ibn ‘Uṯmān 'Ibn Qanbar (d. 177 / 793)
  - Textual genre: grammatical treatise
  - Research areas: morpho-syntax, pragmatics, phonology and phonetics
  - Additional information: see Section 2.2.2.4
- *Ma‘ānī l-Qur'ān*

  - English translation of the title: The Meanings of the Koran
  - Author: AL-FARRĀ', 'Abū Zakariyyā Yaḥyà 'Ibn Ziyād (d. 207 / 822)
  - Textual genre: Koranic commentary
  - Research areas: grammar in general and lexicon (as well as theological and legal exegesis)
  - Additional information: among the few extant works by a proponent of the Kūfan school

- *Šarḥ al-Mufaṣṣal*
  - English translation of the title: The Commentary on The Treatise systematically Arranged into Sections [by al-Zamaxšarī]
  - Author: IBN YA'ĪŠ, Muwaffaq al-Dīn 'Abū l-Barā' Ya'īš (d. 643 / 1245)
  - Textual genre: commentary on a grammatical treatise
  - Research areas: morpho-syntax, phonology and phonetics, theory of predication
  - Additional information: based on the grammatical treatise *al-Mufaṣṣal* ('The Treatise systematically Arranged into Sections') by AL-ZAMAXŠARĪ, Jār 'Alļāh 'Abū l-Qāsim Maḥmūd 'Ibn 'Amr (d. 538 / 1144)

- *Šarḥ al-Kāfiyah*
  - English translation of the title: The Commentary on The Sufficient Compendium [by 'Ibn al-Ḥājib]
  - Author: AL-'ASTĀRĀBĀḎĪ, Rāḍī al-Dīn Muḥammad 'Ibn 'al Ḥasan (d. 686 / 1287)
  - Textual genre: commentary on a grammatical treatise
  - Research areas: morpho-syntax, pragmatics, semantics, theory of predication
  - Additional information: he develops an account of the Arabic Case on semantic assumptions, rather than formal ones, as customary in the Arabic linguistic tradition. His work is based on the grammatical treatise *al-Kāfiyah* ('The Sufficient Compendium') by 'IBN AL-ḤĀJIB, 'Abū 'Amr 'Uṯmān 'Ibn 'Umar 'Ibn 'Abī Bakr al-Malikī (d. 646 / 1249).

- *al-Muzhir fī 'ulūm al-luġah wa-'anwā'i-hā*
  - English translation of the title: The Enlightening Treatise on Philology
  - Author: AL-SUYŪṬĪ, Jalāl al-Dīn 'Abū l-Faḍl 'Abd al-Raḥmān. (d. 911 / 1505)
  - Textual genre: philological treatise
  - Research areas: relic forms, rare words and expressions, organization of the lexicon
  - Additional information: compilative work

- *Lisān al-'Arab*
  - English translation of the title: The Language of the Arabs
  - Author: 'IBN MANẒŪR, Muḥammad 'Ibn Mukarram Ibn 'Alī Riḍwān (d. 711 / 1311)

- Textual genre: dictionary (onomasiological arrangement)
- Research areas: lexicon (but Arab Grammarians' observations about Arabic morpho-syntax are also reported)
- Additional information: compilation from some previous major dictionaries (Haywood, 1965, Ch. 7)

2.4 Summary

The Arabic language is better understood in synchrony as made up of Classical Arabic and the modern Colloquial Varieties. This Chapter introduces this state of affairs also from a diachronic perspective, relying upon recent research by Owens (2006). This latter plausibly demonstrates that Classical Arabic is an evolving entity, developing out of Old Arabic, while the modern Colloquial Varieties behave as a whole as an 'atemporal' entity (Owens, 2006, p. 114) insofar as some key-features are concerned - among them, lack of (morphological) Case.
In so doing, Owens's theory highlights the strong relationship between the modern Colloquial Varieties and Old Arabic and holds that at their core the former ones are as old as the latter. His claim is explored here with particular regard to a Bedouin Colloquial Variety, Najdi Arabic, as described e.g. in Ingham (1994), and is given a practical methodological implementation: analysis of the Arabic data of the present dissertation concentrates on the Colloquial Varieties and Old Arabic, assigning Classical Arabic a more modest role.
Chapter 2 follows Owens (2006) also in defining the nature, the chronological limits and the philological sources of Old Arabic and shows that such a definition is totally in line with characterization of the same linguistic entity provided in Semitic studies by Garbini & Durand (1994). Bringing these two theories in comparison has the advantage of accommodating Owens's diachronic theory on *Arabic* into the broader diachronic scenario of Garbini & Durand (1994), which describes the position of Arabic within *Semitic*. In greater detail, they regard Old Arabic as a mixed language resulting from language contact between a Semitic - more precisely, Amorite - parastrate and a pre-Semitic parastrate, a linguistic system very close to Hamitic (Language Contact Hypothesis).
The practical implication of this move, introduced in this Chapter and developed in the next ones, is that the features typical of these parastrates seem to have played at least originally a role sensibly different from that designated by their traditional name. A case in point is the (putative) Case, probably to be rethought of as a copula (Nyberg, 1920), an analytical trend referred to here as Copula Hypothesis.
In greater detail, the key-features the literature has ascribed to Amorite are

- the pronominal / demonstrative system: the bases *d̲-, h-*...
- the particle *la-* (see (169) in the next Chapter)
- the suffix-conjugation

while those ascribed to pre-Semitic are:

- the broken plurals
- the prenominal article *l-*

- the Case (phonologically realized *u, a, i* etc...: see next Chapter)

Finally, Chapter 2 provides in-depth discussion about the source of Old Arabic data, mainly drawn from Arab Grammarians' and Lexicographers' works, and defends their validity, stressing the point that they are framed within a mature scientific / linguistic paradigm.
The Arabic Linguistic Tradition accordingly provides with the modern linguist useful information about the diachronic change and the mental representations concerning Old Arabic, as well as and the judgments of Old Arabic native speakers.
In particular, Arab Grammarians'term *'aṣl* refers both to an old / relic form attested in diachrony and to the pre-transformational representation of a given morpho-phonological / morpho-syntactic form in synchrony ('Deep Structure' , 'phonological representation').

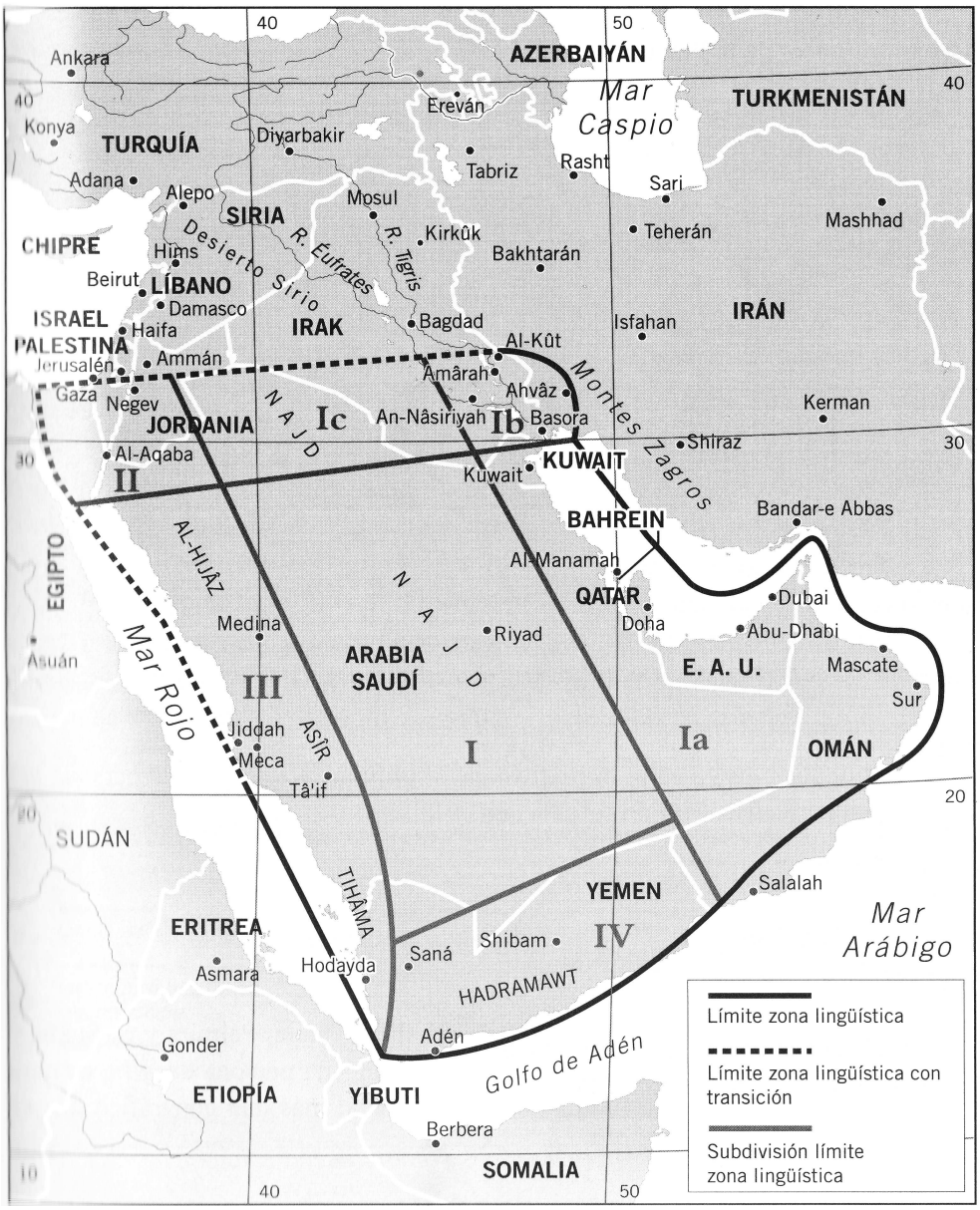


Figure 2.2: Arabian Peninsula: Linguistic Areas (Abboud-Haggar, 2003, p. 119)

Figure 2.3: Najdi Arabic: Dialectal variation (Ingham, 1994, p. xvii)

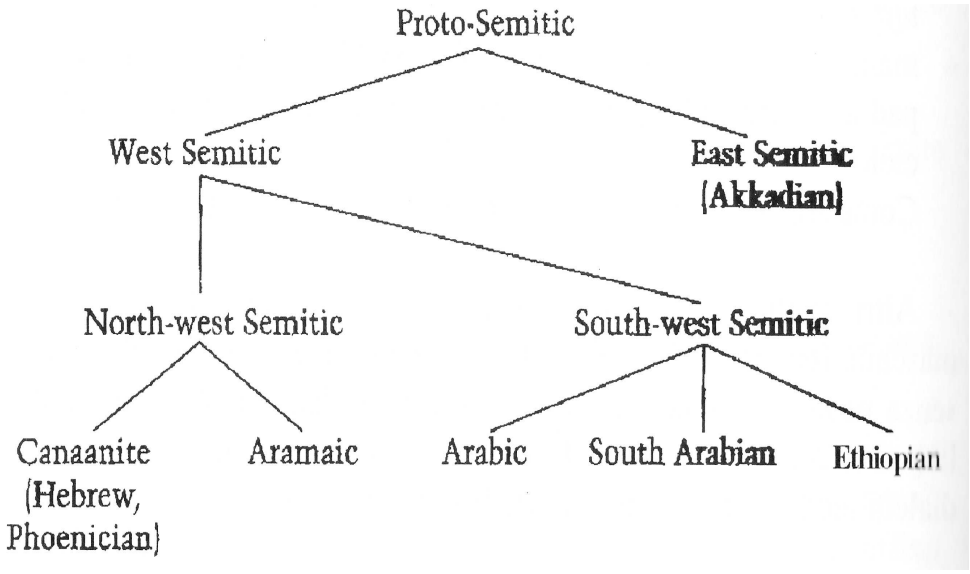


Figure 2.4: The Semitic Languages (Versteegh, 1997a, p. 12)

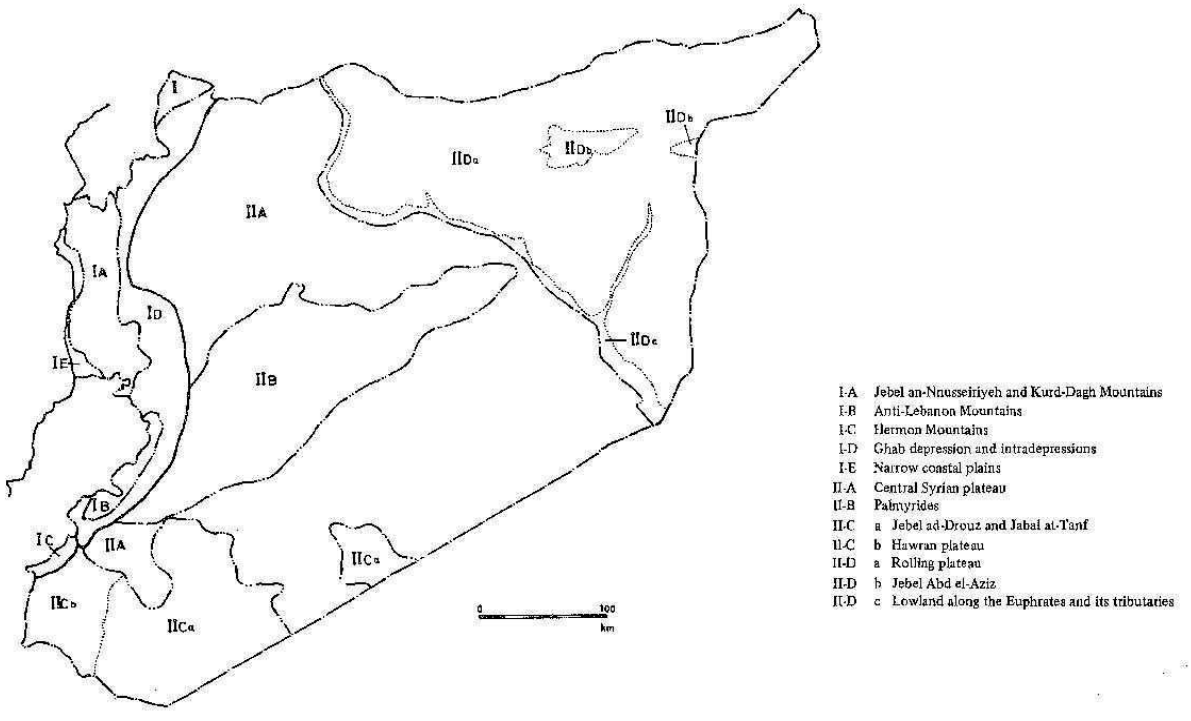


Figure IV-9. Map showing the physiographic divisions of Syria.

Figure 2.5: Syria: geographical map (Suzuki, 1973)

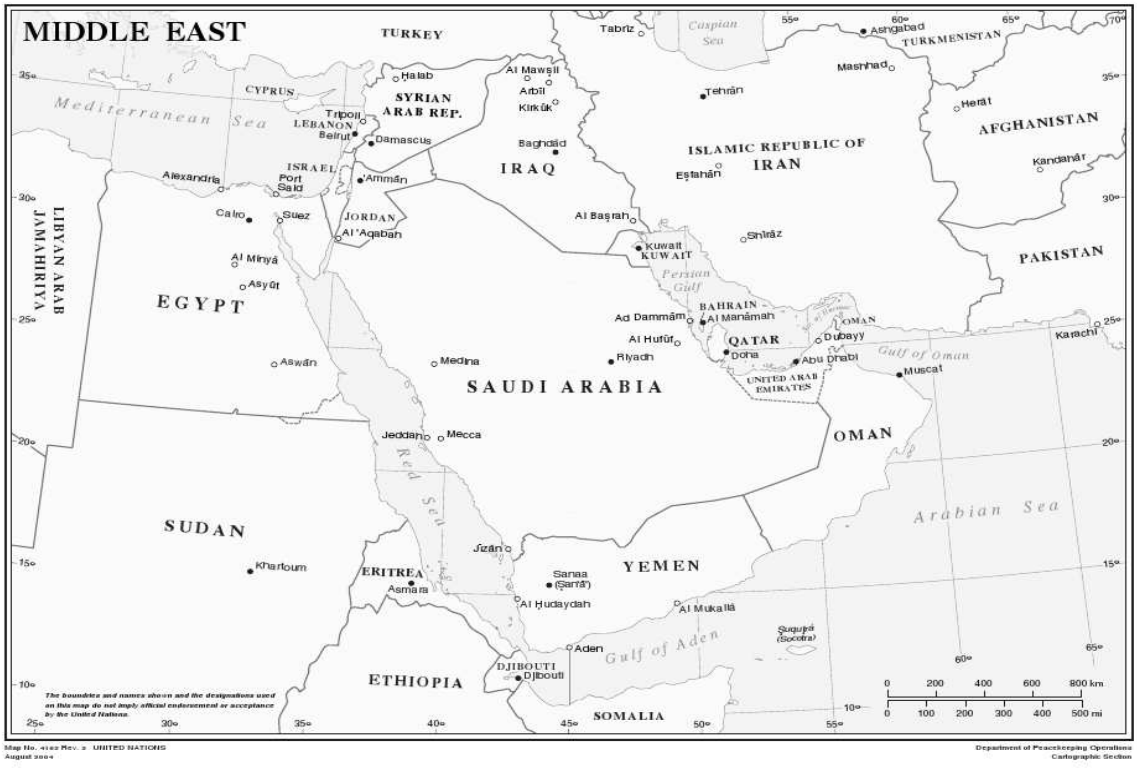


Figure 2.6: Middle East: UN Map No. 4102 (Rev. 3 August 2004)

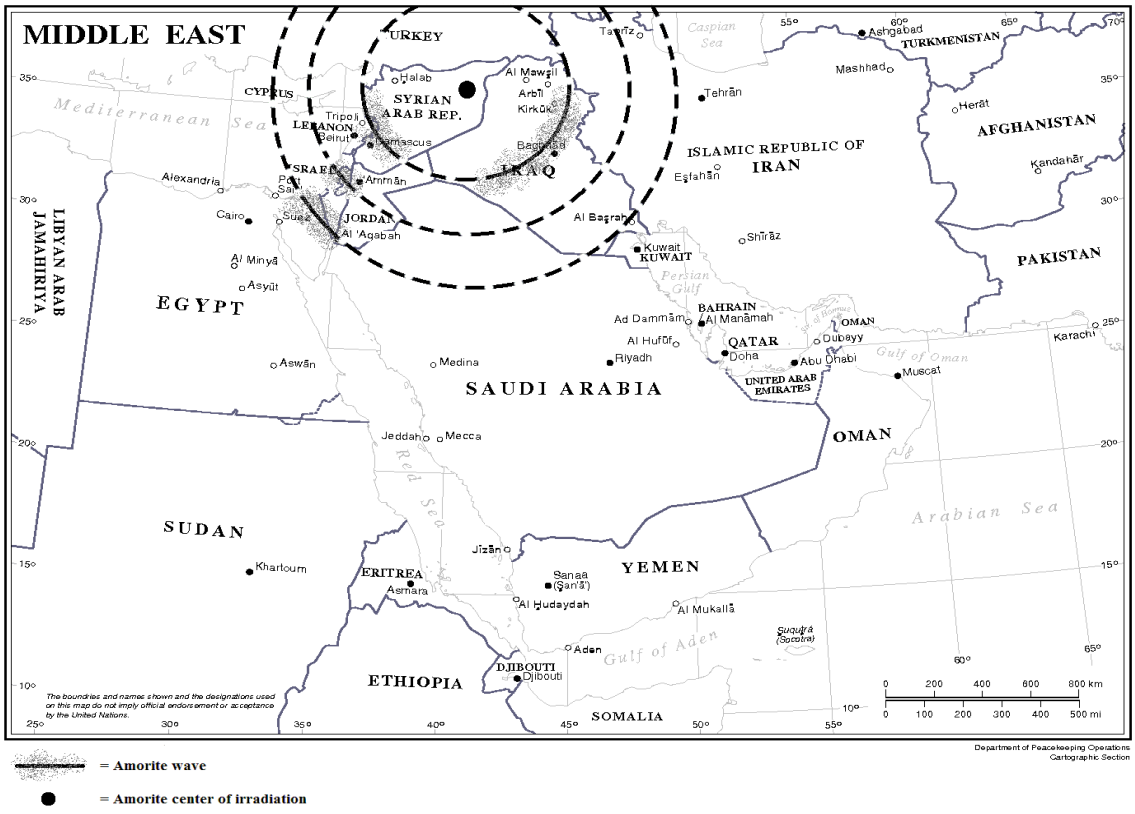


Figure 2.7: Amorite: Centre of irradiation and waves (map based on Fig. 2.6)

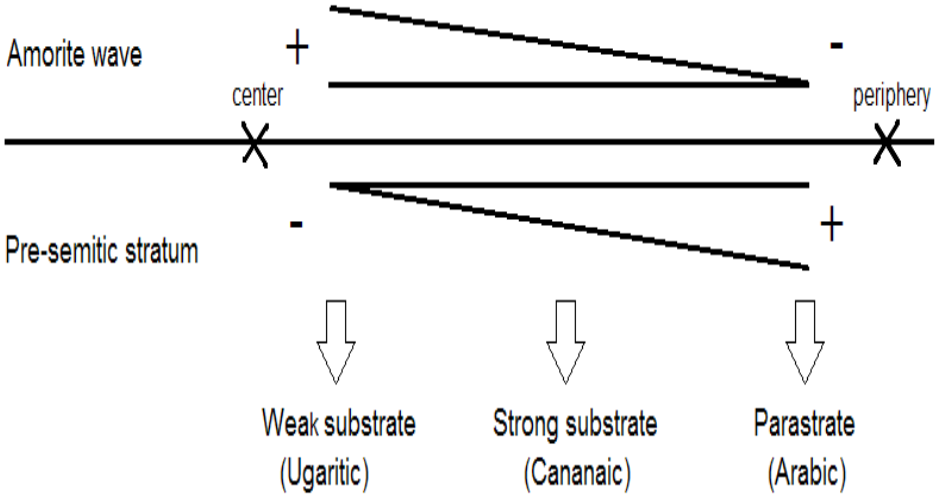


Figure 2.8: Pre-Semitic: Degree of convergence

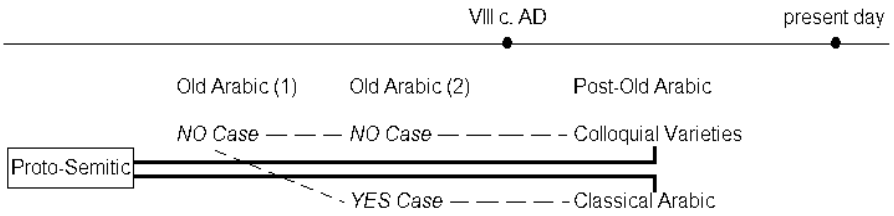


Figure 2.9: Old Arabic and Classical Arabic (Owens, 2006, p. 115 (adapted))

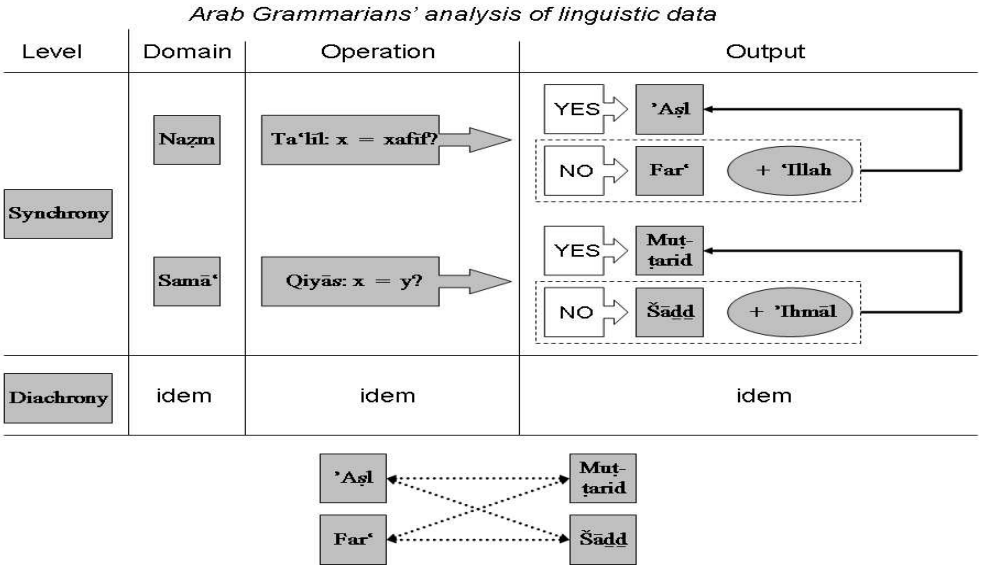


Figure 2.10: The *'aṣl*-based model

Table 2.10: The manifestations of *'aṣl* #2

'Aṣl		Description		Western equivalent	
Diachronic		attested in the usage		Diachronic change	
	'AṢL	FAR‘	‘Illah	Source	*apud*
	REGULAR	REGULAR			
Sample	ṣalla	ṣalṣala	Stem Reduplication	al-Xalīl	(see above)
Gloss	it crinkled	idem			
Cp.	radda	xarxara			
Gloss	he got back	it fluttered			
	IRREGULAR	IRREGULAR			
Sample	min ḏū	munḏu	Univerbation V̄# > V# Vowel Harmony	al-Farra'	Fleisch (1961, vol II, pp. 92, 488), *Šarḥ al-Mufaṣṣal*, vol. IV, *Faṣl ma‘nà muḏ wa-munḏu*, pp. 506 -507
Gloss	from the (time) that	idem, from			
Cp.	-	-			
Gloss	-	-			
	REGULAR	IRREGULAR			
Sample	lam yakun	lam yaku	n-dropping	Sībawayhi	Baalbaki (2006, p. 191)
Gloss	he was not	idem			
Cp.	lam yaqum	-			
Gloss	he did not stand up	-			
	IRREGULAR	REGULAR			
Sample	'istaḥwaḏa	'istaḥāḏa	wa > ā	Ibn Jinnī	Baalbaki (2006, p. 192), *Lisān al-‘Arab*, vol. II, p. 1041, s.v. *Ḥ W Ḏ*
Gloss	he overwhelmed	idem			
Cp.	3 tokens only: 'istarwaḥa 'istaṣwaba 'istajwaba	'ista‘āna			
Gloss	he sighed he approved of he questioned	he asked for help			

Chapter 3

The Arabic Noun Phrase: (rethinking) the received view

3.1 The Arabic Noun Phrase: the state-of-the-art

The received view of the Arabic adnominal markers and of its modification-structure goes back to the seventeenth century (Versteegh, 1997a, p. 3) and in the majority of cases it interprets these phenomenona through the lens of the Western traditional grammar[1].
It is convenient taking the received view as a starting point, as Section 3.2 does, since it is still given a good amount of consideration until today in both the academic research and the didactic practice, so that both the specialized and divulgative literature use to a different degree its categories.
Yet, this should not obscure the fact that the explanatory power of the received view has been questioned over the centuries and since the beginnings of the Arabist studies, scholars have put forward many alternative proposals, both within the received view itself and outside it, in order to account for the same phenomena. Such proposals will be reviewed in the Section 3.3.
Needless to say, all the hypotheses and data presented in what follows are not fresh. Nonetheless, a novel idea will be pursued and evidenced throughout this chapter, namely that all the phenomena concerning the Old Arabic Noun Phrase can be provided an unified account under the notion of specificity (Specificity Hypothesis). In so doing, this chapter aims at extending to Old Arabic a proposal originally developed by Brustad (2000) for the modern Colloquial Noun Phrase.
This theoretical move has two non-trivial consequence: Brustad (2000)'s Specificity Hypothesis, in fact, when applied to Old Arabic

- neutralizes and makes questionable the opposition definite / indefinite in which the received view usually frames Old Arabic (and Classical Arabic) data.

[1]It was in fact Bedwell (1563-1632) and Erpenius (1584-1624) that laid down the foundations for a philological study of Arabic in the modern sense. Nevertheless, as Versteegh (1997a, p. 2) remarks, 'the first analysis of Arabic on the basis of a Greco-Latin model' can be traced back even earlier, to the Spanish scholar Pedro de Alcalà, who in 1505 published a large dictionary of (Spanish) Arabic, the *Vocabulista aravigo en letra castellana* and a manual of Arabic grammar (with a conversation guide for the confessional), the *Arte para ligera mente saber la lengua araviga*.

- makes it possible to set a *pervasive* equivalence between Old Arabic and the modern Colloquial Varieties at least in the Noun Phrase domain, going far beyond the sharing of scattered features discussed in Section 2.2.2.6 and summarized in Table 2.7

3.2 The received view

3.2.1 Case, number, gender, definiteness

The shape of the Arabic Noun Phrase (NP henceforth) depends on the variety taken into account: accordingly, Classical Arabic will be taken here as a starting point for expository purposes, but it is by no means exhaustive of all the properties of the Arabic NP, which will be introduced in due course.

In Classical Arabic the noun has been traditionally analyzed by Western scholars as marked for case (NOM, ACC, OBL), number (SG, DU, PL), gender (M, F) and definiteness (the so-called 'states'): see e.g. Wright (1896, vol. I), Veccia Vaglieri (1937), Fleisch (1961, vol. II) and Hoyt (2006)[2].

To begin with, case is manifested as a set of suffixal case-endings.

Slightly simplifying, the NOM case signals the subject, the ACC case the object as well as many adjuncts, and the OBL case the complement of specification and, if preceded by a preposition, other adjuncts.

While in the SG the three cases are generally (though not always) kept distinct, in the DU, PL, the ACC, OBL cases are expressed by the same marker without exception.

As for number, PL obtains either by lengthening case-endings, which yields the so-called SOUND PL or AL-JAM' AL-SĀLIM or via apophony: this results in the so-called BROKEN PL or JAM' AL-TAKSĪR, whose case-endings are the same of the SG.

It should also be noted that DU exactly behaves like SOUND PL.

Finally, the NP in Classical Arabic displays three forms (or states) that relate to definiteness: indefiniteness, definiteness triggered by the article and definiteness triggered by a complement of specification.

In the SG and in the BROKEN PL, indefiniteness is signalled by the suffix *-n* (NUNATION or TANWĪN) and definiteness by the prefix *l-* (LĀM AL-TA'RĪF), which is traditionally identified with the Indo-european (PRENOMINAL) ARTICLE.

Alternatively, the complement of specification (also called GENITIVE PHRASE) is taken in the Arabic Linguistic Tradition to have the ability of replacing the article to convey definiteness. In this case the NP to be made definite is simply combined with (or ANNEXED TO) a definite or indefinite genitive phrase that always follows it and not only the tanwīn *-n*, but also the article *l-* are barred.

Arabists and Semitists often designate the complex thus formed as CONSTRUCT STATE, while Arab Grammarians call it 'IḌĀFAH (and, more precisely, they refer to the modified NP as MUḌĀF and to the genitive phrase modifying it as MUḌĀF 'ILAY-HI).

On the other hand, the SOUND PL, when indefinite, displays a suffix *-na*, and so does it when made definite by the article *l-*: but, when annexed to a genitive phrase, it loses this suffix. This latter is not singled out as a linguistic entity by Western scholars, while in Arabic Linguistic

[2]See also Levin (1986), Levin (2006) on Arab Grammarians' analysis of the CA / OA noun.

Tradition it is sometimes designated as NŪN AL-JAM‘ (literally, ‘the *-n* of the plural’).
The DU behaves like the SOUND PL, but in this case the suffix is *-ni* instead of *-na*: in the Arabic Linguistic Tradition, the DU suffix is designated as NŪN AL-TAṮNIYAH (literally, ‘the *-n* of the dual’).
Such a terminology reflects some Grammarians’ and native speakers’ judgments, according to which the endings *-na, -ni* are the *locus* of (SOUND) PL and DU, respectively, but their position was not a matter of consensus, since other Grammarians and native speakers maintained that these pieces of information are encoded into the long vowels preceding *-na, -ni* (Goldenberg, 1988).
Only the latter position has been accepted in the Western traditional paradigm (see Section 3.3.2):

(38) Arab Gramm. #1 *-ū + na, -ā + ni* = *-ū* / *-ā* + PL / DU
Received view / Arab. Gramm. #2: *-ū + na, -ā + ni* = PL / DU + *na* / *ni*

Turning to the gender, the M represents the default form and the F obtains by adding a suffix to the SG M, a suffix generally realized as *-at*, the so-called BOUND -T or TĀ’ MARBŪṬAH[3].
Its SOUND PL is formed by lengthening the *a* of *-at*, while the BROKEN PL behaves like its M counterpart.
As for its states, both in the SG and in the PL, they are identical with the SING and the BROKEN PL of the M: that is, they add the *tanwīn -n* instead of *-na* etc... Finally, the case-endings of the F are identical with those of the M.
The Classical Arabic NP thus characterized is exemplified in the Tables 3.1, 3.2, 3.3 by the stems *mu’min-*, ‘believer (M)’, *mu’min-at-*, ‘believer (F)’, *ṭarīq-*, ‘street (M / F in Classical Arabic)’, *ṭuruq-*, ‘streets (M / F in Classical Arabic).
When the NP is found in the construct state, its complement of specification is exemplified by the definite genitive phrase *l-qaryati*, ‘of the town’, but the indefinite genitive phrase *qaryati-n*, ‘of a town’ would be fine as well.

On the whole, the just mentioned declensed forms can be also classified as NON-PAUSAL to distinguish them from the so-called PAUSAL forms, that is forms taken by the NP when it occurs at the end of a breath-group or PAUSE (e.g. the end of the sentence).
The pause results in dropping some word-final material, typically case-endings (when they are short vowels) and the *tanwīn -n* following them. This is why pausal forms are also referred to as caseless forms:

(39) NON-PAUSAL FORM = DECLENSED FORM
PAUSAL FORM = CASELESS FORM

The only exception is the case ACC in the default form M and followed by *-n*, ie the string *-an* but not *-at-an*. This complex instead of being dropped turns into *ā*, although this exception does not always occur (Wright, 1896, vol. II, p. 369).
Generally, also the bound *-t* is dropped, if it is preceded by a short *a*, but sometimes this rule does not apply (Wright, 1896, vol. II, p. 369) .

[3] A DEFAULT value (M, PL etc...) is a value that is automatically assigned to a given linguistic entity (a word, a morpheme etc..) when all the other values are not available, have already been assigned, or cannot be assigned to the linguistic entity at issue for some reason. In other words, the default value is what is left by exclusion of all the other values, and is assigned to the linguistic entity at issue *faute de mieux*.

Table 3.1: The Arabic Noun Phrase: the paradigm #1

INDEFINITE		SG		SOUND PL	
		M	F	M	F
NOM	mu'min-	-u-n	-at-u-n	-ū-na	-āt-u-n
ACC	mu'min-	-a-n	-at-a-n	-ī-na	-āt-i-n
OBL	mu'min-	-i-n	-at-i-n	-ī-na	-āt-i-n
DEFINITE (ARTICLE)		SG		SOUND PL	
		M	F	M	F
NOM	l-mu'min-	-u	-at-u	-ū-na	-āt-u
ACC	l-mu'min-	-a	-at-a	-ī-na	-āt-i
OBL	l-mu'min-	-i	-at-i	-ī-na	-āt-i
DEFINITE (GENITIVE PHRASE)		SG		SOUND PL	
		M	F	M	F
NOM	mu'min-	-u l-qaryati	-at-u l-qaryati	-ū l-qaryati	-āt-u l-qaryati
ACC	mu'min-	-a l-qaryati	-at-a l-qaryati	-ī l-qaryati	-āt-i l-qaryati
OBL	mu'min-	-i l-qaryati	-at-i l-qaryati	-ī l-qaryati	-āt-i l-qaryati

Notice also that the dropped *-t* is retained in writing as *-h*[4].
All the other case endings, as well as the phonological material following them (*-na, -ni*) are not dropped (Owens, 2006, p. 134). This is illustrated in the Tables 3.4, 3.5, 3.6.

What has been dealt with so far is the Classical Arabic NP in its most general manifestation, which is often designated as TRIPTOTIC, ie having a three-fold declension - a term somewhat misleading, because the SOUND PL and DU in both the M, F actually displays only *two* case-endings.
Other minor instances of declension (the so-called DIPTOTIC and MONOPTOTIC NPs) will be illustrated when relevant to the discussion.
Turning to the Colloquial varieties, it is convenient to start with their present-day realizations, because they have been extensively studied in the traditional paradigm (ie within the Old-New dichotomy framework), which is not always the case for their archaic counterparts.
In the modern Colloquial varieties, the shape of the NP is conditioned by dialectal variation: nevertheless, all the modern Colloquial Varieties do share some features, as discussed in Ferguson (1959a), Fischer & Jastrow (1980), Retsoe (1995) and in Abboud-Haggar (2006).
First, they do not exhibit case and their caseless forms are almost identical with the pausal forms of Classical Arabic.
Second, their expression of gender in the F corresponds to the pausal forms of Classical Arabic and, in addition to that, they may drop the *a* of *-at* under certain phonological conditions.
Third, as far as number is concerned, they generally (but not always) preserve the cognate of *ū / īna*, ie *īn* in the construct state and, by the same token, they preserve in this state the *t* of

[4]An account along these lines implies that in the bound *-t -at* is the basic form and *-a(h)* the derived one at some level of representation. The very term *tā' marbūṭah* (= bound *-t*) points to such a state of affairs in synchrony, since as the descriptive term itself suggests OA native speakers conceive this morpheme as a manifestation of *-t*. The same holds for diachrony: following Retsoe (1995) *-at* is likely to be older than its variant *-a(h)*, because the former but not the latter is documented in Akkadian. See Section 7.9.3 for more discussion.

Table 3.2: The Arabic Noun Phrase: the paradigm #2

INDEFINITE		SG		DU	
		M	F	M	F
NOM	mu'min-	-u-n	-at-u-n	-ā-ni	-at-ā-ni
ACC	mu'min-	-a-n	-at-a-n	-ay-ni	-at-ay-ni
OBL	mu'min-	-i-n	-at-i-n	-ay-ni	-at-ay-ni
DEFINITE (ARTICLE)		SG		DU	
		M	F	M	F
NOM	l-mu'min-	-u	-at-u	-ā-ni	-at-ā-ni
ACC	l-mu'min-	-a	-at-a	-ay-ni	-at-ay-ni
OBL	l-mu'min-	-i	-at-i	-ay-ni	-at-ay-ni
DEFINITE (GENITIVE PHRASE)		SG		DU	
		M	F	M	F
NOM	mu'min-	-u l-qaryati	-at-u l-qaryati	-ā l-qaryati	-at-ā l-qaryati
ACC	mu'min-	-a l-qaryati	-at-a l-qaryati	-ay l-qaryati	-at-ay l-qaryati
OBL	mu'min-	-i l-qaryati	-at-i l-qaryati	-ay l-qaryati	-at-ay l-qaryati

-at.

Fourth, following Fischer & Jastrow (1980, p. 89) and Brustad (2000, pp. 45-52) they tend to restrict the applicability of the DU to human or kin terms, or to units of time and, in addition to that, the DU-like ending can designate also a number that range from two to two dozen (the so-called PSEUDO-DU). Remarkably, such an ending may have a distinctive form and be attached to a noun having a specialized meaning.

Fifth, the NP can be combined with a complement of specification in a way highly reminiscent of Indo-european languages: the former is associated with the latter via an intervening preposition-like element and, when definite, it displays the article *l-*, as much as it happens in the English construction *the X of...* .

Notice, however, that the preposition-like element (e.g. *bitā'*, *dyal*) differs from the Indo-european preposition in that it may carry number and gender.

Finally, Bedouin dialects may take the suffix *-in* in the indefinite state SG, if some semantico-syntactic conditions are met, while Sedentary dialects do not.

These properties of the modern Colloquial Arabic NP are illustrated in the Tables 3.7, 3.8, 3.9, 3.10, 3.11 which set forth data from Ingham (1994), Retsoe (1995), Woidich & Heinen-Nasr (2004), Ouhalla (2010).

This traditional presentation does not do justice to more recent findings: the fact that in Najdi Arabic, perhaps the Bedouin dialect *par excellence*, the PL ending is *-ān* rather than *-īn* (Ingham, 2006, p. 329) is but a sample.

This phenomenon as well as other particular properties of the modern Colloquial Arabic NP will be discussed in the Chapters 6, 7.

We are thus left with the old Colloquial varieties, that is with the NP of Old Arabic.

This has been a poorly understood phenomenon until Owens (2006), which basically claims that the Old Arabic NP is to be identified with the pausal / caseless forms of the Classical Arabic NP. Textual research on the most ancient Koranic readings and on Sībawayhi's *al-Kitāb*,

Table 3.3: The Arabic Noun Phrase: the paradigm #3

INDEFINITE	SG		BROKEN PL	
	M / F		M / F	
NOM	ṭarīq-	-u-n	ṭuruq-	-u-n
ACC	ṭarīq-	-a-n	ṭuruq-	-a-n
OBL	ṭarīq-	-i-n	ṭuruq-	-i-n
DEFINITE (ARTICLE)	SG		BROKEN PL	
	M / F		M / F	
NOM	l-ṭarīq-	-u	l-ṭuruq-	-u
ACC	l-ṭarīq-	-a	l-ṭuruq-	-a
OBL	l-ṭarīq-	-i	l-ṭuruq-	-i
DEFINITE (GENITIVE PHRASE)	SG		BROKEN PL	
	M / F		M / F	
NOM	ṭarīq-	-u l-qaryati	ṭuruq-	-u l-qaryati
ACC	ṭarīq-	-a l-qaryati	ṭuruq-	-a l-qaryati
OBL	ṭarīq-	-i l-qaryati	ṭuruq-	-i l-qaryati

in fact, shows that these latter are older than the non-pausal / declensed forms, and had in Old Arabic a general use, not restricted to the pause.
In particular, in these sources the case-distinction between the pausal forms 0 / -$\bar{a}$ (sometimes) and $\bar{u}$ / $\bar{i}na$ (generally) is neutralized, to the effect that the Old Arabic NP becomes almost identical with its modern Colloquial counterpart. These phenomena will be taken up in detail in the Chapters 6, 7.

3.2.2 The modification-structure

In Arabic (regardless of the variety considered) there are four modifiers:

- adjective
- apposition
- relative clause
- genitive phrase (Construct state or Free State)

which generally follow the modified noun[5] they refer to.
This general syntactic pattern was already known in its essence to Arab Grammarians, which subsumed three of the just listed four modifiers under the label TABA', ie 'what follows (the noun)' the modified noun being tellingly designated as MATBŪ', ie 'what is followed (by an adjective, an apposition, a relative clause)' (Wright, 1896, pp. 272-287)[6]:

[5]What is actually modified can be either a noun or a Noun Phrase, since a noun associated with its modifier can be in its turn combined with another modifier, and so on. For the sake of simplicity, however, only the expression '(modified) noun ' will be used in the following.

[6]Recall from Section 3.2.1 that the noun modified by a genitive phrase is designated by Arab Grammarians as MUḌĀF and the genitive phrase itself as MUḌĀF 'ILAY-HI. On a broad reading, this terminology refers not only to the Construct State, but also to the Free State, given that on their view the former is a transformation (*far'*) of the latter (*'aṣl*): see Sections 1.1.3 and 2.3.3.

Table 3.4: The Arabic Noun Phrase: the paradigm #4

INDEFINITE		SG		SOUND PL	
		M	F	M	F
NOM	mu'min-		-ah	-ū-na	-āt
ACC	mu'min-	(-an)	-ah	-ī-na	-āt
OBL	mu'min-		-ah	-ī-na	-āt
DEFINITE (ARTICLE)		SG		SOUND PL	
		M	F	M	F
NOM	l-mu'min-		-ah	-ū-na	-āt
ACC	l-mu'min-	(-ā)	-ah	-ī-na	-āt
OBL	l-mu'min-		-ah	-ī-na	-āt
DEFINITE (GENITIVE PHRASE)		SG		SOUND PL	
		M	F	M	F
NOM	mu'min-	-u l-qaryati	-at-u l-qaryati	-ū l-qaryati	-āt-u l-qaryati
ACC	mu'min-	-a l-qaryati	-at-a l-qaryati	-ī l-qaryati	-āt-i l-qaryati
OBL	mu'min-	-i l-qaryati	-at-i l-qaryati	-ī l-qaryati	-āt-i l-qaryati

(40) Western grammar: Noun + Modifier
Arab Grammarians: matbū‘ + taba‘

3.2.2.1 Adjectives and appositions

The Classical Arabic adjective (or ṢIFAH) is traditionally said (Wright, 1896, vol. II p. 273) to be modelled after the state of the noun it modifies. Thus, the *tanwīn* or the prenominal article found on the noun are 'copied' onto the adjective (in passing, the prenominal article, when 'copied' onto the adjective (42) is sometimes designated in the literature as ADJECTIVAL ARTICLE):

(41) Classical Arabic (Wright, 1896, p. 273)

rajul-u-n karīm-u-n
man-NOM-*n* noble-NOM-*n*

'a noble man'

(42) Classical Arabic (Wright, 1896, p. 273)

l-rajul-u l-karīm-u
the-man-NOM the-noble-NOM

'the noble man'

This results in an opposition definite / indefinite that will be referred to here as definiteness-based distinction or Definiteness Hypothesis.
Modern Colloquial Varieties are taken to deviate from this pattern, so the argument goes, because of the loss of the *tanwīn*. There obtain two constructions, the former found in the Bedouin varieties:

(43) Najdi Arabic (Ingham, 1994, p. 51)

Table 3.5: The Arabic Noun Phrase: the paradigm #5

INDEFINITE		SG		SOUND PL	
		M	F	M	F
NOM	mu'min-		-ah	-ā-ni	-at-ā-ni
ACC	mu'min-	(-ā)	-ah	-ay-ni	-at-ay-ni
OBL	mu'min-		-ah	-ay-ni	-at-ay-ni
DEFINITE (ARTICLE)		SG		SOUND PL	
		M	F	M	F
NOM	l-mu'min-		-ah	-ā-ni	-at-ā-ni
ACC	l-mu'min-		-ah	-ay-ni	-at-ay-ni
OBL	l-mu'min-		-ah	-ay-ni	-at-ay-ni
DEFINITE (GENITIVE PHRASE)		SG		SOUND PL	
		M	F	M	F
NOM	mu'min-	-u l-qaryati	-at-u l-qaryati	-ā l-qaryati	-at-ā l-qaryati
ACC	mu'min-	-a l-qaryati	-at-a l-qaryati	-ay l-qaryati	-at-ay l-qaryati
OBL	mu'min-	-i l-qaryati	-at-i l-qaryati	-ay l-qaryati	-at-ay l-qaryati

bēt-in ṭuwīl
house-*n* tall

'a tall house'

(44) Najdi Arabic (Ingham, 1994, p. 51)

al-bēt al-ṭuwīl
the-house the-tall

'the tall house'

the latter in the Sedentary Varieties:

(45) Egyptian Arabic (Woidich & Heinen-Nasr, 2004, p. 86)

walad šāṭir
boy clever

'a clever boy'

(46) Egyptian Arabic (Woidich & Heinen-Nasr, 2004, p. 86)

l-walad l-šāṭir
the-boy the-clever

'the clever boy'

In some Sedentary Varieties (Levantine Arabic) the article occurs on the adjective but non on the noun:

(47) Levantine Arabic (Pat-el, 2009, p. 33)

sū' əl-'atī'
market the-old

Table 3.6: The Arabic Noun Phrase: the paradigm #6

INDEFINITE	SG		BROKEN PL	
	M / F		M / F	
NOM	ṭarīq-		ṭuruq-	
ACC	ṭarīq-	(-ā)	ṭuruq-	(-ā)
OBL	ṭarīq-		ṭuruq-	
DEFINITE (ARTICLE)	SG		BROKEN PL	
	M / F		M / F	
NOM	l-ṭarīq-		l-ṭuruq-	
ACC	l-ṭarīq-		l-ṭuruq-	
OBL	l-ṭarīq-		l-ṭuruq-	
DEFINITE (GENITIVE PHRASE)	SG		BROKEN PL	
	M / F		M / F	
NOM	ṭarīq-	-u l-qaryati	ṭuruq-	-u l-qaryati
ACC	ṭarīq-	-a l-qaryati	ṭuruq-	-a l-qaryati
OBL	ṭarīq-	-i l-qaryati	ṭuruq-	-i l-qaryati

Table 3.7: The Colloquial Arabic Noun Phrase: the paradigm #7

INDEFINITE		SG		SOUND PL	
		M	F	M	F
EGYPTIAN AR.	mudarris-		-ah	-īn	-āt
DEFINITE (ARTICLE)		SG		SOUND PL	
		M	F	M	F
EGYPTIAN AR.	l-mudarris-		-ah	-īn	-āt
DEFINITE (GENITIVE PHRASE)		SG		SOUND PL	
		M	F	M	F
EGYPTIAN AR.	l-mudarris-	bitā‘ Faṭmah	-(i)t bitā‘it Faṭmah	-īn bitū‘ Faṭmah	-āt bitū‘ Faṭmah

‘the old market’

As for Old Arabic, it behaves like Najdi Arabic in the indefinite state (Owens, 2006, pp. 102-106 and Ch. 4), basically because the pausal forms of Classical Arabic - described in Wright (1896, vol. II, pp. 368-372) - have a generalized usage in Old Arabic.
In the definite state, Old Arabic behaves like Levantine Arabic, following Feghali (1928, p. 125) [7]:

(48) Old Arabic (based on Wright 1896, vol. II, p. 273)

rajul-u-n karīm
man-NOM-*n* noble

‘a noble man’

[7] Feghali (1928)’s description has been refined and given a theoretical account by Pennacchietti (1968), Testen (1998) and Pat-el (2009): see Section 3.3.3 for details.

Table 3.8: The Colloquial Arabic Noun Phrase: the paradigm #8

INDEFINITE		SG		SOUND PL	
		M	F	M	F
EGYPTIAN AR.	mudarris-		-ah	-ēn	-it-ēn
DEFINITE (ARTICLE)		SG		SOUND PL	
		M	F	M	F
EGYPTIAN AR.	l-mudarris-		-ah	-ēn	-it-ēn
DEFINITE (GENITIVE PHRASE)		SG		SOUND PL	
		M	F	M	F
EGYPTIAN AR.	l-mudarris-	bitā‘ Faṭmah	-(i)t bitā‘it Faṭmah	-ēn bitū‘ Faṭmah	-itēn bitū‘ Faṭmah

Table 3.9: The Colloquial Arabic Noun Phrase: the paradigm #9

INDEFINITE	SG	BROKEN PL
	M / F	M / F
MOROCCAN AR.	ktab	ktub
DEFINITE (ARTICLE)	SG	BROKEN PL
	M / F	M / F
MOROCCAN AR.	le-ktab	le-ktub
DEFINITE (GENITIVE PHRASE)	SG	BROKEN PL
	M / F	M / F
MOROCCAN AR.	le-ktab dyal Nadia	le-ktub dyalu Nadia

(49) Old Arabic (Wright, 1896, vol. II, p. 232)[8]

bāb-u l-ṣaġīr
gate-NOM the-little

‘the little gate’

In Old Arabic, however, there occurs also another construction, which is characterized on the level of the form by the lack of both *l-* and *-n* and on the level of the meaning by an ambiguous definite / indefinite meaning, which appears to put into crisis the Definiteness Hypothesis:

(50) Old Arabic (*al-Muzhir fī ‘ulūm al-luġah wa-’anwā‘i-hā*, vol. II, *ḏikr al-’alfāẓ llatī stu‘milat ma‘rifatan lā tadxulu-hā l-’alif wa-l-lām wa-‘aksu-h*, p. 157)

wa-hāḏā ’usam-at-u ‘ādiy-an
and-this the.lion-*at*-NOM coming.back-ACC

‘and this is the lion, which is coming back’

(51) Old Arabic (*Šarḥ al-Mufaṣṣal*, vol. I, *Faṣl ’ism al-jins al-muxtaṣṣ bi-l-ḥayawān*, p. 113)

hāḏā ’usam-at-u maqbūl-an
this the.lion-*at*-NOM approaching-ACC

[8]The example is quoted in pausal form, along the lines of Wright (1896, vol. II, p. 233) and of Owens (2006, pp. 102-106 and Ch. 4). It should be added for the sake of completeness that according to Arab Grammarians the modifier is marked for OBL as if it were a complement of specification: this issue will be taken up in Chapter 6 (Section 7.3).

Table 3.10: The Colloquial Arabic Noun Phrase: the paradigm #10

INDEFINITE	SG		BROKEN PL	
	M / F		M / F	
SEDENTARY (EGYPTIAN AR.)	kitāb-		kotob-	
BEDOUIN (NAJDI AR.)	kitāb-	-in	kotob-	-in

Table 3.11: The Colloquial Arabic Noun Phrase: the paradigm #11

(IN)DEFINITE	SG		DUAL	
	SG		DU	
PROPER DU (MOROCCAN AR., TLEMCEN)	dra‘-		dər‘-	-ayn
	‘arm, fathom’		‘two fathoms’	
PSEUDO-DU (MOROCCAN AR., TLEMCEN)	dra‘-		dər‘-	-īn
	‘arm, fathom’		‘arms’	

‘this is a lion, which is approaching[9]’

The translation proposed in (50, 51) faithfully reproduces Arab Grammarians’ and Lexicographers’ positions, who assign both the definite / indefinite meaning to the noun *’usamah* ‘lion’ when followed by an adjectival modifier[10].
On one side, in fact, the Late Grammarian (and compiler) Al-Suyūṭī (d. 911 / 1505) reporting data from the Lexicographers Ibn al-Sikkīt (d. 246 / 860) and Al-Tibrīzī (d. 502 / 1109), devotes a section of his work *al-Muzhir fī ‘ulūm al-luġah wa-’anwā‘i-hā* to ‘the expressions that were used as definite in spite of the lack of the article *l-*’ (*(ḏikru) l-’alfāẓi llatī-stu‘milat ma‘rifatan lā tadxulu-hā l-’alifu wa-l-lām*) and states that:

> “...Ibn al-Sikkīt collected these [expressions] into a chapter of his work *’Iṣlāḥ [al-manṭiq]* and so did Al-Tibrīzī in his *Tahḏīb [kitāb al-’alfāẓ li-Ibn al-Sikkīt]*. They say: *hāḏihi ḏukā’u ṭāli‘ah* [= this is the rising sun], where [*ḏukā’*] is one of the names of the sun and it is definite (*ma‘rifah*); *hāḏa ’usam-at-u ‘ādiy-an* [= this is the lion, which is coming back], where [*’usamah*] is one of the names of the lion and it is definite (*ma‘rifah*). These are the expressions they report, although I did not mention all of them, for they are legions...”.
> (*al-Muzhir fī ‘ulūm al-luġah wa-’anwā‘i-hā*, vol. II, *ḏikr al-’alfāẓ llatī stu‘milat ma‘rifatan lā tadxulu-hā l-’alif wa-l-lām wa-‘aksu-h*, p. 157)[11]

[9]The English translation proposed here is based on Fleisch (1961, vol. I, p. 346)’s French rendering of the same expression: ‘c’est un lion qui s’avance’.

[10]The adjectival modifier at issue has two relevant properties. First, it does not agree in case with the noun it refers to. Second, it is translated as an intransitive participle. As far as the former property is concerned, it is taken to be negligible here, because in Old Arabic the ACC case is commutable with the NOM case, the different realization being a matter of dialectal variation rather than of meaning, as discussed at the end of Section 2.2.2.4 (see in particular the issue of the so-called *mā hijāzī*, ie the two different manifestations of the negative construction (30, 31)). It is worth noting in this connection that Fleisch (1961, vol. I, p. 346) glosses over the discrepancy between case-endings and translates (51) as a standard modification-structures (see fn. 9). For the latter property, see fn. 45.

[11]‘*...‘aqada la-hā bnu l-sikkīti fī l-’iṣlāḥi wa-l-tibrīziyyu fī tahḏībi-hi bāban qālā fī-hi [...] wa-hāḏihi ḏukā’u ṭāli‘ah ’ismun li-l-šamsi wa-hya ma‘rifah wa-hāḏā ’usam-at-u ‘ādiy-an ’ismun li-l-’asadi wa-hwa ma‘rifah wa-hāḏā mā ḏakarā-hu wa baqiyat ziyādatun ‘alà ḏalika...’*

On the other side, the Late Grammarian 'Ibn Ya'īš (d. 643 / 1245) states in the following passage, summarized in Fleisch (1961, vol. I, pp. 345-346), that similar expressions are definite in form but indefinite in meaning:

> "...what demonstrates the definiteness (*ta'rīf*) of these entities is the fact that there occurs after them an indefinite [adjective] (*nakirah*) which has a circumstantial meaning (*ḥāl*) and in fact you say *hāḏā 'usam-at-u maqbūl-ā* [= this is a lion, which is approaching] as well as *ra'aytu ṯu'ālata muwalliyā* [= I saw a fox, while it was getting close]: if these entities had been indefinite, they would have not been followed by a modifier having a circumstantial meaning [because as a rule this kind of modifier always require a definite noun]. Know that an entity of this kind is definite, as we have just said, but its definiteness is only a formal marking (*'amrun lafẓiyyun*) because from the standpoint of meaning (*min jihati l-ma'nà*) it is indefinite (*nakirah*) given that it has scope (*šiyā'*) over any [individual] of the species and does not define an individual in particular (*'ixtiṣāṣ*) to the exclusion of another...".
> (*Šarḥ al-Mufaṣṣal*, vol. I, *Faṣl 'ism al-jins al-muxtaṣṣ bi-l-ḥayawān*, p. 113)[12]

The same ambiguous behavior is displayed by the morphemic complex *wāḥəd l-* that in the Western Colloquial Varieties works (to a certain extent) like the Indo-european indefinite article, already sketchily introduced in the Section 2.2.1.2 and in the Table 2.2.
This complex, in fact, is a combination of the numeral *wāḥəd* 'one' and of the article *l-*, regarded by the received view as the marker of definiteness *par excellence* but used in this case in a totally opposite context. Moroccan Arabic well exemplifies such a state of affairs:

(52) Moroccan Arabic (Edzard, 2006, p. 189)

wāḥd əl-mra
one the(??)-woman

'a woman'

The careful reader will have noticed that the modern Colloquial (52) not only shares with Old Arabic (50, 51) the ambiguity between definiteness / indefiniteness, but also exhibits another anomalous property, namely the fact that the modifier seems to precede the modified noun. This raises a counterexample to the general pattern (40) illustrated in Section 3.2.2.
The discussion of the former (apparent) anomaly, which challenges the Definiteness Hypothesis, is postponed to Section 3.2.2.2 for reasons that will become clear there, while the latter one is dealt with in greater detail in the following, starting with Classical Arabic.
In this language, the following (apparent) counterexample to (40) is found:

(53) Classical Arabic (Wright, 1896, vol. II, p. 229)

'awwal-u yawm-in
first-NOM day-OBL

[12] '*...wa mimmā yadullu 'alà ta'rīfi hāḏihi l-'ašyā'i 'anna-hu yaqa'u ba'da-hā l-nakiratu ḥālan ka-qawli-ka hāḏā 'usam-at-u maqbūl-ā wa-ra'aytu ṯu'ālata muwalliyā wa-law kānat nakirātin lam yaqa' l-ḥālu ba'da-hā wa-'lam 'anna hāḏihi l-'ašyā'a ma'arifu 'alà mā ḏakarnā 'illā 'anna ta'rīfa-hā 'amrun lafẓiyyun wa-hya min jihati l-ma'nà nakirātun li-šiyā'i-hā fī kulli wāḥidin mina l-jinsi wa-'adami-xtiṣāṣi-hā šaxṣan bi-'ayni-hi dūna ġayri-hi...*'

'the first day'

Constructions like (53) are explained by Arab Grammarians and Western Arabists (Wright, 1896, vol. II, p. 229) as instances of Construct State, where the adjective is to be seen as a NP or as substantivized adjective, like the numeral in (53), and the NP following the adjective as a complement of specification, as much as the English expression *the best of....*
On this view, the adjective turns out to be the modifiee, and the NP the modifier, in compliance with the general behavior of the Arabic modification-structure.
This interpretation is supported by three facts. First, the morphological pattern of the adjective is typical either of a Noun or of an adjective tending to undergo substantivization, given that *'awwal-*, for example, has the pattern 'aCCaC usually denoting in Arabic a superlative (cp. *'afḍal*, '(the) best (of)'.
Second, the NP following *'awwal-* is marked by a bare OBL, which is the case-ending that in Arabic typically signals the complement of specification.
Third, the entire expression is said to be equivalent to:

(54) Classical Arabic (Wright, 1896, vol. II, p. 229)

l-yawm-u l-'awwal-u
the-day-NOM the-first-NOM

'the first day'

The apparently exceptional typology includes not only (positive and superlative) adjectives, but also numerals (67).
This is hardly surprising in the light of the fact that Old and Classical Arabic numerals are regarded by the Arab Grammarians as nouns, and adjectives, according to them, do not have a categorial status, being rather an intermediate category between a noun and a verb (Guillaume, 1992) - thus, when following the NP they refer to, these modifiers can be understood as appositions (BADAL) (Fleisch, 1961, vol. I, p. 371):

(55) Noun + Adjective_{Noun} = Noun + Apposition

(56) Noun + Numeral_{Noun} = Noun + Apposition

Therefore, Arab Grammarians' analysis points to an appositional origin of the adjective in diachrony, an hypothesis that according to Fleisch (1961, pp. 370-371) is supported by three facts, also reported by Fischer (2006a, pp. 17, 19)[13]:

- morphologically, in Classical and Old Arabic it is indistinguishable from the noun
- semantically, in Old Arabic its nominal nature is in some cases still transparent
- syntactically, in Old Arabic it is not agreed with the modified noun

The first phenomenon can be exemplified as follows:

[13]The same position is already hinted at in Wright (1896, vol. II, p. 229), which recognizes that the appositional modifier of (58) is the 'older construction' but takes this phenomenon to be restricted to the adjective denoting materials and in fact dubs it 'apposition of material'. Nevertheless, examples like (59) and (97, 98) plausibly demonstrate that the phenomenon is general.

(57) Old / Classical Arabic (Fleisch, 1961, vol. I, p. 370)

CaCāC-$_i$	'anān-$_i$	jabān-$_i$
Morphological pattern	donkey.F	lazy

The following pieces of data illustrates the second phenomenon[14]:

(58) Old Arabic (Fleisch, 1961, vol. I, p. 371)

l-ṣanam-u	l-ḏahab
the-idol-NOM	the-gold

'the golden (lit. gold) idol'

(59) Old Arabic (Fleisch, 1961, vol. I, p. 371)

naḥnu	ḥarb-u-n	li-man	ḥāraba	wa-silm-u-n	li-man	sālam
we	war-NOM-*n*	to-who	made.war	and-peace-NOM-*n*	to-who	lives.in.peace

'we are enemy (lit. war) to who makes war on us and friends (lit. peace) to who lives in peace with us'

As for the third phenomenon, it is illustrated in (97), (98) below.
On these grounds, the adjectival modification is realigned with more familiar instances of appositional modification, reported in the traditional grammar of Classical Arabic:

(60) Classical Arabic (Wright, 1896, vol. II, p. 286)

jā'a-nī	'ax-ū-ka	Zayd
came-me	brother-NOM-you	Zayd

'your brother Zayd came to me'

The demonstrative as well usually (but not always) precedes the noun, as illustrated in (65, 66). According to Arab Grammarians, this is because semantically the modified noun (MATBŪ') is the demonstrative, not the noun, this latter rather acting as a modifier (TABA'), that is, in modern terms, as an apposition (Wright, 1896, vol. II, p. 277)[15]:

(61) Demonstrative$_{Noun}$ + Noun = Noun + Apposition

The just sketched account of the (alleged) modifiers preceding the noun holds also for the modern Colloquial varieties and Old Arabic, since they display similar behaviors, as shown in (63 - 67) and discussed in Jones (2005, Ch. 6, Ch. 14). In particular, considering again

[14] An alternative interpretation of (59) would regard *ḥarbun* and *silmun* not as nouns functioning as adjectives, but as nouns linked with *naḥnu* by means of a metaphor (Antonella Ghersetti, pc).

[15] The construction:

(1) Noun + Demonstrative$_{Noun}$

does occur in Classical and Old Arabic, but it preferably applies when the noun is a proper name. In other cases the construction 'very rarely' occurs: see Wright (1896, vol. II, p. 277) for Classical Arabic and Rabin (1951, p. 76) for Old Arabic. As for the modern Colloquial Varieties, the just mentioned construction is grammatical (see (64)), but generally native speakers liken the postnominal demonstrative to the so-called REINFORCER, that is to emphatic particles like *here* in the sub-standard English expressions like *this book here* and *-ci* in French expressions like *ce-livre ci*. See e.g. Ihsane (2003, p. 273) and Harrell et al. (1965, p. 71) who explicitly state this for Moroccan Arabic.

the indefinite construction featuring the modifier *wāḥəd* 'one' typical of the Western Colloquial Varieties and exemplified in (52), Versteegh (1997a, p. 164) and Edzard (2006, p. 189) after him observe that it is modelled after the demonstrative construction. On these grounds, *wāḥəd* is better seen along the lines of (61) as a modifiee rather than as a modifier, and the noun associated with *l-* following it the other way around:

(62) wāḥəd$_{Noun}$ + Noun = Noun + Apposition

This results in a generalized pattern Noun + Modifier / Head + Dependent summarized in the Table 3.12.

Moreover, evidence culled from the modern Colloquial varieties (pending further research for

Table 3.12: The Arabic NP: the unified structural pattern

Noun^x	Modifier^y
Head^x	Dependent^y
MATBŪ'^{x}	TABA'^{y}
Noun	Adjective$_{Noun}$
Noun	Noun
Noun	Num$_{Noun}$
wāḥəd$_{Noun}$	Noun
Dem$_{Noun}$	Noun
(Noun	Dem$_{Noun}$)
MUḌĀF^x	MUḌĀF 'ILAY-HI^y
Adjective$_{Noun}$	Noun
Num$_{Noun}$	Noun

Classical and Old Arabic) plausibly shows that in Arabic the relative ordering of the modifiers, when postnominal, is the mirror-image of English (Hetzron, 1978; Fassi-Fehri, 1993, 1999, 2000; Cinque, 2003; Ihsane, 2003; Shlonsky, 2004): this is exemplified in (64).

Fassi Fehri claims that Modern Standard Arabic behaves like the Colloquial Arabic varieties (64) as far as the relative order of the modifiers is concerned (Fassi-Fehri, 1999, 2000). Although Fassi Fehri's material is not immune from criticism, it is reproduced in the following (65 - 67) for the sake of illustration of the Old and Classical Arabic modification-structure, since it well exemplifies the just discussed different word-orders of the numeral and the demonstrative relative to the noun:

(63) Moroccan Arabic (Ihsane, 2003, p. 271)

had l-wəld
this the-boy

'this boy'

(64) Moroccan Arabic (Ihsane, 2003, p. 269)

lə-wlad z-zwiin-in l-xams-a hadu
the-boys the-handsome-PL the-five-*at* these

'these five handsome boys'

(65) Modern Standard Arabic (Fassi-Fehri, 1999, p. 149)

l-ṣuḥuf-u l-jadīd-at-u l-ṯalāṯ-u hāḏīhi
the-newspapers-NOM the-new-*at*-NOM the-threeNOM this-F

'these three new newspapers'

(66) Modern Standard Arabic (Fassi-Fehri, 1999, p. 150)

hāḏīhi l-ṣuḥuf-i l-jadīd-at-u l-ṯalāṯ-u
this-F the-newspapers-NOM the-new-*at*-NOM the-threeNOM

'these three new newspapers'

(67) Modern Standard Arabic (Fassi-Fehri, 2000, p. 86)

hāḏīhi l-ṯalāṯ-u l-ṣuḥuf-i l-jadīd-at-u
this-F the-threeNOM the-newspapers-GEN the-new-*at*-NOM

'these three new newspapers'

The typology of the article and of the *tanwīn* discussed so far are summarized in the Tables 3.13, 3.14, where Classical Arabic is shortened into CA, Old Arabic into OA and Modern Colloquial Varieties into MCV[16].

All these examples show that in the High Variety, Low Variety and Old Arabic the adjective

Table 3.13: *l-* in the Arabic NP: typology

	l Noun l Adj	Noun l Adj	Noun Adj
CA	X		
OA	X	X	X
MCV	X	X	X

Table 3.14: *-n* in the Arabic NP: typology

	Noun n Adj n	Noun n Adj	Noun Adj
CA	X	X	
OA		X	X
MCV		X	X

follows the noun, thus complying with the general pattern (40) stated in the previous Section. Another aspect of the adjectival / appositional modification in need of discussion is agreement. This is traditionally said to take place between the noun and the adjective by 'copying' the adnominal markers of case, number, gender, definiteness *-u(:), -i(:), -a(:), -ay, -n, l-, -t, -na, -ni* onto the adjective: Wright (1896, vol. II, p. 273) is illustrative of such a kind of explanation, which at least in part goes back to Arab Grammarians, as discussed in Kahle (1975, pp. 10-12) (see also pages 120, 121, 123, 123 below). Borrowing a term (admittedly anachronistically) from the generative literature, such an approach is designated here as Feature Copying Hypothesis.

[16]The Table 3.14 abstracts away from the so-called Diptoptic 'Declension' of Classical and Old Arabic: see Section 3.2.1.

But, as far as the adnominal markers of gender and PL number *-n, -t, -na* are concerned, Classical Arabic, Old Arabic (= old Colloquial varieties) and the modern Colloquial varieties are not identical, as discussed at length in Hashemi (2006b) and Acquaviva (2008, Part II, Ch. 7).

It is well-known that in Classical Arabic and modern Colloquial Varieties animacy (humanness) plays a key-role in the agreement. If the noun denotes a human referent (either M or F), the 'adjective' agrees with the noun in both number and gender (70 - 73), while if the noun denotes a non-human referent, the 'adjective' is given a totally different 'PL' default agreement, consisting of the SING F ending *-at* (bound *-t*), irrespective of the (grammatical) gender of the modified noun (74 - 75)[17].

Notice in this connection that adjectives, very much as the nouns, when pluralized have not only a SOUND PL form, but also a BROKEN PL form. Then, the expression of full agreement can be morphologically realized onto the adjectives in both the forms (70b, 70c, 71b, 71c, 72b, 72c, 73b, 73c): see e.g. Fleisch (1961, vol. I, pp. 288-289) and, more recently, Abu Chacra (2007, p. 145) .

Sometimes, the F PL seems to be regarded not as a proper PL, but as a COLL noun. Heavily simplifying, this latter is a noun which has a SG form, but has a PL meaning, since it refers not to one entity, but to *many* entities that can be still identified as such because they are *separate* (cp. English words like *group, club, fish*).

This can be easily seen contrasting a COLL with a MASS NOUN, that is a noun like English *water, furniture* denoting *many* entities that however are *no* longer felt as *separate*: in a *group*, persons are separate entities while in *water*, drops are not. According to Zabbal (2002) and Acquaviva (2008) this difference can be appreciated by combining a COLL and a MASS NOUN with predicates like *count* or *distinguish*, since the resulting sentence is grammatical in the former case but not in the latter:

(68) a. English (Acquaviva, 2008, p. 226)

He counted the fish one by one

b. Moroccan Arabic (Acquaviva, 2008, p. 226)

ḥseb l-ḥut waḥd-a b wahḥd-a
counted.he the-fish one-*at* by one-*at*

'he counted the fish one by one'

(69) English (Acquaviva, 2008, p. 226)

*He counted the furniture one by one

The matter, however, is more complicated, because the strange-looking discrepancy between the form and the meaning arises from the fact that the PL entities at issue albeit viewed as many and separate, are only *in part identifiable*: their identity, so to speak, gets lost in the group, because the sum is perceptually more salient than its parts. As Acquaviva (2008, p. 229) puts it, in a way highly reminiscent of Enç (1991) (see below), entities of a COLL 'are not individuated enough to fix their identity, for instance, by letting one correspond to *a*, another to *b*, and a third to *c*'.

[17] See fn. 3 for the notion of 'default'.

Once again a contrast well illustrates the point: while the proper PL *cows* refers to entities that can be totally identified, a COLL like *cattle* refers to more anonymous entities that according to Acquaviva (2008, p. 229) 'are better paraphrased as 'cattle[...]-property in multiple instances' rather than as 'cows[...]".
Accordingly, PL, COLL, MASS for the purposes of the present work can be understood along the lines of the Table 3.15[18].

It ensues that if the entities denoted by the COLL noun are humans, they are no longer

Table 3.15: PL, COLL, MASS nouns

	PL	COLL	MASS
separateness	YES	YES	NO
identity	YES	PARTIAL	NO

identifiable as such: accordingly, the COLL noun does not trigger the full agreement on the adjective, but the default agreement in *-at* (Brustad, 2000, p. 52ff.). A familiar example is the English sentence *It is a large club* coexisting with *They are a large club.*
According to the received view, a case in point for Classical Arabic is the above mentioned F PL treated as a COLL noun (see (76, 78)), while the M PL is said not to display this behavior (Kahle, 1975, p. 11), although nothing bars it in principle[19].
The just outlined agreement phenomena are summarized in the Table 3.16 and exemplified in the following (for the sake of simplicity, data are drawn from Classical Arabic only, but they can be taken to be illustrative also of the modern Colloquial Varieties):

(70) HUMAN, SOUND PL, M

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʿallim-u l-mašhūr-u
the-teacher-NOM the-famous-NOM

'the famous teacher (M) '

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʿallim-ū-na l-mašhūr-ū-na
the-teacher-NOM-*na* the-famous-NOM-*na*

[18]Looking at this table, one may object that the values 'Separateness' and 'Identity' are not mutually consistent because the former is discrete, the latter non-discrete. This objection loses most of its strength if one considers that the non-discrete value 'partial identity' expressed in the table will be more accurately recast below as SPECIFICITY along the lines of Enç (1991, p. 7). He in fact shows that this value in formal semantics can be decomposed into two discrete values. For the sake of simplicity, a more informal definition of specificity will be adopted here, but it is important to bear in mind that througout this work it is thought of as a discrete value.

[19]No satisfactory explanation for this phenomenon is available in the literature, as pointed out by Garbini & Durand (1994, p. 97). Some sociologically-oriented explanations, like Fleisch (1961, vol. I, p. 326) - in its turn developing Brockelmann (1908) - have it that F PL nouns are more likely to be regarded as less animate and hence as a COLL and that M PL nouns resist such a treatment, because females were less important than males in the Ancient Semitic civilization and / or in the Medieval Arabo-Islamic civilization. However, one may observe that discrimination between men and women characterized in the Ancient and Middle Ages the Old World as a whole, not only the just mentioned societies and thus one may wonder why in the same material conditions *not even one* of the languages of the Old World other than Arabic (say, Latin or Old French) did not develop the same agreement patterns. Furthermore, in purely linguistic terms, Arabic gender seems to be detached from biological sex, as discussed in Hashemi (2006b) and in the Section 3.3.2.

'the famous teachers (M) '

c. Classical Arabic (based on Abu Chacra 2007, p. 145)

l-mu'allim-ū-na l-su'adā'-u
the-teacher-NOM-*na* the-happies-NOM

'the happy teachers (M) '

(71) HUMAN, SOUND PL, F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mu'allim-at-u l-mašhūr-at-u
the-teacher-*at*-NOM the-famous-*at*-NOM

'the famous teacher (F) '

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mu'allim-āt-u l-mašhūr-āt-u
the-teacher-*āt*-NOM the-famous-*āt*-NOM

'the famous teachers (F) '

c. Classical Arabic (based on Abu Chacra 2007, p. 145)

l-mu'allim-āt-u l-su'adā'-u
the-teacher-*āt*-NOM the-happies-NOM

'the happy teachers (F) '

(72) HUMAN, BROKEN PL, M

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-rajul-u l-mašhūr-u
the-man-NOM the-famous-NOM

'the famous man'

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-rijāl-u l-mašhūr-ū-na
the-men-NOM the-famous-NOM-*na*

'the famous men'

c. Classical Arabic (based on Abu Chacra 2007, p. 145)

l-rijāl-u l-su'adā'-u
the-men-NOM the-happies-NOM

'the happy men (M) '

(73) HUMAN, BROKEN PL, F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mar'-at-u l-mašhūr-at-u
the-woman-*at*-NOM the-famous-*at*-NOM

'the famous woman (F) '

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-nisā'-u l-mašhūr-āt-u
the-women-NOM the-famous-*āt*-NOM

'the famous women'

c. Classical Arabic (based on Abu Chacra 2007, p. 145)

l-nisā'-u l-su'adā'-u
the-women-NOM the-happies-NOM

'the happy women (F) '

(74) NON-HUMAN, SOUND PL, M / F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-ḥammām-u l-naẓīf-u
the-bath-NOM the-clean-NOM

'the clean bath'

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-ḥammām-āt-u l-naẓīf-at-u
the-bath-*āt*-NOM the-clean-*at*-NOM

'the clean baths (F) '

(75) NON-HUMAN, BROKEN PL, M / F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-sūq-u l-naẓīf-u
the-market-NOM the-clean-NOM

'the clean market'

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-'aswāq-u l-naẓīf-at-u
the-markets-NOM the-clean-*at*-NOM

'the clean markets'

(76) HUMAN → 'GROUP OF...'(= NON-HUMAN NOUN), SOUND PL, F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mu'allim-at-u l-mašhūr-at-u
the-teacher-*at*-NOM the-famous-*at*-NOM

'the famous teacher (F) '

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mu'allim-āt-u l-mašhūr-at-u
the-teacher-*āt*-NOM the-famous-*at*-NOM

'the famous teachers (F) '

(77) HUMAN → 'GROUP OF...'(= NON-HUMAN NOUN), SOUND PL, F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mu'allim-at-u l-mašhūr-at-u
the-teacher-*at*-NOM the-famous-*at*-NOM

'the famous teacher (F) '

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mu'allim-āt-u l-mašhūr-at-u
the-teacher-*āt*-NOM the-famous-*at*-NOM

'the famous teachers (F) '

(78) HUMAN → 'GROUP OF...'(= NON-HUMAN NOUN), BROKEN PL, F

a. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mar'-at-u l-mašhūr-at-u
the-woman-*at*-NOM the-famous-*at*-NOM

'the famous woman (F) '

b. Classical Arabic (based on Veccia Vaglieri 1937, p. 83)

l-nisā'-u l-mašhūr-at-u
the-women-NOM the-famous-*at*-NOM

'the famous women'

Table 3.16: CA (and MCV): animacy-driven agreement patterns

				Adjective		
				Strict Agr		Deflected Agr
				SOUND PL	BROKEN PL	
Noun	Human	M	SOUND PL	X	X	
			BROKEN PL	X	X	
		F	SOUND PL	X	X	X
			BROKEN PL	X	X	X
	Non-Human	M / F	SOUND PL			X
			BROKEN PL			X

This is not the whole story, however: according to Ferguson (1989), Belnap & Gee (1994), Acquaviva (2008), in modern Colloquial Varieties there is another factor involved in agreement, namely SPECIFICITY (also referred to in the literature as INDIVIDUATION).
This term denotes a kind of definiteness which, without going into details, is weaker than definiteness proper: both these notions share to more or less degree the property of enabling the hearer to identify the referent denoted by a given noun. As Enç (1991) puts it:

> "Definiteness and specificity of NPs are clearly related phenomena. Both definites and specifics require that their discourse referents be linked to previously established discourse referents, and both indefinites and nonspecifics require that their discourse referents not be linked to previously established discourse referents. What distinguishes these notions is the nature of the linking. The linking relevant for definite NPs is the identity relation. [...] Thus, specificity involves a weaker, looser relation to already established referents than definiteness".
> (Enç, 1991, p. 9)

The semantic solidarity between definite and specific NPs can be reflected in the form. For example, in English both of them can be marked by the same article *the* (*a* being also a viable choice in some contexts), and are opposed as such to the indefinite, which is not.
As for Arabic, specificity can be brought into the picture via (79), which contrasts data from both English and modern Colloquial Arabic (Levantine Arabic):

(79) a. Levantine Arabic (Cowell, 1964, p. 424)

l-kətəb maa b-i-həmm-uu-(h)
the-book.PL NEG PRS-3SG-interest-they-him

'the books don't interest him'

b. Levantine Arabic (Cowell, 1964, p. 424)

l-kətəb maa bə-t-həmm-o(h)
the-book.PL NEG PRS-3F.SG-interest-him

'books don't interest him'

The example shows that in the modern Colloquial varieties, the contrast specific vs. indefinite obtains not through the opposition lack vs. presence of article as English does, but via full agreement vs. default agreement (the verbal ending 3PRS F SG being highly reminiscent of *-at*). Such a state of affairs seems to hold not only for the agreement between the noun and the verb, but also for that between the noun and the adjective, as it will be evident shortly.

To summarize the discussion so far, Classical Arabic only has humanness / animacy-driven agreement, while in the modern Colloquial Varieties this kind of agreement coexists with specificity-driven agreement.

Both the kinds of agreement are triggered by a semantic property (be it humanness / animacy or specificity), that is they are SEMANTICALLY CONDITIONED, and, in addition to that, the high degree of the semantic property involved is expressed by full agreement, while its low degree by default agreement. The former is also indicated in the Arabistic literature as STRICT AGREEMENT, the latter as DEFLECTED AGREEMENT (Ferguson, 1989).

It has been said that in synchrony the two kinds of semantically conditioned agreement are not mutually exclusive and may co-occur, as the modern Colloquial Varieties show. Nevertheless, humanness / animacy-driven agreement and specificity-driven agreement are not on the same footing in them, the crucial variable for assignment of strict or deflected agreement being specificity rather than humanness, a phenomenon observed by Belnap & Shabaneh (1994) and further discussed in Brustad (2000, Ch. 2) and Acquaviva (2008).

In diachrony, specificity-driven agreement seems to be older than humanness / animacy-driven agreement (notice in passing that this is a position defended within the 'orthodox' Old-New dichotomy): Ferguson (1989), Belnap & Shabaneh (1994) and Belnap & Gee (1994) demonstrate, in fact, that the former, but not the latter, is also found in Old Arabic.

Furthermore, Khan (1984)'s account of specificity in Semitic, as well as its further refinements focusing on agreement by Brustad (2000, pp.21-25) and Acquaviva (2008, pp. 213, 219), show that on diachronic level these two kinds of agreement are not unrelated, since along a scale of specificity (INDIVIDUATION HIERARCHY), *humans* appear to be more *individuated* (= *specific*) than non-human beings:

(80) Individuation Hierarchy (Khan, 1984, p. 470)
1st > 2nd > 3rd > Human > Inanimate

A non-trivial consequence of the Individuation Hierarchy is that it seems to constrain application of the specificity-driven agreement. Humans, in fact, occupy a high rank in the scale of specificity and therefore the noun denoting them, when pluralized, cannot be treated as generic referents, so that they trigger strict agreement on the adjective, but not deflected agreement (86, 87).

Notice, however, that the Colloquial Varieties and Old Arabic, differ from Classical Arabic in

the expression of PL nouns: while in the latter both SOUND PL and BROKEN PL are regarded as such, in the former ones the SOUND PL only is unequivocally plural. Hence, it is the SOUND PL that typically triggers strict agreement on the adjective (86, 87).

On the other hand, BROKEN PL has an ambiguous behavior, because it can be thought of either as a real PL, ie as a sum of identifiable entities (see Table 3.15) thus triggering strict agreement on the adjective (88b, 89b and the sentence *They are a large club*), or as a COLL, ie as a sum of *partially* identifiable entities, which enables it to trigger only deflected agreement on the adjective modifying it (88c, 89d and the sentence *It is a large club*). Recall from the discussion surrounding (78a - 79) that technically speaking *partially identifiable* is more properly defined as *specific*.

The double nature of the BROKEN PL in Old Arabic is plausibly due to the fact that the BROKEN PL diachronically develops out of a COLL, the former being no more than the latter having undergone some process of affixation, lengthening, gemination etc..., a view shared by both traditional (Fleisch, 1961, vol. I, pp. 309, 363-376, 470-471) and less traditional (Garbini & Durand, 1994, p. 99) views (see also Section 3.3.2).

The same behavior is still found in the modern Colloquial Varieties, either because of convergent development, as the partisans of the Old-New dichotomy contend (Ferguson, 1989), or because of common descent, as an account à la Garbini & Durand (1994) / Owens (2006) would claim. On the other hand Classical Arabic, being chronologically later than Old Arabic (and perhaps quite artificial), has lost memory of this original state of affairs and then does not follow this particular agreement pattern, as already seen above.

Turning to non-human beings, they are lower than humans in the Individuation Hierarchy (80) and then nouns denoting them trigger deflected agreement on their adjectives (90b, 91b).

But, by virtue of the same low position along the scale of specificity, they are not affected by the constraint that applies to the humans. In fact, they can be also promoted, so to speak, to the rank of humans if they are conceived as particularly individuated (= specific), to the effect that deflected agreement of the adjective is replaced by strict agreement, as shown in (90d, 91d) - cp. the conceptualization of a domestic animal in English, which is specifically referred to as *he* or *she* rather than as an undifferentiated *it*, assigned to animals in general.

Once again, Old Arabic and modern Colloquial Varieties resemble in this respect: Ferrando (2006, p. 55-56) reports the alternation of strict / deflected agreement involving non-human beings for Old Arabic, while Blanc (1970, p. 49) observes the same phenomenon in the modern Colloquial Varieties[20].

Furthermore, both the scholars characterize this alternation in terms of specificity, as expected, since according to Ferrando (2006, pp. 40, 52) and Blanc (1970, pp. 52-53), strict agreement occurs in contexts of small number, which are by definition 'individual, specific' ('enumerative' in Blanc's terms: see Brustad (2000, p. 45, fn. 3) for the identification of the opposition: specific vs. non-specific with that enumerative vs. non-numerative).

Both the scholars adopt the expression 'the big house' to exemplify the alternation under discussion:

[20]Notice that Blanc (1970) updates Ferguson (1959a)'s description of the Low Variety facts, where only the deflected agreement (82a) is reported. By the same token, Ferrando (2006) improves Fleisch (1961, vol. I, p. 309), which mentions for Old Arabic only the strict agreement (81b). Further data on Old Arabic strict agreement in Fischer (2006a, p. 19 and references therein). Notice also that the examples cited immediately below inevitably involve a certain amount of idealization.

(81) a. Old Arabic (Ferrando, 2006, p. 56)

buyūt kabīr-ah
houses big-*at*

'big houses'

b. Old Arabic (Ferrando, 2006, p. 56)

buyūt kibār
houses bigs

'big houses'

(82) a. Modern Colloquial Varieties (Blanc, 1970, p. 49)

biyūt kibīr-a
houses big-*at*

'big houses'

b. Modern Colloquial Varieties (Blanc, 1970, p. 49)

biyūt kuḍāṛ
houses bigs

'big houses'

It could be noted at this point that something is missing in Ferrando (2006)'s and Blanc (1970)'s account. Recall, in fact, that in purely morphological terms the deflected agreement triggered by the noun on the adjective has just one realization, ie *-at*, but the strict agreement triggered by the noun on the adjective has *two* possible realizations: BROKEN PL and SOUND PL. Yet, as the just mentioned examples show, Ferrando and Blanc do not mention the strict agreement via SOUND PL on the adjective.

In addition to that, the strict agreement via BROKEN PL on the adjective they discuss requires a closer look, since it has been seen above that the BROKEN PL, when found on the noun, has an ambiguous reading, working either as a real PL or as a COLL. Then, the same behavior is expected for the adjective (*a fortiori* if this latter is itself a noun, as discussed at the beginning of this Section).

Considering the issue of the BROKEN PL first, the expectation is met: on one side, in fact, Ferrando (2006, pp. 45-47) remarks that morphological patterns like CiCāC, used in Old Arabic for nouns and adjectives, already in Sībawayhi's *al-Kitāb* are taken not only to denote a small number, ie specific entities (see above), but also to denote a large number, which Ferrando (2006, p. 40) qualifies as 'collective, non-differential' ; on the other side in Blanc (1970, p. 50) it is observed that in the modern Colloquial varieties strict agreement via BROKEN PL, typically 'enumerative', overlaps and in some varieties even replaces the deflected agreement, which, as seen above, surfaces as *-at* and performs a 'non-numerative' (= COLL) function.

The expectation is also realized for the strict agreement on the adjective via SOUND PL, as evidenced by the following Old Arabic agreement pattern, reported by Kahle (1975, p. 129). Here the adjectival SOUND PL surfaces as *-āt*:

(83) Old Arabic (*Koran*, 6:141)[21]

wa-huwa llaḏī 'anša'a jann-āṯ-i-n ma'rūš-āṯ-i-n
and-he who produced garden-*āt*-ACC-*n* trellised-*āt*-ACC-*n*

[21]The English translation of the Koran used here is by Arberry (1955), unless otherwise stated.

'It is He who produces gardens trellised'

Evidence for the same phenomenon in the present-day Low Variety comes from both a Bedouin variety (Ingham, 1994) and a Sedentary variety (Brustad, 2007), where strict agreement on the adjective behaves as in Old Arabic:

(84) a. Najdi Arabic (Ingham, 1994, p. 63)

byūt-in zēn-ah
houses big-*at*

'good houses'

b. Anazi Arabic[22] (Brustad, 2007, p. 16)

bgarāt guwiyy-āt
cows strong-*āt*

'strong cows'

(85) a. Levantine Arabic (Cowell, 1964, p. 201)

banadōr-āt māwiyy-e
tomato-*āt* juicy-*at*

'juicy tomatoes'

b. Levantine Arabic (Cowell, 1964, p. 201)

banadōr-āt māwiyy-āt
tomato-*āt* juicy-*āt*

'juicy tomatoes'

Because of the substantial identity between Old Arabic and the modern Colloquial Varieties, the agreement patterns of the former, drawn from Kahle (1975, p. 11), are taken in the following examples to be illustrative of the latter as well. This makes it possible to gloss over the situation of high linguistic variation found in the modern Colloquial Varieties which, however, have been significantly illustrated in (81 - 85) and are cited as examples when particularly relevant[23]. For the sake of comparison with Classical Arabic, the examples are modelled after Veccia Vaglieri (1937) and the reader is referred to Kahle (1975, pp. 77-129) for examples taken from the primary sources:

(86) HUMAN, SOUND PL, M

a. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʻallim-u l-mašhūr-u
the-teacher-NOM the-famous-NOM

'the famous teacher (M) '

b. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʻallim-ū-na l-mašhūr-ū-na
the-teacher-NOM-*na* the-famous-NOM-*na*

'the famous teachers (M) '

c. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

[22] A variety of Najdi Arabic spoken by the Anizah tribes of the Syrian Desert: see Section 2.2.1.2

[23] The reader is referred to Belnap (1999) for more data about the Modern Colloquial Varieties.

l-muʻallim-ū-na l-kibār-u
the-teacher-NOM-*na* the-bigs-NOM

'the old teachers (M) '

(87) HUMAN, SOUND PL, F

a. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʻallim-at-u l-mašhūr-at-u
the-teacher-*at*-NOM the-famous-*at*-NOM

'the famous teacher (F) '

b. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʻallim-āt-u l-mašhūr-āt-u
the-teacher-*āt*-NOM the-famous-*āt*-NOM

'the famous teachers (F) '

c. Egyptian Arabic (Woidich & Heinen-Nasr, 2004, p. 41)

(l-)ṭ̣ālib-āt (l-)hulandiyy-īn
(the-)student-*āt* (the-)Dutch-*īn*

'the Dutch students (F) '

d. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-muʻallim-āt-u l-kibār-u
the-teacher-*āt*-NOM the-bigs-NOM

'the old teachers (F) '

(88) HUMAN, BROKEN PL→ 'GROUP OF...'(= NON-HUMAN NOUN), M

a. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-rajul-u l-mašhūr-u
the-man-NOM the-famous-NOM

'the famous man'

b. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-rijāl-u l-mašhūr-ū-na
the-men-NOM the-famous-NOM-*na*

'the famous men'

c. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-rijāl-u l-mašhūr-at-u
the-men-NOM the-famous-*at*-NOM

'the famous men'

d. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-rijāl-u l-kibār-u
the-men-NOM the-bigs-NOM

'the old men'

(89) HUMAN, BROKEN PL→ 'GROUP OF...'(= NON-HUMAN NOUN), F

a. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-mar'-at-u l-mašhūr-at-u
the-woman-*at*-NOM the-famous-*at*-NOM

'the famous woman (F) '

b. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-nisā'-u l-mašhūr-āt-u
the-women-NOM the-famous-*āt*-NOM

'the famous women'

c. Anazi Arabic (Brustad, 2007, p. 16)

(l-)niswān (l-)zēn-īn
(the-)women (the-)good-*īn*

'(the) good women'

d. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-nisā'-u l-mašhūr-at-u
the-women-NOM the-famous-*at*-NOM

'the famous women'

e. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-nisā'-u l-kibār-u
the-women-NOM the-big-NOM

'the old women'

(90) NON-HUMAN: SPECIFIC OR GENERIC, SOUND PL, M / F

a. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-ḥammām-u l-naẓīf-u
the-bath-NOM the-clean-NOM

'the clean bath'

b. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-ḥammām-āt-u l-naẓīf-at-u
the-bath-*āt*-NOM the-clean-*at*-NOM

'the clean baths (F) '

c. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-ḥammām-āt-u l-niẓāf-u
the-bath-*āt*-NOM the-cleans-NOM

'the clean baths (F) '

d. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-ḥammām-āt-u l-naẓīf-āt-u
the-bath-*āt*-NOM the-clean-*āt*-NOM

'the clean baths (F) '

(91) NON-HUMAN: SPECIFIC OR GENERIC, BROKEN PL, M / F

a. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-sūq-u l-naẓīf-u
the-market-NOM the-clean-NOM

'the clean market'

b. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-'aswāq-u l-naẓīf-at-u
the-markets-NOM the-clean-*at*-NOM

'the clean markets'

c. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-'aswāq-u l-niẓāf-u
the-markets-NOM the-cleans-NOM

'the clean markets'

d. Old Arabic (based on Veccia Vaglieri 1937, p. 83)

l-'aswāq-u l-naẓīf-āt-u
the-markets-NOM the-clean-*āt*-NOM

'the clean markets'

The Table 3.17, reproducing Kahle (1975, p. 11), provides a synoptic overview of the specificity-driven agreement patterns.

Restricting attention on the nouns denoting humans, the SOUND PL, M if compared with the

Table 3.17: MCV and OA: specificity-driven agreement patterns

				Adjective		
				Strict Agr		Deflected Agr
				SOUND PL	BROKEN PL	
Noun	Human	M	SOUND PL	X	X	
			BROKEN PL	X	X	X
		F	SOUND PL	X	X	
			BROKEN PL	X	X	X
	Non-Human	M / F	SOUND PL	X	X	X
			BROKEN PL	X	X	X

other instances of PL seems to give rise to the stablest agreement pattern, in that it does not allow deflected agreement (= agreement in *-at*) on the adjective modifying it. In other words, the following SYSTEMATIC (yet unexplained) GAP appears to occur in Arabic:

(92) * Noun: SOUND PL, M + Adjective:*-at*

On the other hand the SOUND PL, F allows it in Classical Arabic, and so do the BROKEN PL, M / F in Old Arabic as well as in the Modern Colloquial Varieties. Moreover, all the instances of PL of nouns denoting humans, in all the examined languages, allow two kinds of strict agreement on the adjective, this latter displaying either the SOUND PL and the BROKEN PL.

Such a state of affairs is illustrated in the Table 3.18, where the symbol ⇒ indicates the systematic gap.

Remarkably, the agreement pattern triggered by the SOUND PL, M denoting *humans* is echoed in that triggered by the DU in general, regardless of its gender and of the nature of the entities denoted.

Whatever its explanation (see Section 3.3.2), this common behavior has at least in part to do with the fact that the DU shares with the SOUND PL, M some salient morphosyntactic properties, which qualify it as a kind of 'SOUND' DU. As Fleisch (1961, vol. I, pp. 284-285) remarks, both

Table 3.18: Noun denoting humans: distribution of deflected agreement

			Strict Agr		Deflected Agr
			SOUND PL	BROKEN PL	
CA	⇒ SOUND PL	M	X	X	
	SOUND PL	F	X	X	X
	BROKEN PL	M	X	X	
	BROKEN PL	F	X	X	
OA	⇒ SOUND PL	M	X	X	
	SOUND PL	F	X	X	
	BROKEN PL	M	X	X	X
	BROKEN PL	F	X	X	X
MCV	⇒ SOUND PL	M	X	X	
	SOUND PL	F	X	X	
	BROKEN PL	M	X	X	X
	BROKEN PL	F	X	X	X

of SOUND PL, M and 'SOUND' DU are expressed by an *n*-like suffix, drop this latter when it is followed by a complement of specification, and lengthen the vowel preceding it (see also Section 3.2.1).

Evidence from Old Arabic has led Brockelmann (1908) and Fleisch (1961, vol. I, pp. 284-285, 298-299) to go even further and claim that the suffix of the DU and of the SOUND PL (*nūn al-taṯniyah, nūn al-jam'*, respectively: see Section 3.2.1) are one and the same suffix, the former deriving from the latter via dissimilation (*-ā-ni* < *-ā-na* etc...).

The relevant data are reported in the following passage by the Late Grammarian 'Ibn Ya'īš:

> "...some Arabs pronounce the DU-ending *-ni* as *-na* when preceded by the ACC, OBL ending [*-ay-*], so that they treat [the sequence] */-a/y-*, which has morphemic status (*ġayra lāzimatin*), as if it were non-morphemic (*l-lāzimati*), namely as [the sequence] */-a/y-* found in words like *'ayna* [= when?], *kayfa* [= how?] and thus they say: *marartu bi-l-zaydayna* [= I passed by the two Zayds, instead of *marartu bi-l-zaydayni*] and *ḍarabtu l-zaydayna* [= I hit the two Zayds, instead of *ḍarabtu l-zaydayni*]. The grammarians of the school of Baghdad report these [data] and mention in this connection a line by [the poet] Ḥumayd Bin Ṯawr [...] Moreover, some Arabs pronounce the DU-ending *-ni* as *-na* when preceded by the NOM ending [*-ā-*]...".
>
> (*Šarḥ al-Mufaṣṣal*, vol. III, *al-Muṯannà*, pp. 190-191)[24]

It is thus expected that the 'SOUND' DU very much as the SOUND PL, M triggers on the adjective modifying it two kinds of strict agreement (see (70b), (70c), (86b), (86c)), making it display both the DU counterpart of the SOUND PL, that is a kind of 'SOUND' DU, as well as the DU

[24] *'...mina-l-'arabi man yaftaḥu nūna l-taṯniyati fi ḥāli l-jarri wa-l-naṣbi wa yujrī l-yā'a wa-'in kānat ġayra lāzimatin majrà l-yā'i l-lāzimati fī naḥwi 'ayna wa-kayfa fa-yaqūlu marartu bi-l-zaydayna wa-ḍarabtu l-zaydayna ḥakà ḏalika l-baġdādiyyūna wa-'anšadū ḥumayda bna ṯawr [...] wa-qad fataḥa-hā ba'ḍu-hum fī mawḍi'i l-raf'i...'*

counterpart of the BROKEN PL, that is a kind of 'BROKEN' DU[25].
As a matter of fact, both the expectations are met, although they appear to be mutually exclusive (complementary distribution): while, in fact, in Old and Classical Arabic the adjective agrees with the DU noun instantiating a 'SOUND' DU ending (93, 94), in the Modern Colloquial Varieties the adjective is marked for the 'BROKEN' DU in the same context (95) [26]:

(93) Classical Arabic (Wright, 1896, vol. II, p. 273)

'imr'-at-ā-ni ḥasan-at-ā-ni
woman-*at*-NOM-*ni* handsome-*at*-NOM-*ni*

'two handsome women'

(94) Old Arabic (*Koran*, 2:233)

wa-l-wālid-āt-u yurḍi'-na 'awlād-a-hunna ḥawl-ay-ni
and-the-parent-*āt*-NOM breastfeed-they children-ACC-they year-ACC-*ni*
kāmil-ay-ni
complete-ACC-*ni*

'Mothers shall suckle their children two years completely'

(95) Modern Colloquial Varieties (Blanc, 1970, p. 49)

bit-ēn kuḅār
house-DU bigs

'two big houses'

Accordingly, the distribution of the agreement patterns of the DU can be summarized as in the Table 3.19.

In spite of the fact that the DU does not trigger deflected agreement on the adjective modifying it, this agreement pattern is found in a category closely connected with it, that is the the so-called PSEUDO DU. This latter, in fact, albeit resembling the DU in its form (see Section 3.2.1), according to Blanc (1970, pp. 51-52) semantically behaves as a BROKEN PL and hence, like this latter (see 82), is able to mark the adjective modifying it for both the BROKEN PL if the

[25]The fact that the BROKEN PL adjective in the modern Colloquial Varieties works as DU may lead one to expect that the BROKEN PL noun also behaves the same way. This phenomenon is not discussed here because it has been rarely studied within the traditional paradigm, Blau (1979a) standing as an isolated case, but it is documented in Old Arabic and was already known to Arab Grammarians (*Šarḥ al-Mufaṣṣal*, vol. III, *Faṣl ja'l al muṯannà 'alà lafẓ al-jam'*, p. 209ff.), which report Koranic examples like:

(1) Old Arabic (*Koran*, 66:4)

'in ta-tūb-ā 'ilà -llāh-i fa-qad ṣaġ-at qulūb-u-kum-ā
if 2SG-repent-both to God-OBL and-already inclined-*at* hearts-NOM-you-both

'If you two repent to God, yet your hearts certainly inclined'

[26]Ferguson (1959a) was the first to observe this generalized agreement pattern for a very large sample of Modern Colloquial Varieties, and Blanc (1970) has further enriched it, thus confirming Ferguson's generalization. Blanc (1970, p. 51) also points out that sometimes in the Modern Colloquial Varieties there occurs strict agreement between a DU and an adjective marked for 'SOUND' DU, but he explicitly states that this is 'elevated or semi-literary usage', which leads to think that such an agreement pattern mimics the High Variety.

Table 3.19: OA, CA, MCV: agreement patterns of the DU

			Strict Agr		Deflected Agr
			'SOUND' DU	BROKEN PL	
CA	'SOUND' DU	M	X		
		F	X		
OA	'SOUND' DU	M	X		
		F	X		
MCV	'SOUND' DU	M		X	
		F		X	

reading is specific ('enumerative') and the ending *-at* if the reading is generic or COLL ('non-numerative'). This gives rise for the PSEUDO DU to contrasts like (96a - 96b), which parallel (82a - 82b):

(96) a. Egyptian Arabic (Blanc, 1970, p. 51)

rigl-ēn simīn-a
leg-PSEUDO DU fat-*at*

'fat legs'

b. Egyptian Arabic (Blanc, 1970, p. 51)

rigl-ēn sumān
leg-PSEUDO DU fats

'fat legs'

At the present state of research, evidence like (96) is limited to modern Colloquial Arabic Varieties, which results in an incomplete knowledge of the distribution of the agreement patterns of the PSEUDO DU, summarized in the Table 3.20.

To this, one may add with Blanc (1970, p. 46) that in some Colloquial Varieties not only the

Table 3.20: OA, CA, MCV: agreement patterns of the PSEUDO DU

			Strict Agr		Deflected Agr
			'SOUND' DU	BROKEN PL	
CA	'SOUND' DU	M	?	?	?
		F	?	?	?
OA	'SOUND' DU	M	?	?	?
		F	?	?	?
MCV	'SOUND' DU	M		X	X
		F		X	X

meaning but also the form of the PSEUDO DU has somehow to do with the PL rather than with the DU. In (Urban) Moroccan Arabic, for example, the ending of the PSEUDO DU is the same of the SOUND PL, ie *i(:)n* (e.g. *režl-i(:)n* is the equivalent of the Egyptian Colloquial Arabic (96) *rigl-ēn* 'foots, legs') and in fact Harrell (1962, p. 105) considers the PSEUDO DU as a SOUND PL, rather than as an independent category.

The picture that emerges is that the received view regards the PSEUDO DU of the Modern

Colloquial Varieties as a border-line category, sharing the nature of DU, BROKEN PL, SOUND PL.

One at this point may wonder whether the received view argues the same for Old and Classical Arabic: a look on the literature on the subject (Fleisch, 1961; Moscati et al., 1980; Lipinski, 1997; Hetzron, 1997; Eksell, 2006) shows that the answer is in the negative. The only notable exception is a recent study by Retsoe (1995), which, although mentioned in the 'non-orthodox' work by Owens (2006, p. 72), seems to have gone otherwise unnoticed[27].

Be that as it may, one may agree with Brustad (2000, p. 46ff.) that the agreement phenomena manifested by DU and PSEUDO DU, at least as documented in the modern Colloquial Varieties, clearly involve specificity, ie they are specificity-driven as much as it happens for the PL.

Finally, it is a well-established fact (Hashemi, 2006b, pp. 162-163 and references therein) that modern Colloquial Varieties and Old Arabic share another kind of agreement phenomenon, that is the total lack of agreement on the adjective, irrespective of the gender and number of the modified noun. This is arguably due to the appositional nature of the 'adjective' (see above). Adopting Brustad (2000, p. 51)'s terminology, lack of agreement will be referred to here as AGREEMENT NEUTRALIZATION.

For the sake of simplicity, Old Arabic data are used here to exemplify also the modern Colloquial Varieties[28].

They are reported by the Late Grammarian 'Ibn Ya'īš:

> "...and expressions like *rajul-u-n 'adl-u-n* [= 'a righteous man'] are described as verbal nouns [...] and such verbal nouns in their most typical form (*'aṣl*) cannot be marked for F, PL, DU, in spite of the fact that they occur in a domain of agreement: this is why you say: *hāḏā rajulun 'adlun* [= 'this is a righteous man'] [...], *marartu bi-'imra'atin 'adlin* [= I passed by a righteous woman] etc...".
> (*Šarḥ al-Mufaṣṣal*, vol. II, *Faṣl al-waṣl bi-l-maṣdar*, pp. 236-238)[29]

That is:

(97) Old Arabic (*Šarḥ al-Mufaṣṣal*, vol. II, *Faṣl al-waṣl bi-l-maṣdar*, pp. 236-238)

a. rajul-u-n 'adl-u-n
man-NOM-*n* righteous-NOM-*n*
'a righteous man'

b. rijāl-u-n 'adl-u-n
men-NOM-*n* righteous-NOM-*n*
'righteous men'

c. rajul-ā-ni 'adl-u-n
man-NOM-*ni* righteous-NOM-*n*
'two righteous men'

[27] Moscati et al. (1980, p. 93), Lipinski (1997, p. 244), Eksell (2006) - perhaps the most up-to-date account on the PSEUDO DU - dismiss the category of PSEUDO DU as a DU that has undergone some kind of development (desemantization etc...): see Section 3.3.2 for further discussion.

[28] The reader is referred to Hashemi (2006b, pp. 162 and references therein) for data from the Modern Colloquial Varieties. As for Old Arabic, already the traditional grammar recognizes the agreement-less pattern: see e.g. Wright (1896, vol. II, p. 274), Fischer (2006a, and references therein).

[29] '*...wa-yūṣafu bi-l-maṣādiri ka-qawli-him rajul-u-n 'adl-u-n [...] wa-l-'aṣlu 'anna-hā maṣādiru lā tuṯannà wa-lā tujamma'u wa-lā tu'annaṯu wa-'in jarat 'alà muṯannàn 'aw majmū'in 'aw mu'annaṯ taqūlu hāḏā rajul-u-n 'adl-u-n [...], marartu bi-'imra'atin 'adlin ...*'

(98) Old Arabic (*Šarḥ al-Mufaṣṣal*, vol. II, *Faṣl al-waṣl bi-l-maṣdar*, pp. 237-238)

a. 'imr'-at-u-n 'adl-u-n
woman-*at*-NOM-*n* righteous-NOM-*n*
'a righteous woman'

b. nisā'-u-n 'adl-u-n
women-NOM-*n* righteous-NOM-*n*
'righteous women'

c. 'imr'-at-ā-ni 'adl-u-n
woman-*at*-NOM-*ni* righteous-NOM-*n*
'two righteous women'

The Table 3.21 illustrates the distribution of the three agreement patterns discussed so far (notice that these latter include also DU and PSEUDO DU in the Modern Colloquial Varieties). In particular, the table clearly shows that in the modern Colloquial Varieties all these agreement patterns are found, a state of affairs referred to as EQUIVOCAL AGREEMENT by Ferguson (1989).
Other phenomena of agreement take place when the numeral is combined with the noun.

Table 3.21: Agreement in Arabic: semantics

	CA	MCV	OA
Animacy-driven agreement	X	(X)	
Specificity-driven agreement		X	X
Agreement neutralization		X	X

This kind of agreement is subject to the so-called GENDER POLARIZATION, by virtue of which a F numeral quantifies a M noun and vice versa: contrast the sequence *lə-wlad (M)... l-xams-ah (F)*, 'the five boys' of (64) with the sequence *l-ṣuḥuf-u (F) ... l-ṯalāṯ-u (M)* 'the three newspapers' of (65).
This is admittedly a simplified account of gender polarization, since it leaves out of the picture differences among Old Arabic, Classical Arabic and Modern Colloquial Varieties, which are partly rooted in syntax: e.g., as Ferguson (1959a, pp. 624-625) remarks, in the modern Colloquial Varieties the numeral, when prenominal, agrees in gender with the noun, hence with no gender polarization. Therefore, the interested reader is referred to the just mentioned work by Ferguson, as well as to Hetzron (1967).
To conclude this Section, the received view makes it clear that, taking apart gender polarization, which is not dealt with in the present work, specificity plays a key-role in the agreement phenomena instantiated in Old Arabic and in the modern Colloquial Variety - an account indicated here as Specificity Hypothesis.
Anticipating the discussion of the Section 3.3, it should be also noted that, on the other hand, the received view is not totally convincing as far as the Feature Copying Hypothesis is concerned. In Old Arabic and in the Modern Colloquial Varieties especially, the adnominal markers of case, number, gender, definiteness are not totally copied from the noun onto the adjective, contrary to what is traditionally assumed: sometimes they are found on the former but not on the latter (case, number), sometimes on the latter but not on the former (article).

3.2.2.2 Construct State and Free State

The noun modified by a genitive phrase is typically instantiated in Arabic by the Construct and the Free State, already described and exemplified in the Section 3.2.1.
Both the constructions are found in Old Arabic, Classical Arabic and in the Low Variety, albeit the Free State is used in Old and Classical Arabic only under certain circumstances (see immediately below) and the Construct State is lacking in some Colloquial Varieties, e.g. Shamaliya Arabic, a variety spoken in the northwest region of Morocco, as per Ouhalla (2010) and Cilician Arabic, a variety spoken in south-central Turkey (Owens, 2006, p. 113).
While in the received view the Free State has been identified with the Indo-european possessive construction and therefore paid limited attention, the Construct State has been the primary focus of its research and three properties have been ascribed to such a construction in the Arabist and Semitist (as well as generative) literature (Benmamoun, 2006):

- no modifier can intervene between the noun and its genitive phrase, so that this latter is the leftmost and closest constituent relative to the noun
- the thus formed complex behaves like a single word
- the state (= (in)definiteness marking) of the modified noun depends on the state of the genitive phrase, or, in other words, the (in)definiteness marking of the genitive phrase scopes over both the genitive phrase and the modified noun. Borrowing the term from the generative literature, such a property is designated here as (IN)DEFINITENESS SPREADING

The first property is clearly shown in (99 - 101):

(99) Old Arabic (*Koran*, 1:1 *et passim*)

bi-sm-i -llāh-i l-raḥmān-i l-raḥīm
in-name-OBL God-OBL the-merciful-OBL the-compassionate

'In the Name of God, the Merciful, the Compassionate'

(100) Classical Arabic (Wright, 1896, p. 222)

yad-u-hu l-yumnà
hand-NOM-his the-right

'his right hand'

(101) Najdi Arabic (Ingham, 1994, p. 51)

bēt al-bint al-ṭuwīl
house the-girl the-tall

'the tall house of the girl'

The second property is evidenced by the fact that in the Modern Colloquial Varieties the noun modified by a *nominal* genitive phrase, if bearing the ending *-at*, does not drop this latter (102)[30]. But this is exactly the form typically shown by the noun when it is modified

[30]This analysis entails that in the Construct State there applies a process of dropping in which *-at* is the input form and *-a(h)* its output. Yet, in principle one may also assume that the totally opposite process is at work: say, a process of *-t* insertion converting *-a(h)* into *-at*. This hypothesis, however, can be rejected on the basis of both native speakers' judgments and diachronic considerations, as already pointed out in fn. 4.

by a *pronominal* genitive phrase, that is a personal pronoun which in the Modern Colloquial Varieties is clearly a suffix attached to the noun and giving rise to a word-like unit (103):

(102) Moroccan Arabic (Benmamoun, 2006, p. 479)

madras-t Nadya
school-*at* Nadya

'Nadia's school'

(103) Moroccan Arabic (Benmamoun, 2006, p. 479)

madras-t-i
school-*at*-me

'my school'

The third property, since Wright (1896, vol. II, p. 225) has been exemplified with (Old and) Classical Arabic data like:

(104) Classical Arabic (Wright, 1896, vol. II, p. 199)

ḥarr-u l-šams-i
heat-NOM the-sun-OBL

'the heat of the sun'

(105) Old / Classical Arabic (Wright, 1896, vol. II, p. 225)

bint-u malik-i-n
daughter-NOM king-OBL-*n*

'a daughter of a king'

According to Wright (1896, vol. II, p. 225) Free State is related with (in)definiteness spreading in that it is employed when the speaker needs to characterize the modified NP and the genitive phrase with two different states of definiteness:

(106) Old / Classical Arabic (Wright, 1896, vol. II, p. 225)

bint-u-n li-l-malik-i
daughter-NOM-*n* to-the-king-OBL

'a daughter of the king'

A closer look at (Old and) Classical Arabic data as well as to native speakers'judgments, however, makes it questionable to ascribe indefiniteness spreading to the Construct State (a theoretical move designated here as (In)definiteness Spreading Hypothesis).
The Dutch Arabist De Goeje, in fact, in revising Wright's grammar reports an interesting counterexample to the (In)definiteness Spreading Hypothesis, namely that a noun modified by an *indefinite* genitive phrase can have a *definite* meaning, a phenomenon observed also by Veccia Vaglieri (1937, vol. I, p. 63):

(107) Classical Arabic (De Goeje *apud* Wright (1896, vol. II, p. 226))

'imr'-at-u ḥajjām-i-n
woman-*at*-NOM barber-OBL-*n*

'the wife of a (certain) barber'

De Goeje furthermore remarks (Wright, 1896, vol. II, pp. 198, 261) that the interpretation à la Wright (1896), ie (105) and the just mentioned counterexample are given an unified account if Arab Grammarians' view is adopted, because they regard the 'indefinite' Construct State as conveying, actually, a kind of *partial* definiteness (TAXṢĪṢ) rather than indefiniteness (TANKĪR). The Dutch Arabist proposes to identify Arab Grammarians' notion of *taxṣīṣ* with the concept of partial definiteness (his 'partial determination') on the basis of their definition of *taxṣīṣ*[31], which is reproduced in the following:

> "...the Construct State [...] conveys definiteness when [the genitive phrase is] definite and *taxṣīṣ* when [the genitive phrase] is indefinite.[...]
> The meaning of *taxṣīṣ* is according to their [= Arab Grammarians'] definitions : lessening the vagueness occurring in the indefinite nouns. Consider in this connection [the noun] *rajulun* [= a man] in the sentence *jā'a-nī rajulun ṣāliḥ* [= there came to me a pious man]: [the noun *rajulun*] if uttered in isolation by the speaker (*bi-waḍ'i l-wāḍi'i*) refers to any individual of the [human] kind but when you say *ṣāliḥ* you lessen [its] vagueness and restrict [its] reference...".
> (*Šarḥ al-Kāfiyah*, vol. II, *al-'Iḍāfah al-ma'nawiyyah, al-Na't*, pp. 206, 287)[32]

A strong argument Arab Grammarians bring in favor of such an interpretation is syntactic diagnostics. An indefinite noun (*l-nakirah*) cannot be topicalized, but an (allegedly) 'indefinite' Construct State can - this contrast, on their view, implies that the 'indefinite' Construct State is not such, being rather in a condition of *taxṣīṣ*:

> "...Know that the vast majority of grammarians [agree] on that the topic must be a definite noun or an indefinite noun which has undergone a kind of *taxṣīṣ*...".
> (*Šarḥ al-Kāfiyah*, vol. I, *Musawwaġāt al-'ibtidā'bi-l-nakirah*, p. 231)[33]

Ideed, a sentence like (109), featuring an indefinite topic is regarded by the Late Grammarian 'Ibn Hišām (d. 761 / 1359), *apud* Gully (1995, p. 244) as 'an extraordinary occurrence' (*min xawāriqi l-'ādah*).
Wright (1896, vol. II, pp. 261-262) himself is aware of this fact when he states that the topic of a sentence 'cannot , according to Arab Grammarians, consist of an indefinite noun or one which is not qualified by an adjective, or an expression equivalent to an adjective' . Such a state of affairs can be exemplified as follows:

[31] Veccia Vaglieri (1937, vol. I, p. 63), on the other hand, reports the counterexample at issue but claims that Arab Grammarians did not recognize similar cases, because they would have characterized it as indefinite ('Dagli Arabi però l'espressione è considerata in ogni caso indeterminata'): the passage quoted below shows the contrary. The only (notable) exception seems to be the grammarians of the Kūfan school, which take an indefinite noun to remain as such also when modified by some constituent: see al-'Astārābāḏī's passage quoted in the Section 3.2.2.3.

[32] '*... fa-l-ma'nawiyyatu [...] tufīdu ta'rīfan ma'a l-ma'rifati wa-taxṣīṣan ma'a l-nakirah [...] ma'nà l-taxṣīṣi fī-ṣtilāḥi-him taqlīlu l-'ištirāki l-ḥāṣili fī l-nakirāti wa-ḏalika 'anna rajulun [omission of the expected* ACC*-ending sic in the original text] fī qawli-ka jā'a-nī rajulun ṣāliḥ kāna bi-waḍ'i l-wāḍi'i muḥtamalan 'li-kulli faradin min 'afrādi hāḏā l-naw' fa-lammā qulta ṣāliḥ qallalta l-'ištirāka wa-l-'iḥtimāl...'*

[33] '*... 'i'lam 'anna jumhūra l-nuḥāti 'alà 'anna-hu yajibu kawnu l-mubtada'i ma'rifatan 'aw nakiratan fī-hā taxṣīṣan mā...'*

(108) Old / Classical Arabic (Wright, 1896, vol. II, p. 262)

rajul-u-n karīm-u-n ‘inda-nā
man-NOM-*n* noble-NOM-*n* at-us

‘a noble man is with us’

(109) Old / Classical Arabic (Gully, 1995, p. 244)

??? baqar-at-un takallam-at
cow-*at*-NOM-*n* spoke-*at*

‘a cow spoke’

This leads to restrict the third property of the Construct State to the definite Construct State only, and to replace it with the following two properties in the case of the ‘indefinite’ Construct State, à la De Goeje / Veccia Vaglieri (1937):

- syntactically, the ‘indefinite’ noun is combined with some extra-material (e.g. complement of specification or adjective)
- semantically, the noun has a partially definite reading

But these are exactly two properties that in modern linguistics are taken to be typical of a *specific* NP, since as far as the former property is concerned, ‘it has long been observed that "extra material" in an indefinite NP facilitates a specific interpretation’ (Enç, 1991, p. 22 and references therein), while the identification of weak / partial definiteness with specificity along the lines of Enç (1991) has been already discussed in the Section 3.2.2.1.
Such a state of affairs leads Brustad (2000, pp. 20-21) to claim that the relevant property of the ‘indefinite’ Construct State is specificity, to the effect that not only the two just mentioned properties of the the ‘indefinite’ Construct State boil down to one property - specificity -, but also the definite characterization of the definite Construct State falls within the same category. As Khan (1984, p. 470) and Brustad (2000, p. 22) highlight, definiteness from the perspective of Individuation Hierarchy is no more than a stronger degree of specificity:

(110) Individuation Hierarchy (Khan, 1984, p. 470)
1st > 2nd > 3rd > Human > Inanimate
Definite > Indefinite

This account, which will be henceforth referred to as Specificity Hypothesis (see also the end of the Section 3.2.1), has non-trivial implications.
First, it has an important bearing on the understanding of the ambiguity indefiniteness / definiteness observed in the Old Arabic and modern Colloquial modification-structures (50, 51, 52) dealt with in Section 3.2.2.1, which are now reinterpreted in terms of specificity. Such cases only differ from the constructions discussed so far in that it is the adjective instead of the complement of specification that acts as the extra-material which makes the modified noun specific. This account has been worked out by Brustad (2000, Ch. 1) for the modern Colloquial Varieties and is extended here to Old Arabic on the basis of the textual evidence provided in Section 3.2.2.1. This makes it possible to abandon the Definiteness Hypothesis in favor of the Specificity Hypothesis at least as far as Old Arabic and the Low Variety are concerned.

Second, it derives an additional counterexample to the (In)definiteness Spreading Hypothesis, raised by the German Arabist Nöldeke and included in Wright (1896, vol. II, p. 226).
The counterexample at issue is the opposite of (107), since it involves an (allegedly) 'definite' Construct State, which actually instantiates a noun modified by a *definite* genitive phrase having an *indefinite* meaning:

(111) Old / Classical Arabic (De Goeje *apud* Wright (1896, vol. II, p. 226))

ʾaḥad-u-hum
one-NOM-them

'someone of them'

It goes without saying that this behavior is perfectly natural under the Specificity Hypothesis, where the modified noun is partially definite / specific.
Last but not least, Wright (1896)'s argument that the Free State (e.g. 106) occurs when the states of the modified noun and of its genitive phrase do not coincide loses mots of its force, simply because the Free State syntactically features an 'indefinite' noun combined with some extra-material (a Prepositional Phrase), which renders it specific as much as its counterpart, the Construct State.
It is worth noting in this connection that at least as far as Old Arabic and Classical Arabic are concerned, native speakers' judgments confirm the strong degree of similarity holding between the Free State and the Construct State (here accounted for in terms of specificity), because they conceive the Free State (106) as the *ʾaṣl* of the Construct State (105), or, in modern terms, they conceive the Construct State (105) as a transformation of the Free State (106) (see the Sections 2.3.3, 3.2.1).
Notice in passing that Wright (1896, vol. II, pp. 124-125) partially contradicts himself and recognizes such a relationship when he deals with the so-called MEASURE-CONSTRUCTION (see immediately below).
There at least two aspects to this transformational relationship that the received view has taken up.
The former aspect concerns Arab Grammarians' contention that the 'indefinite' Free State derives not only an 'indefinite' Construct State, but also a '*definite*' Construct State, as discussed in Section 3.2.1 and illustrated in (11), (12). In this regard, Bohas et al. (1990, p. 63) and later on Kouloughli (1999, pp. 59-60) take the (alleged) discrepancy in definiteness ('statut déterminatif') between the 'indefinite' (11) and the 'definite' (12) as a crucial argument to reject the transformational relationship between them advocated by Arab Grammarians. Nevertheless, this criticism evaporates as soon as the Specificity Hypothesis is adopted, where such a discrepancy does not exist.
The latter aspect revolves around the *ʾaṣl* of the Construct State, which is not necessarily a Free State.
Wright (1896, vol. II, pp. 124-125), in fact, states that in (Old and) Classical Arabic a noun that denotes an unit of measurement is usually modified by a noun denoting the measured substance (cp. English *two kilos of bread*), thus giving rise to the so-called MEASURE CONSTRUCTION and he also points out that generally this semantic relationship is translated in syntax as the so-called TAMYĪZ-construction, where the modified noun takes the indefinite state

and bears the case-ending required by the context, while its modifier is in the same state and assigned ACC:

(112) Old / Classical Arabic (Wright, 1896, vol. II, p. 125))

'qafīz-u-n burr-a-n
qafīz-NOM-*n* wheat-ACC-*n*

'one *qafīz* (unit of measurement) of wheat'

Sometimes however, he adds, the Construct State replaces the TAMYĪZ-construction:

(113) Old / Classical Arabic (Wright, 1896, vol. II, p. 125))

'qafīz-u burr-i-n
qafīz-NOM wheat-OBL-*n*

'one *qafīz* (unit of measurement) of wheat'

This results in an alternation *tamyīz*-construction / Construct State that, according to Carter (1972), goes far beyond the measure-construction (Carter, 1972, p. 492), ranging over at least twenty grammatical structures (Carter, 1972, pp. 490-494)[34]. Noteworthily all of these alternations are already documented in Sībawayhi's *al-Kitāb*, ie in Old Arabic: this leads Carter (1972, p. 489) to conclude that 'it can properly be claimed that annexation and *tanwīn-naṣb* [= Construct State and TAMYĪZ-construction] are structurally equivalent[35]'.
Even more interesting is the fact that from Sībawayhi's onward the two constructions are not taken to be created equal. Carter (1972, p. 486), in fact, demonstrates that in the *Kitāb* a Construct State like (112) (his (3)) is derived from a *tamyīz*-construction like (113) (his (2)), and the Late Grammarian al-'Astārābāḏī couches this relationship in the *'aṣl*-based model, regarding the Construct State as a transformation of the *tamyīz*-construction, at least when its constituents are two nouns, the former denoting a thing and the latter denoting the material of which the thing is made:

> "...[the modified noun] that does not denote a unit of measurement is any item that has been assigned a particular name once that it has been taken out of [the material it is made of] and that, accordingly, can be modified by [a noun denoting] such a material: in fact, it is pefectly fine to mention the material [when referring] to these items: for example [you say, using a *tamyīz*-construction] *xātimun ḥadīdan, bābun sājan, ṯawbun xazzan* [= an iron ring, a teak door, a silk dress] - nevertheless, in these expressions the OBL-ending [of the Construct State instead of the ACC-ending of the *tamyīz*-construction] can also occur, more frequently than in the measure-constructions. This is so because a unit of *measurement* is semantically vague (*mubham*) and therefore requires a modifier denoting what *is measured* (*mumayyiz*) - now, it is exactly the ACC-ending that signals that a noun is a modifier denoting what is measured and it represents the basic case-ending (*'aṣl*) of the

[34] Thanks to Antonella Ghersetti for pointing me out the pervasiveness of this alternation.

[35] Carter (1972) makes little or no use of the grammatical terminology used in the Arabistic and Semitistic literature for philological reasons. His terminological choice is meant to highlight the fact that Early Arab Grammarians like Sībawayhi often had no explicit terms for grammatical constructions and linguistic phenomena. This by no means implies for Carter (1972) that Early Arab Grammarians lacked a clear understanding of their discipline and of the Arabic linguistic system and in fact Carter (1972) aims at showing the contrary.

tamyīz-construction: in this respect, the ACC-ending differs from the OBL-ending, which is the typical marker of the Construct State. On the other hand, the OBL-ending is preferable in expressions other than the measure-constructions, because the semantic vagueness of the former is not as strong as that of the latter [and hence the semantic conditions that trigger the ACC-marking are not met]. In addition to that, the OBL-ending is easier to pronounce because it is associated with the dropping of the *tanwīn* found on [the modified noun of] the Construct State...".
(*Šarḥ al-Kāfiyah*, vol. II, *al-'Iḍāfah al-ma'nawiyyah, al-Na't*, p. 57)[36]

The discussion so far allows to interpret the alternation Construct State / *tamyīz*-construction as a transformation which differs from the transformation Free State - Construct State only in that the element dropped is just one (the *tanwīn*), not two (the *tanwīn* and the preposition) - contrast (11 - 13) of the Section 3.2.1 with (112 - 114):

(114) (112) → TRANSFORMATION → (113)

(112) = N-*n* NP-ACC

TRANSFORMATION = Delete -*n*, Case-reassignment

(113) N NP-OBL

Staying with the Free State a little longer, it could be noticed that this construction in the Modern Colloquial Varieties (see (8) and the Tables 3.7, 3.7) stands as an alternative to the Construct State much more than it usually happens in Old and Classical Arabic (Owens, 2006, pp. 111-113 and references therein). In some cases, it is even the only construction available (Ouhalla, 2010). Without going into details, it is important to notice that Brustad (2000, pp. 70ff.) recognizes that in the Modern Colloquial Varieties specificity is a key-factor in the usage of the Free State, and more generally, in the choice between these latter and the Construct State, to the effect that the Specificity Hypothesis plausibly holds valid not only for the Construct State and the Free State in Old and Classical Arabic, but also for the Low Variety, as illustrated in the Table 3.22.

It has been also seen that the Specificity Hypothesis applies to an even broader domain,

Table 3.22: Construct State and Free State: semantics

	Specificity
CA	X
OA	X
MCV	X

since the extra-material that makes a noun specific is not only a complement of specification, but also an adjective (recall from Section 3.2.2.1 that this latter is equivalent in Arabic to an apposition)- as far as the adjective is concerned, however, the Specificity Hypothesis only

[36] *'...wa-ġayru l-miqdāri kullu far'in ḥaṣala la-hu bi-l-tafrī'i-smun xāṣṣun yalī-hi 'aṣlu-hu wa yakūnu bi-ḥayṯu yaṣuḥḥu 'iṭlāqu l-'aṣli 'alay-hi naḥwa xātimun ḥadīdan, bābun sājan, ṯawbun xazzan wa-l-xafḍu fī hāḏā 'akṯaru min-hu fī l-maqādīri wa ḏalika li-'anna l-miqdāra mubhamun muḥtājun 'ilà mumayyizin wa-naṣbu l-tamyīzi naṣṣun 'alà kawni-hi mumayyizan wa-hwa l-'aṣlu fī l-tamyīzi bi-xilāfi l-jarri fa-'inna-hu 'alamu l-'iḍāfati fa-hwa fī ġayri l-miqdāri 'awlà li-'anna 'ibhāma-hu laysa ka-ibhāmi l-miqdāri ma'a 'anna l-xiffata ma'a l-jarri 'akṯaru li-suqūṭi l-tanwīni wa-l-tanwīni bi-l-'iḍāfati...'*

accounts for modern Colloquial Varieties and Old Arabic, as summarized in the Table 3.23.
The conclusion is a sobering one: a more accurate look at Arab Grammarians' work, as well

Table 3.23: Adjective (/ Apposition): semantics

	Specificity	Definiteness-based distinction
CA		X
OA	X	
MCV	X	

as consideration of the data from the modern Colloquial Varieties implies a return to a *status quo ante*, by virtue of which the Specificity Hypothesis supplants (In)definiteness Spreading Hypothesis and the Definiteness Hypothesis.

3.2.2.3 Relative Clause

In Classical Arabic, the relative clause (ṢILAH or, in Sībawayhi's *al-Kitāb*, ḤAŠW) is accounted for along the lines of the Feature Copying Hypothesis, ie is said to be subject to the same copying phenomena affecting the adjective (although such an explanation runs into the same problems encountered in connection with this latter).
Thus, if the noun the relative clause refers to is definite, ie preceded by an article, the relative clause too is preceded by the article-like particle *llaḏī*, often referred to as RELATIVE MARKER or AL-'ISM AL MAWṢŪL (Wright, 1896, vol. II, p. 317)[37].
In addition to that (Wright, 1896, vol. I, p. 271) this particle is agreed with the modified noun in number, gender and (not always) in case according to the agreement patterns discussed in connection with the adjective (see Section 3.2.2.1 and Table 3.16).
On the other hand, if the noun the relative clause refers to is indefinite, the relative clause is juxtaposed to it, with no insertion of the relative marker - notice, furthermore, that the expected copying of the (alleged) marker of the indefinite state *-n* onto the relative clause does not take place.
At any rate, the received view generally takes the just mentioned opposition between definite and indefinite to be the most prominent distinction within the Classical Arabic relative clause, while the opposition between restrictive and non-restrictive relative clauses, usually observed in the Indo-european languages, is not regarded as pertinent in Classical Arabic.
This is why the relative clauses illustrated in the following can have both a restrictive and a non-restrictive reading (for the time being, suffice it to say that the absence vs. the presence of a comma distinguishes the former from the latter, pending further discussion in Chapters 5 and 6[38]:

(115) Classical Arabic (Wright, 1896, vol. II, p. 318)

l-malik-u	llaḏī	ya'dilu	
the-king-NOM	RELM	(,)	3M.SG-act.with.justice

[37] See also Mughazy (2006) which, however, concentrates on data from Modern Standard Arabic and the Low Variety.
[38] Wright (1896, vol. II, pp. 317-318) assigns (115) a restrictive reading and (116) a non-restrictive reading. The other interpretations are mine.

'The king (,) which acts with justice'

(116) Classical Arabic (Wright, 1896, vol. II, p. 317)

marar-tu bi-rajul-i-n ya-nāmu
passed-I by-man-OBL-*n* (,) 3M.SG M-sleep

'I passed by a man (,) who was sleeping'

The relative clause, regardless of its being definite or indefinite, instantiates in its structure a pronoun that refers back to the noun that precedes it, the so-called AL-ḌAMĪR AL-'Ā'ID (ANAPHORIC PRONOUN).
The anaphoric pronoun is always understood if it plays the role of a subject denoting the 3PRS SG[39] and in fact in the examples (115, 116) it is not evident and is taken by Arab Grammarians to occur after the verb. Their analysis relies upon the constrast between the SG and the PL / DU, where the subject, according to them, follows the verb and is phonologically realized within the strings *-ū-na, -ā-ni* respectively, although the exact location of the subject in these latter is not a matter of consensus (see (38)).
This contrast is discussed in greater detail in Fleisch (1961, vol. II, p. 136-143), Goldenberg (1988), Bohas et al. (1990, pp. 64-68), Versteegh (1997b, Ch. 5), Peled (2006) and can be summarized as follows:

(117) SG ya- nāmu$_{sleep}$ $HE_{SUBJECT}$ (he sleeps)
PL ya- nām-$_{sleep}$ ū-na$_{SUBJECT}$ (they sleep)

(118) SG ya- nāmu$_{sleep}$ $HE_{SUBJECT}$ (he sleeps)
DU ya- nām-$_{sleep}$ ā-ni$_{SUBJECT}$ (they two sleep)

Turning to the modern Colloquial Varieties, they differ from Classical Arabic in that they display a generalized relative marker not inflected for number, gender, case: e.g. *illi, yalli, halli, il, al, di* or the alike (Brustad, 2000; Vicente, 2006) - on the other hand, they appear at first glance to manifest the same definite / indefinite distinction found in Classical Arabic:

(119) Najdi Arabic (Ingham, 1982b, p. 252)

ana lāzim axxi āl ḏufīr alli amīr-ham ibn swēṭ
I necessary I.make.subjects.of Āl Dhafīr RELM leader-them Ibn Suwayṭ

'I must make subjects of the Āl Dhafīr (tribal clan) whose leader is Ibn Suwayṭ'

(120) Najdi Arabic (Ingham, 1982b, p. 254)

malḥa t-rāwiź wagt al-'aṣur ḏilla-h.
mount 3F.SG-dance time the-late.afternoon shadow-him

'A mount whose shadow dances in the late afternoon'

This example is taken from a Bedouin variety, but the same holds for more studied Colloquial Varieties like Egyptian, Levantine and Moroccan Arabic, as the data collected by Brustad (2000, p. 92) show (cp. also (121) immediately below).
Furthermore, the received view generally maintains that the just mentioned varieties instantiate

[39]The anaphoric pronoun is also optionally dropped when it works as object.

a subtype of definite relative clause modifying a noun not marked by the article and having properties similar to the Construct State, like the retention of the *t* of-*at* (see Section 3.2.2.2 and 102, 103), as the following contrast, reported by Haddad & Kenstowicz (1980) and Ouhalla (2004) mades adamantly clear[40]:

(121) Levantine Arabic (Ouhalla, 2004, p. 288)

1-baTT-a illi ʔakal-naa-ha
the-duck-*at* RELM ate-we-her

'the duck we ate'

(122) Levantine Arabic (Ouhalla, 2004, p. 291)

baTT-it illi ʔakal-naa-ha
duck-*at* RELM ate-we-her

'the duck we ate'

Interpreting a construction like (122) as 'a paraphrase of' (121) (Ouhalla, 2004, p. 291) is not always possible, since in some cases, reported by Feghali (1928, p. 311) for Levantine Arabic and by Brustad (2000, p. 93ff.) for Egyptian Arabic, the definite relative clause modifying a noun not marked by the article has actually an unambiguously non-definite meaning (Brustad 2000, p. 93 precises that such sentences cannot represent performance errors):

(123) Levantine Arabic (Brustad, 2000, p. 94)

fī-[h] wāḥd-e yalli bə-tzzakar-ha fī-ha ʾəsm-ha
in-it one-*ah* RELM PRS-I.remeber-her in-her name-her

'there's one (F) I remember that has her name'

(124) Egyptian Arabic (Brustad, 2000, p. 93)

fī-[h] tamsiliyy-a illi kān-u bi-y-gib-ū-ha fi t-tilivizyōn illi hiya
in-it serial-*ah* RELM were-they PRS-3SG-bring-they-her in the-TV RELM she
bi-t-ʾūl ḥabak-it ya 'umda
PRS-3SG.F-say was.necessary-she O Mayor

'there's a serial that they used to show on TV that says: is it really necessary, Mayor'

In all the examples above, the non-definite reading of the construction under scrutiny is forced by the fact that the modified noun is the subject of an existential construction (123, 124). According to Brustad (2000, p. 96) the non-definite reading of this construction naturally follows from the fact that the modified noun is not marked by the article and therefore can be interpreted not only as definite but also specific (recall that the indefinite interpretation is ruled out by the presence of the 'extra-material', ie the relative clause). In other words, Modern Colloquial Varieties witness specificity (110) at work in the relative clauses as much as it is in the agreement phenomena and in the Construct State / Free State.
It seems safe here to argue that the same holds for Old Arabic, on the basis of the following piece of data:

[40]Thanks to Jamal Ouhalla for pointing me out this contrast.

(125) Old Arabic (*Koran*, 104:1-2)

wayl-u-n li-kull-i humaz-at-i-n lumaz-at-i-n llaḏī jama‘a
woe-NOM-*n* to-every-OBL slanderer-*at*-OBL-*n* backbiter-*at*-OBL-*n* RELM collects
māl-a-n
wealth-ACC-*n*

‘Woe unto every backbiter, slanderer, who has gathered riches’

reported by the Arab Grammarians:

> “...some grammarians of the school of Kūfah [cp. Section 2.3.1] accept as grammatical (*’ajāza*) an indefinite noun modified by a definite relative clause, when this latter expresses blame or praise, bringing as evidence God’s word - be He exalted! - : Woe unto every backbiter, slanderer, who has gathered riches [*Koran*, 104:1-2)...”. (*Šarḥ al-Kāfiyah*, vol. II, *Al-ḥaqīqī wa-l-sababī min al-na‘t*, p. 308)[41]

Additional evidence comes from a passage of *Tārīx al-’umam wa-l-mulūk* (‘The History of Nations and Kings’) by al-Ṭabarī (d. 310 / 923), *apud* Reckendorf (1895, vol. I, p. 600, fn. 2), where, according to Reckendorf, an indefinite noun is modified by a relative clause introduced by *llaḏī* (‘der determinirte Relativsatz scheinbar zu einem indet[erminirte] Substantiv gehört’):

(126) Old Arabic (Reckendorf, 1895, vol. I, p. 600, fn. 2)

min niṣf-i yawm-i-n wa-kāna ma‘nà qawl-i l-nabiyy-i niṣf-u
from half-OBL day-OBL-*n* and-was meaning expression-OBL the-prophet-OBL half-NOM
yawm-i-n llaḏī miqdār-u-hu ’alf-u san-at-i-n
day-OBL-*n* RELM duration-it thousand-NOM year-*at*-OBL-*n*

‘from half a day, and Prophet Muḥammad’s words refer to half of a cosmic day, which lasts a thousand years[42] ’

Under the approach adopted here, the modified noun of (125, 126) is not indefinite, but specific, along the lines of the Arab grammarians other than the Kūfan ones (see Section 3.2.2.2) but this does not affect the main point: in any case, the noun at issue is not definite.
Notice also that (126) is documented about a couple of generations after the *terminus ad quem* of Old Arabic, but its rarity in Classical Arabic points to its relic-nature in this linguistic system.
The ‘specific’ relative clause discussed so far is not the only similarity between Old Arabic and modern Colloquial Varieties. Actually, the relative marker in the former sometimes surfaces

[41] ‘*...wa-’ajāza ba‘ḍu l-kūfiyyīna waṣfa l-nakirati bi-l-ma‘rifati fī-mā fī-hi madḥun ’aw ḏammun-i-stišhādan bi-qawli-hi ta‘alà waylun li-kulli humazatin lumazatin llaḏī jama‘a mālan ’*

[42] The context in which this passage is located is as follows, in Rosenthal (1989, pp. 182-183)’s translation: ‘Thus, (the evidence permitting) a conclusion is as follows: The beginning of the day is the rise of the dawn, and its end is the setting of the sun. Further, the reported tradition on the authority of the Prophet is sound. As we have mentioned earlier, he said after having prayed the afternoon prayer: What remains of this world as compared to what has passed of it is just like what remains of this day as compared to what has passed of it. [...] There is also a sound tradition on the authority of the Messenger of God, as I was told by Ahmad b. ’Abd al-Rahman b. Wahb- his paternal uncle ’Abd-allah b. Wahb- Mu’awiyah b. Salih- ’Abd al-Rahman b. Jubayr b. Nufayr- his father Jubayr b. Nufayr- the companion of the Prophet, Abu Tha’labah al-Khushani: The Messenger of God said: Indeed, God will not make this nation incapable of (lasting) half a day- referring to the day of a thousand years’.

as *llayī, d̲ū* and in this forms is not marked for number, gender and case (Rabin, 1951, pp. 154-155, 203) very much as it happens in the latter.
Another construction shared by Old Arabic and the modern Colloquial Varieties (more precisely, Levantine Arabic) is a type of definite relative clause that is characterized by the lack of the relative marker *lla-d̲ī* on (morpho-)syntactic level, albeit the modified noun is associated with the prenominal article:

(127) Old Arabic (*Koran*, 2:248)

'an ya'tiy-a-kumu l-tābūt-u fī-hi sakīn-at-u-n
that comes-SBJV-you the-ark-NOM in-it peace.of.reassurance-F-NOM-GM

'that the Ark will come to you, in it a Shechina from your Lord'

(128) Levantine Arabic (Feghali, 1928, p. 313)

'ẹn-nâṛ t-ḥarkés-a̲ dâimen b-t-ẹnṭfe
the-fire 2SG-touch-it in.any.instant FUT-3F.SG-get.extinguished

'the fire you touch in any instant will get extinguished'

This construction in Old Arabic appears to be semantically identical with the more common instance of definite relative clause proper, introduced by *lla-d̲ī* (e.g. (115)). In spite of Wright (1896, vol. II, p. 318)'s claim to the contrary (cp. also Brustad (2000, p. 98)), 'the commonly accepted rule that this should occur only when a noun is used in a generalized way (see Wright 2, 318) is not true of Early Arabic [= Old Arabic], where the relative adjective [= the relative marker] is quite often omitted' (Jones, 2005, p. 146).
The same holds for Levantine Arabic, since Feghali (1928, p. 313) paraphrases the relative clause of (128) *t-ḥarkés-a̲* as *'elli t-ḥarkés-a̲*, where *'elli* is the relative marker and explicitly states that sentences like (128) are the Levantine Arabic counterpart of the Old Arabic (127).
Finally, Old Arabic documents a type of relative clause that is also found in Classical Arabic and is characterized by an adjectival predicate (in Arab Grammarians' sense: see Section 3.2.2.1), the so-called EXTENDED ADJECTIVE or NA'T SABABĪ (Fischer, 2006a), which is exemplified by the following data from Classical Arabic (see Sībawayhi's passage quoted immediately below for data from Old Arabic):

(129) Classical Arabic (Wright, 1896, vol. II, p. 283)

jā'a-nī rajul-u-n ḥasan-u-n 'ax-ū-hu
came-me man-NOM-*n* man-NOM-*n*

'there came to me a man whose brother is handsome'

(130) Classical Arabic (Wright, 1896, vol. II, p. 284)

ra'ay-tu zayd-a-n al-ḥasan-a wajh-u-hu
saw-I Zayd-ACC-*n* the-handsome-ACC face-NOM-him

'I saw Zayd, whose face is handsome'

Interpreting similar constructions as a relative clause (Relative Clause Hypothesis henceforth) is an analytical trend well represented in the received view, from Reckendorf (1895, vol. I, p.

597) and Wright (1896, vol. II, p. 284) to Diem (1998) and Kremers (2003). In particular, the most recent works propose to analyze the extended adjective as a transformation of the relative clause, either in a (theory-)neutral meaning of the term (Diem, 1998) or within recent versions of the generative paradigm (Kremers, 2003) (see Section 1.1.3). Notice in passing that Wright (1896, vol. II, p. 284) in a likewise fashion paraphrases examples like (129-130) as relative clauses.
In these works, the Relative Clause Hypothesis is essentially corroborated by the fact that the noun following the adjective is annexed to a pronoun referring back to the modified noun:

(131) Extended Adjective $= \text{Noun}_i$ + Adjective + Noun + Pronoun_i

But the Relative Clause Hypothesis is likely to go back at the beginnings of the Arabic Linguistic Tradition, since Sībawayhi, whose interpretations plausibly reflect the judgments of native speakers of Old Arabic, assigns to the complex made up of the extended adjective and the noun following it the same syntactic status of a clause made up of a verb and a subject, which by definition in Arabic can be only understood as a relative clause when juxtaposed to the noun they refer to (see (115, 116)) :

(132) Extended Adjective $= \text{Noun}_i$ + (*l-* +) Adjective + Noun + Pronoun_i
Relative Clause $= \text{Noun}_i$ + (*llaḏī* +) Verb + Noun + Pronoun_i

As Sībawayhi puts it:

> "...when you use [extended adjectives and interrogative clauses like]: *marartu bi-rajulin ḥasanin 'abawā-hu* [= I passed by a man whose parents are handsome, an instance of extended adjective], *'a-ḥasanun 'abawā-hu* [= Are your parents handsome?], *'a-xārijun qawmu-ka* [= Is your clan getting out?], the structures [
> *ḥasanin + 'abawā-hu*: $\text{handsome}_{PREDICATE}$ + parents-him =
> *ḥasanun + 'abawā-hu*: $\text{handsome}_{PREDICATE}$ + parents-him =
> *xārijun + qawmu-ka*: $\text{leaving}_{PREDICATE}$ + clan-you]
> have the same syntactic status (*manzilah*) of [the affirmative clause] *qāla 'abawā-ka* [= your parents said...], *qāla qawmu-ka* [= your clan said...] [where deflected agreement takes place, that is:
> *ḥasanin + 'abawā-hu*: $\text{handsome}_{PREDICATE}$ + parents-him =
> *qāla 'qawmu-ka*: $\text{said}_{PREDICATE}$ + clan-you].
> Concentrating on the extended adjective], its structure is precisely what you obtain when you say: *qawmu-ka ḥasanūna* [= your clan is handsome] and then you shift [the noun] to the right (*'axxarū*), which results in the same syntactic status of: *'a-ḏāhibun 'abawā-ka* [= are your parents leaving?], *'a-munṭaliqun 'qawmu-ka* [= is your clan departing?], [that is
> *qawmu-ka + ḥasanūna*: clan-you + $\text{handsome}_{PREDICATE}$ → shifting →
> *ḥasanun + qawmu-ka*: $\text{handsome}_{PREDICATE}$ + clan-you =
> *munṭaliqun + qawmu-ka*: $\text{departing}_{PREDICATE}$ + clan-you
> More generally] if you put the name before the predicate and you say either *qawmu-ka ḥasanūna* [= your clan is handsome] or *qawmu-ka munṭaliqūna* [= your clan is departing], is the same [pattern of the affirmative clause that you use] when you

say: *'abawā-ka qālā* [= your parents said...], *qawmu-ka qālū* [= your clan said...], [a pattern where strict agreement takes place] ...".
(*al-Kitāb*, vol. I, *Hāḏā bāb mā jarà min al-'asmā' llatī min al-'af'āl wa-mā 'ašbaha-hā min al-ṣifāt llatī laysat bi-'amal naḥwa l-ḥasan wa-l-karīm wa-mā 'ašbaha ḏalika majrà l-fi'l 'iḏā 'aẓharta ba'da-hu al-'asmā' 'aw 'aḍmarta-hā*, p. 201)[43]

In this connection, Wright (1896, vol. II, p. 284) interestingly remarks that the extended adjective (= adjective + noun) is commutable with a construction where the order of constituents is reversed (= noun + adjective), such a construction being a relative clause:

(133) Old / Classical Arabic (Wright, 1896, vol. II, p. 284)

marar-tu bi- rajul-i-n ḥasan-at-u-n 'umm-u-hu
passed-I by man-OBL-*n* handsome-*at*-NOM-*n* mother-NOM-him

'I passed by a man whose mother is handsome'

(134) Old / Classical Arabic (Wright, 1896, vol. II, p. 284)

marar-tu bi- rajul-i-n 'umm-u-hu ḥasan-at-u-n
passed-I by man-OBL-*n* mother-NOM-him handsome-*at*-NOM-*n*

'I passed by a man whose mother is handsome'

The parallelism is *a fortiori* possible because, as Sībawayhi himself highlights in the just quoted passage, the word order of the clause in Arabic can manifest itself as VSO with deflected agreement (see (32) in Section 2.2.2.5), which is exactly the pattern found in the extended adjective, a similarity that can be illustrated as follows:

(135) VSO: Verb + Noun → DEFLECTED AGREEMENT
Extended Adjective: Adjective + Noun → DEFLECTED AGREEMENT

Recall however from the end of Section 2.2.2.5 that the similarity between the VSO clause and the extended adjective, if any, is expected to be in Old Arabic richer than in Classical Arabic, because Old Arabic VSO clause allows not only deflected agreement but also strict agreement, the so-called VSO-type *'akal-ū-nī l-barāġīṯ* (see (33)). This leads to predict for this language the following pattern too:

(136) VSO: Verb + Noun → STRICT AGREEMENT
Extended Adjective: Adjective + Noun → STRICT AGREEMENT

The prediction is confirmed, as (137) shows (the significance of this piece of data will become clear shortly):

(137) Old Arabic (*Šarḥ al-Kāfiyah*, vol. II, *l-Ḥaqīqī wa l-sababī min al-na't, Natā'ij li-ma taqaddama*, p. 308)

qāma rajul-u-n qā'id-ū-na ġilmān-u-hu
stood man-NOM-*n* sitting.down-NOM-*na* slaves-NOM-him

[43] *'...wa-ḏalika qawlu-ka marartu bi-rajulin ḥasanin 'abawā-hu 'a-ḥasanun 'abawā-hu' a-xārijun qawmu-ka fa-ṣāra hāḏā bi-manzilati qāla 'abawā-ka wa-qāla qawmu-ka 'alà ḥaddi man qāla qawmu-ka ḥasanūna 'iḏā 'axxarū fa-yaṣīru hāḏā bi-manzilati 'a-ḏāhibun 'abawā-ka 'a-munṭaliqun 'qawmu-ka fa-'in bada'ta bi-l-'ismi qabla l-ṣifati fa-qulta qawmu-ka ḥasanūna wa-qawmu-ka munṭaliqūna ka-mā taqūlu 'abawā-ka qālā wa-qawmu-ka qālū...'*

'A man, whose slaves are sitting down, is standing up'

On Arab Grammarians' view, however, Relative Clause Hypothesis is not restricted to the extended adjective alone, but involves the simple adjective as well, which is said to be a complex made up of an adjectival predicate and of the well known counterpart of the noun, ie a pronoun (ḌAMĪR), which in this case refers back to the modified noun. Such an account allows them to give the adjectival constructions an unified account.

The received view is sometimes aware of this generalized Relative Clause Hypothesis and in fact Wright (1896, vol. II, p. 284) and Bohas et al. (1990, p. 66) depict the parallelism between the extended adjective and the simple adjective along the following lines:

(138) a. Adjective = (Noun$_i$) + Adjective + Pronoun$_i$
Extended Adjective = (Noun$_i$) + Adjective + Noun + Pronoun$_i$

b. Adjective = Relative Clause
Extended Adjective = Relative Clause

Generally speaking, however, the received view does not adopt the interpretation (138) in its full extent, preferring instead to liken the Arabic adjective to the Indo-european one and thus conceiving it as a word, rather than as a (relative) clause. As a corollary, the extended adjective is given a marginal theoretical status, being considered in the received view as no more than an adjective associated with a complement of limitation or the alike.

Nevertheless, the Relative Clause Hypothesis seems to survive in the received view under a weak formulation, worked out by Reckendorf (1895, vol. I, p. 601), Brockelmann (1910, p. 123), Barth (1913, p. 157) and, more recently, by Fleisch (1961, vol. II, p. 86), Haddad & Kenstowicz (1980), having it that *llaḏī* is to be analyzed as the article *l-* plus some morphemic material *laḏī*. On this view, there exists in Arabic a structural parallelism between adjectives and relative clauses consisting of the absence vs. presence of the article *l-* intervening between the modified noun and the modifier[44]:

(139) a. Adjective = *l-* Noun + *l-* + Adjective
Relative Clause = *l-* Noun + *l-* (*laḏī*) + Adjective

b. Adjective ≈ Relative Clause
Extended Adjective = ???

(139) makes adamantly clear that in the received view the extended adjective (and in particular the alleged complement of specification) is not integrated within the Relative Clause Hypothesis. What has plausibly led the foundational texts of the received view to adopt a weaker version of the Relative Clause Hypothesis, giving the extended adjective a marginal theoretical status, is not lack of knowledge about Arab Grammarians' treatises - the just mentioned Wright (1896) and De Sacy (1831, vol. II, part IV) clearly show the contrary. Rather, this choice is motivated by the fact that adopting the fully-fledged Relative Clause Hypothesis leaves a major issue unsettled, namely the correct identification of the subject pronoun of both the simple adjective and the extended adjective.

This issue is part and parcel of the more general problem of the subject pronoun, already touched upon in (38) and in (117), where it was discussed that according to the Arab Grammarians:

- the subject pronoun is overt in the DU, PL, and follows the verb, but its exact location within the post-predicate morphemic material *-ū-na, -ā-ni* is hard to find (38)

[44] A view shared by less traditional threads of research: see Pennacchietti (1968, pp. 84-85).

- the subject pronoun is covert in the SG, and is taken to follow the verb on the basis of the syntactic position of the DU, PL (117)

Now Arab Grammarians, when taking into account instances of extended adjective displaying strict agreement like (137) point out that:

- the overt / DU, PL subject pronoun that seemingly follow the adjective, cannot be longer regarded as such because it intervenes between the adjective and the noun that already acts as the overt subject of this latter: otherwise, two subjects would co-occur, which is semantically impossible

Al-'Astārābādī's passage quoted below is a good illustration of this line of reasoning. To put it into context, the Late Grammarian, after recognizing that the extended adjective agrees with the noun it modifies in two out of seven features, that is case and definiteness, first deals with the parallelism between the extended adjective and the simple adjective, then proceeds to discuss the problematic nature of the pronoun associated with the extended adjective. In so doing he raises the just mentioned criticism about the co-occurrence of the nominal and pronominal subject, as well as two other kinds of criticism:

> "...the extended adjective is like the verb insofar the remaining five features are concerned: SG, DU, PL, M, F - thus, you agree it looking at the subject noun that follows it [not to the modified noun that precedes it], so that if the subject is SG, DU, PL you agree the extended adjective in the SG [via the expected deflected agreement and so on] [...].
> If you understand the validity of this position, you can easily see that also the former kind of adjective [= the simple adjective], which describes a particular condition of the modified noun, involves taking into account the role of the subject and is like the verb insofar the remaining five features are concerned, because it incorporates a pronoun as its subject (*li-'anna fā'ila-hu ḥīna'iḏ l-ḍamīru l-mustakinu fī-hi*), which refers back (*l-rājiʿ ilà*) to the modified noun: the verb, in fact, when it is the predicate of a subject pronoun (*'usnida 'ilà*) undergoes suffixation of *ā* if DU and of *ū* if PL denoting an animate being (*l-ʿāqil*) [and so on...].
> Therefore, you say: [...]*rajul-ā-ni ḍārib-ā-ni* [= two hitting men][45], *rijāl-u-n ḍārib-ū-na* [= some hitting men], [...] as much as you say when using a verb: *ya-ḍrib-ā-ni* [= both of them hit], *ya-ḍrib-ū-na* [= they hit] [...].
> It follows that a sentence like *qāma rajulun qāʿidun ġilmānu-hu* [= A man, whose slaves are sitting down, is standing up, exhibiting deflected agreement between the adjective and the following noun] is fine, while a sentence like *qāma rajulun qāʿidūna*

[45] Arab Grammarians make no difference between the adjective and the participle of an intransitive verb (see e.g. De Sacy (1831, vol. II, Ch. XII)): in these examples, the latter stands for the former because the conjugation of the verb *ḍaraba* 'to hit' is the paradigm that Arab Grammarians typically resort to in order to exemplify and describe linguistic facts - an usage due to Hellenist influence, since *ḍaraba* is the Arabic translation of Ancient Greek *tuptein*, which was the most popular and frequently used verb in Ancient Greek grammars (Versteegh, 1977, pp. 42-43). On these grounds, the occurrence of the intransitive participle *ḍārib-* instead of the adjective - say, *ḥasan* - does not affect the point made here. See also Sībawayhi's passage quoted immediately above, where intransitive participles like *leaving, departing* are commutable with the same adjective and Section 2.3.1 for the issue of the Greek influence on the Arabic linguistic thinking.

ġilmānu-hu [which means the same, but exhibits strict instead of deflected agreement] is borderline grammatical (*ḍa'īf*). Yet, the same sentence can be made acceptable (*yajūzu*) by saying: *qāma rajulun qu'ūdun ġilmānu-hu* [that is by replacing the SOUND PL *qā'idūna* with the BROKEN PL *qu'ūdun*].

On one side, the extended adjective is like the verb and hence *qā'idun ġilmānu-hu* [= whose slaves are sitting down, exhibiting deflected agreement between the adjective and the noun following it] is fine as much as it is: *yaq'udu ġilmānu-hu* [= his slaves are sitting down, exhibiting deflected agreement between the verb and the subject following it] [...] while: *(jā'a-nī rajulun) qā'idūna ġilmānu-hu* [= (there came to me a man) whose slaves are sitting down, exhibiting strict agreement between the adjective and the noun following it] is borderline grammatical because it has the same syntactic status (*manzilah*) of: *yaq'udūna ġilmānu-hu* [= his slaves are sitting down, exhibiting strict agreement between the verb and the subject following it] and in similar cases suffixing the endings of the DU, PL *ā, ū*, to the verb which works as predicate of the explicit DU, PL subject is borderline ungrammatical because these endings too play the role of the subject when affecting the verb in the vast majority of cases [to the effect that two subjects obtain, which is not desirable] - nonetheless, as described at the end of [Sībawayhi's] *al-Kitāb*, extended adjectives like *qā'idūna ġilmānu-hu* [= whose slaves are sitting down, exhibiting strict agreement between the adjective and the noun following it] are more acceptable than borderline grammatical VSO-clauses like *yaq'udūna ġilmānu-hu* [= his slaves are sitting down, exhibiting strict agreement between the verb and the subject following it] because of the following contrast: while depriving the verb of the endings of the DU, PL *ā, ū* [when no explicit subject occurs] is borderline grammatical, as it will be seen later, depriving the adjective of the same endings is possible, e.g. *ra'aytu qā'idayni* [= I saw [both of them] sitting down, where the DU-ending *-ay* replaces *-ā* and *ra'aytu qā'idīna* [= I saw [them] sitting down, where the PL-ending *-ī* replaces *-ū*] - this happens because these endings when found on the adjective [are not subjects but] markers (*'alamah*) of the DU, PL, as described at the beginning of [Sībawayhi's]] *al-Kitāb*: in fact, if they had been subjects they would have not turned [from *-ā*, *-ū*] into the ACC / OBL-ending *ay, ī*.

[...] On the other side, the sentence *qāma rajulun qu'ūdun ġilmānu-hu* improves the grammaticality of *qāma rajulun qā'idūna ġilmānu-hu* in spite of the fact that both *qā'idūna* and *qu'ūdun* are PL only because when you derives a BROKEN PL from an adjective, the form that you get does not follow the pattern of the verb (*xaraja lafẓan 'an muwāzanati l-fi'l*) nor is it constrained by its rules of well-formed agreement (*munāsabati-hi*) [that impose the deflected agreement to avoid the co-occurrence of two subjects]- this is so because a PL verb is never realized as a BROKEN PL [and hence using this latter in making the PL of the adjective does not violate any of the rules of well-formed agreement of the verb]...".

(*Šarḥ al-Kāfiyah*, vol. II, *l-Ḥaqīqī wa l-sababī min al-na't, Natā'ij li-mā taqaddama*, pp. 308-309)[46]

[46] *'...l-Ḥaqīqiyyu wa l-sababiyyu min al-na't [...] hāḏā l-sababiyyu fī l-xamsati l-bawāqī 'ay-l-'ifrādi wa-l-taṯniyati wa-l-jam'i wa-l-taḏkīri wa-l-ta'nīṯi ka-l-fi'li 'ay yunẓaru 'ilà fā'ili-hi fa-'in kāna l-fā'ilu mufradan 'aw*

Of particular relevance here is the fact that the construction that falsifies the subject-analysis of the DU, PL-endings suffixed to the extended adjective, namely (137), here repeated as (140):

(140) Old Arabic (*Šarḥ al-Kāfiyah*, vol. II, *l-Ḥaqīqī wa l-sababī min al-na't, Natā'ij li-ma taqaddama*, p. 308)

qāma rajul-u-n qā'id-ū-na ġilmān-u-hu
stood man-NOM-*n* sitting.down-NOM-*na* slaves-NOM-him

'A man, whose slaves are sitting down, is standing up'

is a clear instance of strict agreement paralleling the Old Arabic VSO-type *'akal-ū-nī l-barāġīṯ*, a parallel that is even stronger if one considers that Sībawayhi interprets the DU, PL-endings of the Old Arabic VSO-type *'akal-ū-nī l-barāġīṯ* as 'markers' (*'alamah*) (whatever the exact nature of such entities) rather than subject pronouns (*ḍamīr* (Goldenberg, 1988; Peled, 2006), that is exactly as the endings of the extended adjective exhibiting strict agreement in (140).

It follows that on a diachronic perspective dismissing the extended adjective so easily, as the received view does, is not totally correct, *a fortiori* when the extended adjective displays strict agreement, because it reflects the usage of Old Arabic. In fact Al-'Astārābāḏī in the just quoted passage hesitates to assign this particular kind of extended adjective a marginal status, first defining it as a borderline grammatical construction (*ḍa'īf*) but later conceding that it is more acceptable than a borderline grammatical construction (*'aqall min ḍa'f*).

On the other hand, the theoretical move of the received view, outlined in (139) has the advantage of circumventing a non-trivial interpretive problem arising from the parallelism between the extended adjective and the simple adjective. Once that the overt DU, PL endings of the extended adjective are no longer regarded as subject pronouns following their predicate, in fact, Arab Grammarians' unified account forces to adopt the same interpretation for the same endings occurring in the simple adjective. But at this point there is no more compelling evidence to assume that the syntactic position of a subject pronoun must follow its predicate and hence the covert SG subject pronoun may in principle either follow or precede its predicate:

(141) DU, PL = Noun$_i$ + Adj + *ūna, āni*: Pronoun$_i$
SG = Noun$_i$ + Adj + 0: Pronoun$_i$

(142) DU, PL = Noun$_i$ + *Pronoun$_i$*??? + Adj + *ūna, āni*: Marker$_i$ + *Pronoun$_i$*???
SG = Noun$_i$ + *Pronoun$_i$*??? + Adj + *Pronoun$_i$*???

muṯannàn 'aw majmuū'an 'ufrida-l-sababiyyu [...] wa-law naẓarta ḥaqqa l-naẓari la-wajadta l-'awwala wa-hwa l-waṣfu bi-ḥāli l-mawṣūfi 'ayḍan fī l-xamsati l-bawāqī manẓūran 'ilà fā'ili-hi wa-kā'inan ka-l-fi'li-'anna fā'ila-hu ḥīna'iḏ l-ḍamīru l-mustakinu fī-hi l-rājiʿu 'ilà mawṣūfi-hi wa-l-fi'lu 'iḏā 'usnida 'ilà l-ḍamīri yalḥaqu-hu l l-'alifu fī l-taṯniyati wa-l-wāwu fī jam'i l-muḏakkari l-'āqil [...] fa-li-ḏalika qulta [...] rajul-ā-ni ḍārib-ā-ni rijāl-u-n ḍārib-ū-na [...] ka-mā taqūlu fī l-fi'li [...] ya-ḍrib-ā-ni ya-ḍrib-ū-na [...] natā'iju li-mā taqaddama wa-min ṯamma ḥasuna qāma rajulun qā'idun ġilmānu-hu wa-ḍa'ufa qā'idūna wa-yajūzu qu'ūdun ġilmānu-hu 'ay min jihati 'anna l-sababiyya fī l-xamsati l-bawāqī ka-l-fi'li ḥasuna qā'idun ġilmānu-hu ka-mā ḥasuna yaq'udu ġilmānu-hu [...] wa-ḍa'ufa jā'a-nī rajulun qā'idūna ġilmānu-hu li-'anna bi-manzilati yaq'udūna ġilmānu-hu wa-liḥāqu 'alamatay l-taṯniyati wa-l-jam'i fī l-fi'li l-musnadi 'ilà ẓāhiri l-muṯannà wa-l-majmuū'i ḍa'īfun ka-mā yajī'u fī 'āxiri l-kitābi lakinna ḍa'fa qā'idūna ġilmānu-hu 'aqallu min ḍa'fi yaq'udūna ġilmānu-hu li-'anna l-'alifa wa-l-wāwa fī l-fi'li fā'ilun fī l-'aġlabi l-'akṯari wa-tajrīda-humā 'alamatay l-taṯniyati wa-l-jam'i ḍa'īfun ka-mā yajī'u bi-xilāfi l-'alifi wa-l-wāwi fī muṯannà l-'ismi wa-majmūʿi-hi fa-'inna-humā ḥarfāni wuḍi'ā 'alamatayni li-l-muṯannà wa-l-majmūʿi kamā maḍà fī 'awwali l-kitāb wa-law kānā fā'ilayni lam yanqalibā fī ḥālatay l-naṣbi wa-l-jarri wa-naḥwa ra'aytu qā'idayni qā'idīna [...] wa-'inna-mā qāma rajulun qu'ūdun ġilmānu-hu wa-'in kāna qu'ūdun jam'an ka-qā'idūna li-'anna-ka 'iḏā kassarta l-'isma l-mušābiha li-l-fi'li xaraja lafẓan 'an muwāzanati l-fi'li wa-munāsabati-hi li-'anna l- fi'la lā yukassaru...'

To summarize the discussion of this Section, there occur in Classical Arabic, Old Arabic and in the modern Colloquial Varieties several kinds of relative clauses, whose typology is summarized in the Table 3.24[47].

The table shows that Classical Arabic and the modern Colloquial Varieties share one out of

Table 3.24: The Arabic Relative Clause: typology

	l Noun Relm Agr RC	l Noun Relm RC	l Noun RC	Noun Relm RC	Noun n RC	Noun RC
CA	X				X	
OA	X	X	X	X	X	
MCV		X	X	X	X	X

six types of relative clauses, while Classical Arabic and Old Arabic share two out of them: in particular, in Classical Arabic, these constructions realize an opposition definite / indefinite. Old Arabic and the modern Colloquial Varieties share *four* out of six types of relative clause, and in both of them the relevant property is specificity rather than the just mentioned definiteness-based opposition, as illustrated in the Table 3.25. At least in these languages, therefore, the Specificity Hypothesis supplants the Definiteness Hypothesis, as much as it happens for the other kinds of modifiers, that is the complement of specification and the adjective (recall from Section 3.2.2.1 that this latter is equivalent in Arabic to an apposition).

In addition to the abovementioned types of relative clause, both the adjective and the so-called

Table 3.25: The Arabic Relative Clause: semantics

	Specificity	Definiteness-based distinction
CA		X
OA	X	
MCV	X	

extended adjective have been likened by Arab Grammarians and modern scholars to the relative Clause (Relative Clause Hypothesis), but their constituency is not totally clear, which has led the received view to opt for a weaker claim, namely that the adjective structurally parallels the relative clause, the extended adjective being left out of such a parallelism.

3.3 Some open issues

The previous Sections have discussed the five main assumptions that the received view has been entertaining in connection with the Arabic NP, namely

- Specificity Hypothesis
- Relative Clause Hypothesis (weak formulation)
- Feature Copying Hypothesis
- Definiteness Hypothesis

[47] The Table 3.24 abstracts away from the so-called Diptoptic 'Declension' of Classical and Old Arabic: see Section 3.2.1.

- (In)definiteness Spreading Hypothesis

and have plausibly demonstrated that the Specificity Hypothesis, first hinted at in Feghali (1928) as well as in Khan (1984), then developed and championed by Brustad (2000) for the modern Colloquial Varieties, can be extended to Old Arabic (and, in some cases, to Classical Arabic, albeit this is less relevant for the present work).
It has been also shown that the Specificity Hypothesis is compatible with the Relative Clause Hypothesis, while it rules out - at least as far the Low Variety and Old Arabic are concerned - the remaining hypotheses worked out in the received view, because they are based one way or the other on the opposition definite / indefinite.
Furthermore, the Specificity Hypothesis has turned out to be in line with - if not even pioneered by - Arab Grammarians' interpretation of Old Arabic data, and in particular with the interpretive tool of TAXṢĪṢ, as already recognized in Brustad (2000, p. 21).
Therefore, the received view will be taken into account in this work insofar as the following hypotheses are maintained:

- Specificity Hypothesis
- Relative Clause Hypothesis (weak formulation)

Nevertheless, in the literature other hypotheses have been elaborated around the Arabic NP, either within the received view itself, but less central to the academic research and / or the didactic practice, or outside it. They are reviewed in the next Sections.

3.3.1 Case

As already partially discussed in Chapter 2, interpretation of case falls into two categories, which can be labelled as follows:

- Meaningful Case Hypothesis
- Meaningless Case Hypothesis

The former approach can be further divided into:

- Standard Hypothesis
- Alternative Hypothesis
- Copula Hypothesis

The latter into:

- Epenthesis Hypothesis
- Language Contact Hypothesis

The Meaningful Case Hypothesis basically has it that case-endings in Arabic (and Akkadian) play some semantic roles (subject, object etc...), and are the grammaticalization of some pronominal stems (Fleisch, 1961, vol. I, p. 277).
In particular, the Standard Hypothesis is the normal currency in the received view (Wright,

1896; Moscati et al., 1980, etc...) and interprets the semantic roles (allegedly) played by the case-endings along the lines of the Indo-european languages.
The Alternative Hypothesis, put forward by Rabin (1969) is an attempt to depart from the Indo-european paradigm by deriving the semantics of the Arabic (and Akkadian) case-system from the notion of state, taken to be specific of Semitic, and from the Jespersenian concepts of nexus and junction, taken to be language-independent.
As for the Copula Hypothesis, it has been proposed by Nyberg (1920), which regards the case-endings of a particular kind of declension (the Diptotic one) as pronominal stems performing the function of a copula:

(143) N + Case_1 + Modifier + Case_2, Case_2 = copula

The principal difficulty with this account is that the copula is taken to derive from the (Old) Arabic SG, M pronoun *-hu(:)*, if NOM, and from the (Old) Arabic SG, F pronoun *-ha(:)*, if ACC / OBL but then there arises the question why the same referent is agreed sometimes for the M and sometimes for the F.
Apart from Owens (2006, Ch. 3, Ch. 4)'s arguments against case (e.g. case-endings are practically commutable: see Section 2.2.2.4), a strong criticism about the Meaningful Case Hypothesis has been raised by Corriente, which makes two observations about Arabic case (Corriente, 1971, 1973). First, the contexts where a difference of meaning can be effected by a change of case-ending alone (FUNCTIONAL YIELD) are very few, and in fact such a difference is made through word order and prepositions, as it happens in the Romance languages. Second, Arabic case-endings are characterized by lack of allomorphy, since they tack onto the end of a stem with little morpho-phonological interaction either with this latter or with other morphemic material which may be suffixed to them.
This has recently led Owens (2006) to dismiss the Meaningful Case Hypothesis as an instance of 'logical matrix' and to defend the Meaningless Case Hypothesis, or more precisely, the Epenthesis Hypothesis: case-endings are not functionally integrated into the grammar because they arise from epenthetic vowels.
Nevertheless, Tosco, *apud* Owens (2006, p. 101, fn. 22) signals that the Epenthesis Hypothesis is flawed in a major way: on a typological and diachronic viewpoint case develops out of previous morphemic material, not from epenthetic material. To circumvent this problem, Owens concedes the possibility that Arabic shows up as the result of linguistic contact, being borrowed from Akkadian (Language Contact Hypothesis).
Once again, the argument is not totally convincing, since case is a grammatical feature that can be borrowed only under certain circumstances, namely iff and only if a case-ending of the language source of the borrowing resembles a functionally related marker in the language target of the contact (Matras, 2009) - a situation not met with in Arabic.
Consequently, the validity of Owens (2006)'s Meaningless Case Hypothesis crucially depends of a more appropriate characterization of the Language Contact Hypothesis.
Recall that in Section 2.2.2.2 it has been suggested that a possible solution to this problem lies in interpreting the case as a connective / copula, an idea that will be implemented in the Chapters 6, 7: this is tantamount to saying that Nyberg's Copula Hypothesis is to be revived, but this move, as seen above, requires a principled explanation of the allomorphy NOM / ACC, OBL.

3.3.2 Gender and number

Gender and number have been regarded in the literature either as separate or interlocked phenomena (Hashemi, 2006b). These two trends of research can be defined as follows:

- Non-unified Hypothesis
- Unified Hypothesis

The Non-unified Hypothesis, following Hashemi (2006b) and Brustad (2000), can be further divided into:

- Sex-based Hypothesis
- Grammar-based Hypothesis
- Noun Class Hypothesis
- Specificity Hypothesis

as far as the gender is concerned, while the following hypotheses have been worked out for the category of number:

- Suffixation Hypothesis
- Agglutination Hypothesis
- Lengthening Hypothesis

to account for the SOUND PL (Hasselbach, 2007b), then:

- Dual Desemantization Hypothesis
- Depluralization Hypothesis

to account for the PSEUDO DU and finally:

- Deverbal Abstract Noun Hypothesis
- Pluralized Collective Hypothesis

to account for the BROKEN PL and the COLL.
The Unified Hypothesis, on the other hand, can be further divided into:

- Weak Unified Hypothesis
- Classifier Hypothesis

The Non-unified Hypothesis relies upon the methods of the Western traditional grammar, singling out the categories of gender and number and describing them separately.
Starting with number, the basic line of reasoning of the Non-unified Hypothesis is giving a prominent role to the SOUND PL and interpreting it along the lines of the Indo-european PL: the SOUND PL is built on the SG, mimicking, e.g., English *dog - dog-s, room - room-s* etc...
The DU as well is accounted for in a similar way. The strong solidarity between these two categories would explain why the suffix of the DU sometimes functions as a PL, that is why the PSEUDO-DU obtains (see Section 3.2.1 and (96)): this is the outcome of analogy proceeding from the SOUND PL to the DU, an account here referred to as Dual Desemantization Hypothesis (see Eksell 2006, Hasselbach 2007b, and references mentioned in Section 3.2.1):

(144) DU > SOUND PL ('PSEUDO DU')

Recently, however, Retsoe (1995) has put forward a different proposal which reverses the directionality of the just described process, claiming that the DU is a semantic specialization of the SOUND PL:

(145) SOUND PL ('pseudo du') > DU

This approach, here designated as Depluralization Hypothesis, is strongly supported by diachronical and typological evidence.
First, in Akkadian and in Old Arabic (Ferrando, 2006) the SOUND PL denotes a small number, ranging from 3 to 10: in particular for Old Arabic this meaning, technically referred to as PAUCAL PL or JAM' AL-QILLAH, is reported by Early Arab Grammarians (Ferrando, 2006).
Now, the lower bound of the small number encoded by a PAUCAL PL is known in typological literature to oscillate between two and three (Corbett, 200, p. 22) and this instability may result precisely in a restructuring of the category of number, which disambiguates between the quantity 'two' and 'three' differentiating between a DU and a PAUCAL PL whose lower bound is three. According to Retsoe (1995) it is precisely this process that is at work in Arabic (and Semitic), as (145) clearly shows.
Second, the PSEUDO DU can be hardly deemed to be an innovation, as the Dual Desemantization Hypothesis holds, since it is already documented in Akkadian, one of the most archaic documented Semitic languages.
For these reasons, the Dual Desemantization Hypothesis is abandoned here in favor of the Depluralization Hypothesis.
Returning to the SOUND PL, three different approaches have been worked out, according to Hasselbach (2007b), to properly characterize the Arabic / Semitic SOUND PL, which nevertheless share a key concern, namely explaining the long vowel that opposes the SG to the (SOUND) PL:

(146) SG: mu'min-u̲-n → (SOUND) PL: mu'min-ū̲-na
SG: mu'min-a̲-t → (SOUND) PL: mu'min-ā̲-t

In addition to that, they have in common the assumption that the DU and the SOUND PL are primitive entities, with the former possibly influenced by the latter (Dual Desemantization Hypothesis).
The Lengthening Hypothesis is championed in the received view by Moscati (1958), Moscati et al. (1980), as well as, in some cases, by the partisans of the Suffixation Hypothesis (see immediately below) and in part, outside it, by Garbini & Durand (1994, p. 100). Its gist is that the long vowel of the PL ending is derived from the short vowel SG via a rule of lengthening. This interpretation is grounded in the fact that vowel-lengthening is a widespread phenomenon in both Classical and Colloquial Arabic (see e.g. Owens (2006, p. 103)), while the main objection that one could raise (Hasselbach, 2007b, pp. 128-129) is that the matching between the short and the long vowel via the lengthening rule is quite loose. In fact, the ending *-āni*, which works either as DU in Arabic or as a particular kind of PL (arguably a PSEUDO DU) in both Arabic and Akkadian, cannot be reduced to a short-vowelled SG *-an* via a lengthening rule, given that the input form of *-āni* is *-un*, if DU, or *-un, -an, -in*, if PL:

(147) SG: mu'min-u̲-n → DU: mu'min-ā̲-ni

In a likewise fashion, the PL ending *-īn* has as its short-vowelled SG counterpart both *-an, -in*, as shown in the Table 3.5.
The Suffixation Hypothesis, advocated by Brockelmann (1908), de Lacy O'Leary (1923), Kurilowicz (1972) and Retsoe (1995), is devised to solve similar problems and contends that PL (and DU) endings include long vowels and are simply tacked as they are to the nominal stem.
This is at the same time the strong and weak point of this position which, while not affected by the problems of the Lengthening Hypothesis seems to loose most of the explanatory power of this latter. Hasselbach (2007b, p. 130) herself espouses the Suffixation Hypothesis but recognizes that such a move implies considering the F PL ending *-āt* as *not* derived from the F SG ending *-at*, the lengthening rule that allows such a derivation being rejected.
The Agglutination Hypothesis has been elaborated in less traditional trends of Semitistics, making extensive use of comparison between Semitic and Hamitic, and it has been defended by Petracek (1965), Zaborski (1976), Voigt (1997), Voigt (1999), Tropper (2004), Garbini & Durand (1994). On this perspective, the PL-ending is to be identified with the long vowel / glide *-w-* found, for example, in the Classical Arabic NOM (see (146)) and hence the long-vowelled endings other than it, ie *-īn* and *-āt* (see (146)) are reinterpreted as the result of a contraction between *w* and the vowels *i, a*, respectively (notice, however, that the long-vowelled ending *-āni* in (147) is not given the same interpretation).

(148) (SOUND) PL, NOM: mu'min-w-na → mu'min-ū-na
(SOUND) PL, ACC, OBL: mu'min-w + i-na → mu'min-ī-na

(149) (SOUND) PL, NOM: mu'min-w-at → mu'min-āt

This hypothesis has been in part forerun by the Arab Grammarians which interpreted the SOUND PL as the pluralizing morpheme (see (38) in Section 3.2.1).
In greater detail, the morpheme *-w-* and the phonological change are posited on the basis of data from some Semitic and Hamitic languages, where there occur PL endings combining *-w-* with *-n / -t* as well as phonological rules of contraction that affect the glide *-w-* when co-occurring with one or more vowel(s), as observed by Petracek (1965) and Garbini & Durand (1994, p. 100).
Instantiations of such a morpheme plausibly are, in Hamitic, the Lybico-Berber PL-endings *-awan, -iwan, -awin, -iwin, -awat* (Ratcliffe, 1998; Hasselbach, 2007b, and references therein):

(150) Amazigh (Boukhris et al., 2008, p. 47)
tudm-awin 'petits visages'

and in Semitic some forms attested in Syriac (an Aramaic variety) like:

(151) Syriac (Hasselbach, 2007b, p. 125)
lebb-ā / lebb-awāt̲ā 'heart / hearts'
nahr-ā / nahr-əwāt̲ā 'river / rivers'

as well as in Classical Arabic:

(152) Classical Arabic (Hasselbach, 2007b, p. 123)
san-at / san-awāt 'year / years'
fāris / f-awā-ris 'horseman / horseman'

As (152) clearly shows, the same pluralizing morpheme *-w-*, according to the partisans of the Agglutination Hypothesis, occurs not only in the SOUND PL but also in the BROKEN PL (Hasselbach, 2007b, p. 123).

The Agglutination Hypothesis has two advantages. First, since 'there is no doubt that Afroasiatic had a plural marking morpheme {w} or {u}'(Hasselbach, 2007b, p. 129), the existence of such a morpheme is empirically grounded; second, even in the cases in which it is not visible, the underlying processes that opacize it are 'theoretically possible' (Hasselbach, 2007b, p. 129). The greatest obstacle encountered by this Hypothesis are, in Hasselbach (2007b, p. 129)'s opinion, its inability to account for the Diptotic endings, but it should be noted here that this objection loses most of its force if one recalls from the Chapter 2 and Section 3.3.1 that projecting back whatever case-ending to Semitic - not to speak of Hamito-Semitic - is in all likeliness a 'logical matrix'to be avoided. Therefore, among the hypotheses concerning the SOUND PL, the Agglutination Hypothesis seems to be the more solid.

Turning to the BROKEN PL, the Non-unified Hypothesis takes it to be a deverbal abstract noun (something like English *jogg-ing*) that gradually shifted into the role of PL (cp. English *(a bag of) shopp-ing*), through an intermediate stage of concrete COLL (cp. English *tub-ing*), as schematized in (153), adapted from Kurilowicz (1972, p. 139):

(153) deverbal noun > concrete COLL > BROKEN PL

Following Ratcliffe (2006), this position (here designated as Deverbal Abstract Noun Hypothesis) is typical of the received view and in fact goes back to Barth (1904) and is reiterated in much subsequent work, e.g. Fleisch (1961), Kurilowicz (1961), Kurilowicz (1972), Moscati et al. (1980).

The principal difficulty with this account, as Ratcliffe (2006, p. 446) points out, is that *de facto* the most productive BROKEN PL patterns are *not* attested in the usage of verbal nouns or any other non-PL function in Arabic, and more generally, in Semitic. On these grounds (but see also p. 136 for further arguments), the Deverbal Abstract Noun Hypothesis is to be discarded.

As for the gender, the Non-unified Hypothesis seems to have long abandoned the Sex-based Hypothesis advocated by Wright (1896) in favor of the Grammar-based Hypothesis. The sex-based approach, in fact, relies upon *biological* categorization to assign M and F markers to the noun, but this is clearly not the case in Arabic, where both these markers can be assigned to sexless entities and the F can be sometimes associated with a noun denoting a male (e.g. *-at* in *xalīf-at-* 'caliph').

For similar reasons scholars like Brockelmann (1908) and Cuny & Feghali (1924) already at the beginning of the XX century argued for the Grammar-based Hypothesis, according to which nouns are assigned M and F markers on the basis of purely *linguistic* (ie *cognitive*) criteria, like animacy.

This approach, however, is not able to account for the multifunctional usage of the F markers like *-at* etc... which are not only opposed to M, but also used as:

- singulative (see al-'Astārābāḏī's passage immediately below)
- COLL: *kam'-ah*, 'bunch of truffles' (Fleisch, 1961, vol. I, p. 302, fn. 2)
- deverbal abstract noun (see al-'Astārābāḏī's passage immediately below)
- diminutive: *duwayr-ah*, 'petit maison' (Fleisch, 1961, vol. I, p. 326)

- intensive: the just mentioned *xalīfat-*, 'caliph'

Such values, however, cannot be dismissed so easily, because already native speakers of Old Arabic appear to have had a good perception of them (Hashemi, 2006b), as the following passage clearly shows[48]:

> "...the second [kind of *-at*] serves to distinguish the single inanimate entities and the single deverbal abstract nouns from the 'species' they belong to:
> *naxl - naxlah* [= palm grove - palm]
> *tamr - tamrah* [= head of dates - date]
> *baṭṭ - baṭṭah* [= sore of ducks - duck]
> *naml - namlah* [= swarm of ants - ant]
> and in fact the Koran [27:18] reads: *qālat namlatun* [= an ant said], where *namlatun* can be understood as M because *-at* denotes the single entity (*li-l-waḥdati*) and accordingly the *-at* of *qāl-at* [= said] [being the reflex of agreement with the noun marked for the singulative *-at*] is to be understood as a singulative, not as a real F.
> The deverbal abstract nouns can be exemplified as follows:
> *ḍarb - ḍarbah* [= to hit - *a* hit]
> *'ixrāj - 'ixrājah* [= expulsion - *an action of* expulsion]
> *'istixrāj - 'istixrājah* [= extraction - *an operation of* extraction]...".
> (*Šarḥ al-Kāfiyah*, vol. III, *Al-ma'ānī llatī tajī' la-hā l-tā'*, p. 325)[49]

Furthermore, on diachronic level, Old Arabic nouns are generally formed by using different stems for male and female humans and animals, rather than by adding *-at* (Fleisch, 1961; Hashemi, 2006b), as exemplified in Table 3.26, drawn from Hashemi (2006b, pp. 158-159).

This has prompted Arabists and Semitists following both the received view (Fleisch, 1961; Moscati et al., 1980; Drodźik, 1998) and alternative trends of research (Garbini & Durand, 1994) to maintain that multiplicity of functions of -AT (and, more generally, of any Semitic (alleged) F marker) has its origin in its being a kind of morpheme also found in some African languages, the so-called NOUN CLASS. This is a morpheme tacked onto a noun to express the degree of its animacy, size, socio-cultural importance (Aikhenvald, 2000, Ch. 2). Such an interpretation, however, is not compelling in the light of (Khan's /) Brustad's Specificity Hypothesis, since all the values of *-at* can be derived as different specializations of the semantic kernel 'individuated / specific' encoded in the singulative *-at* and at the core of the Individuation Hierarchy (110). For the purposes of the present work, Brustad (2000, p. 24)'s discussion can be represented in the following compositional terms:

(154) singulative *-at*: [specific], relative to the generic group (p. 91)
COLL *-at*: [specific], *qua* concrete, relative to the abstract noun
deverbal abstract noun *-at*: [specific], relative to the generic verb (p. 133)

[48] Notice in particular that al-'Astārābāḏī explicitly states that *-at*, when singulative, can be understood either as M or F, a claim shared with other grammarians (Fleisch, 1961, vol. I, p. 302) and recently confirmed by Drodźik (1998, p. 35) diachronic research.

[49] '*...l-ṯāniyu li-faṣli l-'āḥādi l-maxlūqati wa-'āḥādi l-maṣdari min 'ajnāsi ka-naxlin wa-naxlatin wa-tamrin wa-tamratin wa-baṭṭin wa-baṭṭatin wa-namlin wa-namlatin fa-fī qawli-hi ta'alà qālat namlatun yajūzu 'an yakūna l-namlatu muḏakkaran wa-l-tā'u li-l-waḥdati fa-takūnu tā'u qālat li-tā'i l-waḥdati fī namlatin lā li-kawni-hā mu'annaṯan ḥaqīqiyyan kamā yajī'u wa-l-maṣādiru naḥwa ḍarbun wa-ḍarbatun, wa-'ixrājun wa-'ixrājatun, wa-stixrājun wa-stixrājatun...*'

Table 3.26: Old Arabic: natural gender nouns

F	M	Gloss
'ab	'umm	father / mother
šayx	‘ajūz	old man / woman
ḥiṣān	faras	stallion / mare
ḥimār	'atān	[male / female] donkey
xuzāz	'arnab	[male / female] rabbit
-	murḍi‘	breastfeeding woman
-	ḥāmil	pregnant
-	ṭāliq	divorced woman
-	kā‘ib	buxom
-	‘ānis	spinster

diminutive *-at*: [specific], relative to the affective sphere / property
intensive *-at*: [specific], relative to the social sphere

A further argument against the Noun Class Hypothesis is the just quoted al-'Astārābāḏī's passage, which describes the COLL as the input of a derivational process that yields the output singulative / SG via suffixation of *-at*:

(155) Old Arabic (*Šarḥ al-Kāfiyah*, vol. III, *al-Jam‘*, p. 325)
naml (COLL) → *naml-ah*: (SG, M / F)

Thus, al-'Astārābāḏī (and more generally Arab Grammarians) make it very clear that in Arabic the singulative *-at* is opposed to the COLL, and that in this opposition the former is the MARKED (= complex) term, the latter the UNMARKED (= simple) one. But in the noun class languages the opposite state of affairs is typically found (Aikhenvald, 2000, pp. 50-51 and fn. 29). In other words, the Noun Class Hypothesis is not compatible with native speakers' jugdments and is therefore to be rejected.
To summarize the discussion so far, within the Non-Unified Hypothesis framework the best approach proposed for number is the Agglutination Hypothesis, and for gender the Specificity Hypothesis, which has the merit to recognize in the singulative function (NOMEN UNITATIS, 'ISM WAḤDAH) the primary value of *-at*.
This morpheme has also attracted attention from linguists outside Arabist and Semitist studies and the reason for such interest arguably lies in its *systemic* nature. The singulative *-at*, the alleged *gender*-marker, in fact, differs from its diminutive, intensive etc... counterparts in that it is strictly interlocked with the category of *number*, since it is *systematically* opposed to the COLL, as evidenced by Arab Grammarians' description (= native speakers' judgments) (see (155)).
This relationship has informed recent discussions about a rethinking of Arabic ‘gender’ (and, more generally, the category of gender *per se*) that can be referred to as (Weak) Unified Hypothesis and summarized as follows (see Hashemi 2006b for more details and references):

> “gender has the function of qualitatively more precisely defining a quantity. Gender offers the opportunity to refine the crude perspective of number — singular

versus plural — into distributive versus collective plural. It is this aspect of quantity that links gender so closely to number".
(Unterbeck, 1999, p. xxv)

In so doing, Unterbeck (1999, pp. xxx-xxxi) (explicitly) revives Greenberg's Universal 36, according to which the presence of the category of gender always implies the presence of the category of number (Greenberg, 1963, p. 95) and similar views are advocated by Zabbal (2002); Acquaviva (2008).
But Greenberg (1990b) himself, when exploring in the seventies such a strict correlation in connection with Arabic, implemented this idea in a stronger version, by taking the typical context of occurrence of *-at* to be those Arabic constructions in which the noun associated with *-at* is modified by a numeral (see Section 1.1.3 and in particular (17)).
In this light, *-at* turns out to be a (NUMERAL) CLASSIFIER rather than a NOUN CLASS, the former being distinguished from the latter precisely by the obligatory association with a numeral, as widely documented in the Southeast Asian languages[50]. This approach, here indicated as Classifier Hypothesis, has been subsequently developed in Fassi Fehri & Vinet (2004, 2007), Ouwayda (2009) and was also independently pioneered by Bauer (1912) (although Greenberg does not seem aware of this, since he never cites Bauer's work).
One may wonder at this point how instances of number other than SG and COLL are accommodated within the the Unified Hypothesis.
Starting with the SOUND PL, the Unified Hypothesis by and large reiterates the arguments of the Non-unified Hypothesis.
Both the Weak Unified Hypothesis (Zabbal, 2002; Acquaviva, 2008)[51]) and the Classifier Hypothesis (Greenberg, 1990b, p. 178) assume a plural formation that proceeds from the SG to the (SOUND) PL, thus sharing with the received view the Indo-european logical matrix (*dog - dogs*)[52].
Greenberg (1990b, p. 181) seems to be aware that this conceptualization is problematic for the Classifier Hypothesis, because in the numeral classifier systems the PL does not occur and the SG enters in relationship as *output* with the COLL rather than entering in relationship as *input* with the PL (see (155)).
Greenberg is thus forced to consider the manifestation of the SOUND PL in numeral classifier systems like Arabic as exceptional (and Aikhenvald 2000, p. 120 as 'incipient') :

"we may suggest as a hypothesis that when a language is an exception to the implicational universal that numeral classifiers imply the absence of compulsory plurals, what is involved is a subsystem of such a singular / plural language within which the basic opposition is collective / singulative rather than singular / plural".
(Greenberg, 1990b, p. 181)

(156) illustrates the conundrum raised by Greenberg (1990b):

[50] See page 12 for an informal definition of (NUMERAL) CLASSIFIER and some examples.
[51] More precisely, Acquaviva (2008, p. 200) claims that the SOUND PL is not formed from the SG-stem, but from a stem form 'systematically related (typically identical) to that of the singular' . This detail does not affect the point made here.
[52] This logical matrix would be even stronger in Ratcliffe (2006, and references therein)'s approach, where the process SG → PL is taken to apply also to the BROKEN PL. See also the following footnote.

(156) Expected: COLL → SOUND) PL (Classifier - systems): YES
Logical matrix: SG → (SOUND) PL (Indo-european): NO

Turning to the BROKEN PL, both the Weak Unified Hypothesis (Acquaviva, 2008; Zabbal, 2002) and the Classifier Hypothesis (Greenberg, 1990b, pp. 179-180) retain the idea of the Non-Unified Hypothesis that this category diachronically arises from the COLL ('les pluriels internes sont des collectifs précisés en pluriel': Fleisch 1961, vol. I, p. 309), but the former ones depart from the latter in that the COLL is not taken to derive in its turn from a deverbal abstract noun.
Instead, Acquaviva (2008, pp. 222-24) and Greenberg (1990b, p. 180) reverse the directionality of the process and contend that the deverbal abstract noun originates from the COLL, thus 'bypassing' the received view and espousing Arab Grammarians' view (see al-'Astārābāḏī's passage on p. 133)[53].
This results in *two* separate processes (Pluralized Collective Hypothesis), summarized in (157, 158) which replace the single process posited by the Deverbal Abstract Noun Hypothesis (153):

(157) concrete COLL > BROKEN PL

(158) 'COLL'> Deverbal Abstract Noun

In particular, according to Greenberg (1990b, p. 179) and Acquaviva (2008, pp. 227-228), the development in (157) can be easily tested relative to COUNTABILITY: the COLL cannot be counted, the BROKEN PL can[54].
Thus, insofar as meaning is concerned, the BROKEN PL seems to be more MARKED (ie complex) than the COLL relative to the piece of information (or: SEMEME, FEATURE) 'countability' . Such a state of affairs plausibly impacts on the form of the nouns as well, because in the majority of cases the BROKEN PL is more complex than the COLL, the stem of the former undergoing phenomena of lengthening, gemination and affixation not found in the latter (Ratcliffe, 2006, and references therein):

(159) Old / Classical Arabic (Fleisch, 1961, vol. I, p. 309)

naml COLL: swarm of ants FORM: NO length. MEANING: NO count.
nimāl BROKEN PL: ants FORM: YES length. MEANING: YES count.

Acquaviva (2008, p. 231) uses contrasts like that illustrated in (159) as evidence to claim that the opposition between the COLL and the BROKEN PL underlies a system (SG -) uncountable

[53]Greenberg (1990b) deliberately revives Arab Grammarians' analysis modelling his theory after al-'Astārābāḏī's *Šarḥ al-Kāfiyah*. Zabbal (2002) relies upon Greenberg (1990b) and Acquaviva (2008) in its turn follows Zabbal (2002). On the other hand, Ratcliffe (2006, and references therein)'s position is even more drastic: BROKEN PL would be a primitive entity, created as such rather than developed out of a COLL. This position, however, is not totally convincing, given that Ratcliffe (1998) himself in examining a large sample of Hamitic and Semitic languages is forced to admit throughout the paper that the COLL overlaps with the BROKEN PL, vaguely assuming that 'the singulative / collective pattern is formally similar with' the singular / BROKEN PL pattern (Ratcliffe, 1998, p. 88) and that 'the two classes have merged' (Ratcliffe, 1998, p. 93). Nevertheless, he does not provide a principled explanation of such similarity / merger.

[54]More precisely, a COLL cannot be counted through the usual numeral constructions of Arabic, but sometimes they can be counted via some kind of periphrasis, e.g. via insertion of the preposition *min* of the Free State, as much as it happens to the English COLL *cattle* in the expression *two heads of cattle*. Greenberg (1990b) and Acquaviva (2008) in this connection report data from Wright (1896, vol. II, p. 237). Recall also from Section 3.2.2.1 that even if a COLL cannot be *directly* counted, it can be nevertheless associated with predicates like *count* or *distinguish* (see (68)). Thanks to Paolo Acquaviva for helpful discussion on these issues

- countable and that this latter category may include all kind of PL, ie both BROKEN PL and SOUND PL, *qua* opposed to the COLL.

In so doing, the Weak Unified Hypothesis (Acquaviva, 2008) admittedly develops an account already proposed in the received view by Fleisch (1961, vol. I, p. 308) which, in dealing with the singulative *-at*, *de facto* takes out of the picture the deverbal abstract noun and rather states that in Classical Arabic 'on a eu un système complete: collectif, nom d'unitè, pluriel' regardless of whether this latter is broken or sound:

(160) Old / Classical Arabic (Fleisch, 1961, vol. I, p. 308)

baqar baqar-at baqar-āt
cow.COLL cow-SG cow-PL

(161) Old / Classical Arabic (Fleisch, 1961, vol. I, p. 308)

šajar šajar-at 'ašjār-
tree.COLL tree-SG tree.PL

Interestingly, the Classifier Hypothesis Greenberg (1990b, pp. 178-179) not only interprets the data depicted in (160), featuring a SOUND PL, along the same lines of Acquaviva (2008) / Fleisch (1961), telling *how* they are organized, but also provides a motivated explanation for it, telling *why* they are so organized. the three term system, in fact, is a morpho-semantic architecture typical of the numeral classifier systems (Greenberg, 1990b, pp. 179, 184) (where, in the non-exceptional case, the SG, PL *meanings* are realized combining the Classifier with a numeral / quantifier).

Therefore, the tripartite division à la Fleisch, far from being language-specific, is a cognitive phenomenon, and in fact it partially occurs also in the Indo-european languages, as the contrast between (160, 161) and (162) clearly shows:

(162) English (Greenberg, 1990b, p. 183)

police policemen policeman
UNCOUNTABLE COUNTABLE SG

On these grounds, the Classifier Hypothesis seems to be preferable to the Weak Unified Hypothesis because, all the other things equal, the former has the ability to explain the three term system that remains a descriptive statement in the latter.

Thus, as far as number is concerned we are left with two hypotheses: the Agglutination Hypothesis (a subcase of the Non-Unified Hypothesis) and the Classifier Hypothesis (a subcase of the Unified Hypothesis).

The two hypotheses at issue are not mutually exclusive, at least in one fundamental respect. Upon closer scrutiny, in fact, both of them seem to provide an unified treatment of the SOUND PL and of the BROKEN PL.

On one side, in fact, the Classifier Hypothesis (as well as the Weak Unified Hypothesis) draws a clear *distributional* parallelism (commutability) between the SOUND PL and the BROKEN PL, where this latter is said to be a former COLL (see (160), (161)).

On the other side, the Agglutination Hypothesis draws a clear *structural* parallelism (*w*-affixation) between the SOUND PL and the BROKEN PL, this latter developing out of a COLL.

Although these converging scenarios cannot be regarded as conclusive evidence[55], they seem to

[55]See Chapter 7 for data in support of (164).

conspire to characterize, by transitive property, the SOUND PL as originating from a COLL, as depicted in the following diagram (where ⇳ stands for 'commutable'):

(163) COLL > BROKEN PL: ⇳, *-w-*$_i$
SOUND PL: ⇳, *-w-*$_i$

(164) COLL > BROKEN PL: ⇳, *-w-*$_i$
COLL > SOUND PL: ⇳, *-w-*$_i$

Taking into account gender, the Unified Hypothesis is compatible with the non Unified Hypothesis insofar as the singulative function of *-at* (and its allomorphs) is considered.
The Agglutination Hypothesis and the Classifier Hypothesis, however, cannot be accepted without reservations.
Although, in fact, numeral classifier systems with an agglutinating type[56] are found crosslinguistically (Aikhenvald, 2000, p. 99), some questions remains unaswered:

- where are the expected numeral / quantifier in the three term system (160), (161) (see Section 1.1.3)?
- which is the singulative counterpart of the SOUND PL-ending *-na*?

[56] An AGGLUTINATING (type of) language is a language in which (ideally) many morphemes combine each other to yield a single word (as if they were 'glued each other', hence the name) and each of them conveys one and only one piece of information. To put it differently, if the word at issue is decomposed and each morpheme is analyzed separately, one finds a one-to-one correspondence between the form and the meaning:

(1) Turkish (Comrie, 1981, p. 79)

ev + ler + im + e
house + PL + my + DAT

'to my houses'

As such, an agglutinating language is opposed to an ISOLATING language, where a word is not composed by many morphemes, being itself a single morpheme. Simply put, a word is a morpheme that is kept 'isolated' from the other morphemes (hence the name). The one-to-one correspondence between form and meaning, then, is even more transparent:

(2) Thai (Comrie, 1981, p. 77-78)

khi tôi d̄'ên nhá bạn tôi ...
when I come house friend I ...

'When I came to my friend's house...'

An agglutinating language is also opposed to a FUSIONAL language. In this latter, in fact, many morphemes combine each other to yield a single word, as much as it happens in an agglutinating language, but each morpheme can convey more than one piece of information, to the effect that two or more pieces of information are 'fused' into a single morpheme and a one-to-many correspondence between form and meaning may obtain:

(3) Latin

lup + u + s
wolf + M.SG + NOM

'he-wolf'

This taxonomy serves expository purposes and abstracts away from the relationships holding among the just described types: see Comrie (1981, Ch. 2, 4) and Thornton (2005, Ch. 3, 5, 6, 9) for more details.

- How can the Depluralization Hypothesis (145) be integrated within both the Agglutination Hypothesis and the Classifier Hypothesis? (recall that both of them espouse the Dual Desemantization Hypothesis (144) that however has proven to be unvalid)

To sum up, the Classifier Hypothesis, if combined with the Agglutination Hypothesis appears to account for the Arabic (and Semitic) gender and number better than other approaches. Nevertheless, its validity depends on its ability to answer to the just raised questions.

3.3.3 The article

The interpretation of the Arabic (and Semitic) article can be divided into three main trends, all of them discussed at length in Mascitelli (2006, Ch. 2):

- the Deictic Hypothesis
- the Dissimilation Hypothesis
- the Clitic Hypothesis / Copula Hypothesis

The first approach is the most widespread and is maintained both in the received and in less traditional views: Fleisch (1961) is but a sample for the former, Garbini & Durand (1994) well exemplifies the latter.
It proposes that the Arabic (and Semitic) article develops out of a demonstrative, along the lines of English *that* > *the* and of Italian *ille* > *il*:

(165) (Distal) Dem + N → Art + N

The major problem with this interpretation is that diachronic evidence documents an evolutionary process for the article sensibly different from that observed in the Indo-european languages (cp. the Clitic Hypothesis / Copula Hypothesis immediately below). This leads to consider the Deictic Hypothesis as an instance of logical matrix (see Section 2.2.2.3) and then to discard it.
The second proposal, advocated essentially by Ullendorff (1965) and Lambdin (1971), has it that Arabic (and Semitic) word-initial consonants tend to geminate and that subsequently the geminated consonant undergoes dissimilation under certain phonological circumstances, resulting in the original consonant preceded by *-l*:

(166) #C → #CC → #l-C

This proposal reverses the directionality of the process described in the Arabic Linguistic tradition (in its essence, an assimilation) and recorded in the Arabic writing (see e.g. Wright (1896, vol. I, p. 15)):

(167) #C → #l-C → #CC

The main objection to the Dissimilation Hypothesis, raised by Voigt (1998), is that writing is a conservative phenomen, tending to record words in their etymological forms.
The third interpretation, hinted at in Feghali (1928), and developed independently by Pennacchietti (1968), Testen (1998) and Pat-el (2009), claims that the Arabic (and Semitic) article arises as an element intervening between the noun and its modifier, and that only later does it become prenominal:

(168) N + 'Art' + Modifier → Art + N

Feghali (1928), Pennacchietti (1968), Testen (1998) and Pat-el (2009) share with Pennacchietti (1968) the idea that the context in which the article originates is a modification structure. Pennacchietti (1968), moreover, identifies the modification structure with a predication-structure, to the effect that the article appears to be a copular element. Accordingly, this approach is designated here as Clitic Hypothesis / Copula Hypothesis.
In the case of Arabic, the element at issue was originally asseverative, ie meant to emphasize the syntactic constituent it refers to (STATUS-MARKER), and is to be connected with the particle *la*, which may still perform this original function in Old and Classical Arabic.
In diachrony, the particle *la* is documented from Amorite onwards (Testen, 1998, pp. 110-112, 167-168) (the overall meaning of the expressions it occurs in being open to various interpretations):

(169) Amorite (Testen, 1998, p. 168)

zi-id-ku-la-na-si
justice-*la*-Naśī

'(The) justice (of) Naśī (?)'

Accordingly, *l-* is older than *la*, this latter being no more than the vowelless *l-* that becomes syllabified (Testen, 1998, Ch. 3): in this respect, Testen in part revamps Barth (1913)'s positions (Testen, 1998, pp. 137-138).
Nevertheless, the Clitic Hypothesis / Copula Hypothesis belittles the role of the etymological semantics of the article, given that, as Testen (1998, p. 160) puts it, 'the key to the development of the article from the status-marker lies not so much in its original value as in its location'.
Testen (1998)'s evidence for this claim is primarily syntactic and typological, since he observes that Arabic *l-* originates as a clitic and therefore is affected by WACKERNAGEL'S LAW. This latter states that the site between the first and second elements of a sentence is the prime location for the insertion of clitics, and Testen, after unveiling a strong structural parallelism between the distribution of *la* in the sentence and that of *l-* in the Noun Phrase (pp. 160-163), extends Wackernagel's law from the former to the latter (pp. 163-168).
The same 'structural' idea is pursued by Pennacchietti (1968) and Pat-el (2009) which, however, deal with the Semitic Article in general without focusing on its Arabic manifestation and it is implemented through a different line of reasoning, which is essentially comparative: Pat-el (2009) and Pennacchietti (1968) provide additional philological evidence for the emergence of the 'article' in a clitic position not only for Arabic but also for Hebrew and Phoenician.
Testen and Pat-el argue furthermore that the article becomes prenominal as a consequence of some process of DISPLACEMENT[57], by virtue of which the article instead of following the

[57]This phenomenon, also known as DISLOCATION, can be defined along the lines of Moro (2000) as:

> "a specific property of natural languages: certain constituents are interpreted in a position other than the one where they occur in the sequence of words. Considering the matter informally, let us look at a simple case:
>
> (1) a. John believes that Mary knows [this story]
> b. [Which story] does John believe that Mary knows?
>
> Even though the constituent *which story* does not follow *knows* in the sequence of words in (1b),

noun precedes it. It appears that this is the weakest point of the Clitic Hypothesis / Copula Hypothesis.
On one side, in fact, Pat-el (2009) seems not to explain the reason why displacement applies, while Testen (1998)'s explanation of it is not totally satisfactory. In the latter case, Testen claims that displacement, a *syntactic* phenomenon, takes places to avoid consonantal clustering, a *phonological* phenomenon, but the influence of the segmental level of phonology (consonant, vowels etc...) on syntax seems to be unprobable on typological grounds, as Eisele (2001, p. 328) notices in reviewing Testen's work[58].
On the other side, the displacement they posit aims also at deriving a construction like (44), here repeated as (170) and characterized by a double occurrence of the article:

(170) Najdi Arabic (Ingham, 1994, p. 51)

al-bēt al-ṭuwīl
the-house the-tall

'the tall house'

that is:

(171) N + Art + Modifier → Art + N + Art + Modifier

Yet, displacement in the natural languages does not retain memory of its previous position, leaving a copy of itself in this latter: simply put, in an English sentence like *its wheels can spin*, the auxiliary *can* undergoes displacement in its interrogative counterpart, but this results in the sentence *can its wheels spin?*, not in the sentence *can its wheels <u>can</u> spin?*. A sentence of this kind is in fact documented only under very particular circumstances, namely (Radford, 2004, pp. 155-157) in the speech that children use during language acquisition (MOTHERESE) - which is clearly not the case of Arabic (and Semitic) article.
On these grounds, the Clitic Hypothesis / Copula Hypothesis can be accepted in the former part of its formulation (article as an originally intervening element, along the lines of Wackernagel's law):

(172) N + 'Art' + Modifier

but not in the latter (article as prenominal via displacement):

(173) ??? N + Art (+ ...) → Art + N (+ ...)

3.3.4 Tanwīn

Two major analytical trends can be found in the research on the tanwīn, indicated here as:

- the Semantic Hypothesis

it is interpreted as the object of *knows* paralleling the case in (1a)".
(Moro, 2000, p. 1)

It is customary to refer to the position of the constituent before displacement as its TRACE, notationed *t*. In the generative paradigm, the normal practice is to call displacement MOVEMENT but in some of its branches also the term RAISING occurs. These issues will be taken up in greater detail in Chapter 4 .

[58] Whereas *non*-segmental properties of phonology (e.g. ordering of segments, ie their linearity and syllabicity) does impact on syntax, as shown by Moro (2000). This topic will be delved into in Parts II and III.

- the Structural Hypothesis

The Semantic Hypothesis, in its turn, can be split into two broad categories:

- the Specificity Hypothesis
- the Definiteness Hypothesis

By the same token, two main approaches are pursued within the Structural Hypothesis framework:

- the Demarcative Hypothesis
- the Copula Hypothesis

As discussed at length in the previous Sections, the Specificity Hypothesis is due to Brustad and in this particular case it argues (Brustad, 2000, pp. 27ff.) that the specificity characterizing expressions like (108) is encoded in the tanwīn.
This claim, however, seems to be at odds with Brustad (2000, p. 21)'s claim that in the same expressions specificity obtains by simply adding a modifier to a noun (which, remarkably, in some of the modern Colloquial Varieties examined by her, does not exhibit -*n* at all).
The Definiteness Hypothesis is discussed at length in Fleisch (1961, vol. I, pp. 342-345) and, more recently, in Mascitelli (2006, Ch. 2).
It has been worked out in the received view and is championed by Kurilowicz (1950), Kurilowicz (1972, Ch. VII) and Fleisch (1961, vol. II, p. 342-345). They contend that -*n* was originally a marker of definiteness and that subsequently turned into a marker of indefiniteness because of the restructuring of the system. After the introduction of the article, this latter takes over the function of the marker of definiteness, and therefore -*n* loses this value and is assigned a new function that opposes it to the article, namely that of marker of indefiniteness (this process is known in the structuralist tradition as POLARIZATION):

(174) N + 0 DEF → N + n = DEF → l + N = DEF
N + 0 INDF → N + 0 = INDF → N + n = INDF

All the forms and the meanings depicted in (174) are attested in Arabic and Epigraphic South-Arabian.
Recent work by Kouloughli (2001), Kouloughli (2007) pursues the same line of research, but on synchronic perspective: while no attempt is made to establish the historical evolution of -*n*, this scholar works out an exhaustive taxonomy of the several semantic nuances that both the definite and the indefinite -*n* display in Arabic.
Among the problems raised by the Semantic Hypothesis approach is the fact that in diachrony the indefinite bare noun is expected to behave like a relic (being replaced in its function by the noun associated with -*n*), but such an expectation is not met, because the bare noun is still productive as indefinite form in both Old Arabic and Classical Arabic (pausal form) as well as in the modern Colloquial varieties.
Another problematic issue is the definite value of *n*. Its alleged archaicity does not go back before Epigraphic South-Arabian, and therefore partisans of the Semantic Hypothesis have also claimed that -*n* is to be etymologically connected (via dissimilation or the alike) with a morpheme -*m* found in Akkadian, which would be a marker of definiteness. But even accepting

such a connection, it has been clear since the thirties to scholars like Gelb (1930) and von Soden (1955) that *-m* in Akkadian can be also interpreted as being neither indefinite nor definite.
Gelb (1930), and many other after him (Garbini & Durand, 1994, and references therein), in particular, rely upon evidence from Akkadian to maintain that *-m* has no semantic import, and that it is associated with the noun to make it recognizeable and identifiable as such in a given speech act - technically speaking, *-m* has a DEMARCATIVE FUNCTION.
Similar positions therefore illustrate the Structural Hypothesis and, more precisely, the Demarcative Hypothesis.
As for the Copula Hypothesis, it has been recently explored by Owens (2006) and holds that the oscillation of *-n* between a definite and indefinite meaning is only a deceiving effect of the fact that this element in and by itself has no *semantic* import, its function being merely *syntactic*, ie that of signalling the relationship between a noun and its modifier.
In this light, definiteness / indefiniteness (or, more precisely, specificity) are a function of the context in which *-n* is found, and this latter is to be identified with a copula acting in the noun phrase rather than in the sentence, and known in the typological literature under various labels: LINKER, LIGATURE, ASSOCIATIVE PARTICLE and so on.
Notice also that the Copula Hypothesis is not incompatible with the Specificity Hypothesis.
Textual research by Ayoub (1991, 2006) demonstrates that Arab Grammarians (and hence native speakers) interpret *-n* more along the lines of the Structural Hypothesis than of the Semantic Hypothesis: in particular, analysis of the primary sources seems to lead Ayoub to espouse the Demarcative Hypothesis (since *n* signals the 'nominalitè du nom', ie distinguishes it from the other parts of discourse).
To summarize, diachronic and syntactic considerations make the Structural Hypothesis preferable to the Semantic Hypothesis. Nonetheless, taking apart the issue of the relationship between *-n* and *-m*, which is of little interest here, the Structural Hypothesis raises a non-trivial problems: how does a copular element turn into a marker of definiteness and / or indefiniteness?
Another problem lies in the compatibility between the two kinds of Structural Hypothesis: why does the noun associated with the copula *-n* stand alone under certain circumstances (Demarcative Hypothesis), if the copula by definition requires a modifier - predicate (Copula Hypothesis)?
Clearly, the validity of the Structural Hypothesis rests on its ability to get around the just mentioned problems.

3.3.5 Adjective

The analytical trends of the Arabic adjective can be subsumed under three main heads:

- Adjective Hypothesis
- Apposition Hypothesis
- Relative Clause Hypothesis

The Adjective Hypothesis is adopted not only in the traditional accounts of the Arabic Noun Phrase, but also in the most recent generative paradigm (Cinque, 2003; Shlonsky, 2004, and references therein).
It basically holds that the Arabic adjective is to be likened to the Indoeuropean one, except

for a different choice of the agreement markers to be copied onto it (e.g. in Italian, but not in Classical Arabic, the number marker is copied onto the adjective when this latter is PL and inanimate; in Classical Arabic, but not in English, the article is copied onto the adjective etc...). In other words, the Adjective Hypothesis crucially relies upon the Feature-Copying Hypothesis which however does not seem to be an appropriate characterization of the Arabic Adjective, as discussed at length in Section 3.2.2.1.

The Apposition Hypothesis and the Relative Clause Hypothesis have also been dealt with in the same Section.

Of all the just mentioned proposals, the Relative Clause Hypothesis, which has been recently couched within a generative framework by Kremers (2003), seems to be the most adequate one: This is so not only because it can provide an unified account of the Relative Clause and the adjective, but also because it gives a principled explanation to the presence of the adjectival article *l-*, which is likened to a COMPLEMENTIZER[59] (Kremers, 2003, pp. 7, 9).

It should be also noted that the Relative Clause Hypothesis, once adopted, absorbs the Apposition Hypothesis, because of the traditional claim that the apposition is basically a subtype of (non-restrictive) relative clause.

This approach, however, leaves some issues unsettled. Apart from the identification of the syntactic position of the covert pronoun, Kremers (2003)'s interpretation of the adjectival article *l-* is stipulative, in the sense that it is not corroborated by some syntactic diagnostics.

It is also expected that the characterization of this type of relative clause be compatible with that of other types, namely the relative clause introduced by *llaḏī, illi* etc... and, on Bravmann (1961)'s view, the Construct State (see Section 3.3.6).

3.3.6 Construct State and Free State

The research about these phenomena can be subsumed under three main heads:

- Copula Hypothesis
- Transformational Hypothesis
- Drift Hypothesis

In the Transformational Hypothesis framework, two main analyses can be distinguished:

- Grammaticalization Hypothesis
- Relative Clause Hypothesis

The Copula Hypothes, argued for by Pennacchietti (1968) maintains that the particle intervening between the modified noun and its complement of specification is a copular element (see Chapter 1).

To this one may add (Buccellati, 1972) that the Copula Hypothesis is a taxonomic approach given that it regards the Construct State and the Free State as two unrelated constructions - in passing, a position widespread also in the generative paradigm (Benmamoun, 2006).

The advantage of Pennacchietti's proposal is its empirical coverage: interpreting the particle introducing the Construct State as a copula provides a unified explanation for constructions found not only in Semitic but also in Hamitic (see (7 - 10) in Section 1.1.3):

[59] This issue will be dealt with in Chapter 6.

(175) Noun (+ Copula) + Genitive Phrase

His account, however, is flawed in an important respect: no attempt is made to explain the native speakers' judgment that the Construct State and the Free State are deeply related, as already discussed in Section 1.1.3 .
According to the Transformational Hypothesis, the Free State undergoes some kind of operation and the outcome of this latter is the Construct State:

(176) Free State → operation X → Construct State

Generally, the operation at issue is the deletion of the element(s) intervening between the modified noun and the complement of specification modifying it (see Section 3.2.2.2).
The level this operation applies to can be either synchronic (ie the speaker's mind) or diachronic. The former interpretation, the Grammaticalization Hypothesis, was first entertained by Arab Grammarians (see Section 3.2.2.2) in dealing with Old / Classical Arabic data, and has been recently evidenced in the generative paradigm by Ouhalla (2010) on the basis of data from modern Colloquial Varieties.
According to Arab Grammarians, the modified noun is in the indefinite state and accordingly what is dropped is not only a preposition, but also the ending *-n, -na, -ni* (a statement that makes sense insofar as Old / Classical Arabic is taken into account).
Recently, Ouhalla (2010) puts forward that the modified noun enters the structure in the definite state, ie preceded by the article, which is subsequently dropped, while the preposition associated with the complement of specification is not phonologically realized for reasons of economy.
Slightly simplifying, in the Free State the particle intervening between the two nouns (e.g. *dyal*: see Table 3.9 in Section 3.2.1) is marked for number / gender via some affixes (e.g. *-u* in *dyal-u*), but the affixes by definition need a stem to which they can be attached: accordingly, the stem (e.g. *dyal*) is given a phonological representation to host the affix *u* etc... (cp. the affix *-es* in the English sentence *does John read the book?*, where *-es* requires the dummy-support *do-*: **-es John read the book?*, in fact, is ungrammatical).
Last but not least, Ouhalla ascribes the presence of markers of gender / number on the particle at issue to its being in diachrony a former noun:

(177) Noun (+ Preposition$_{>Noun}$) + Genitive Phrase

On the other hand, in the Construct State the particle intervening between the two nouns is not marked for number / gender: thus, there occurs no affix in need of a stem hosting it, and then the particle is not phonologically realized.
Turning to the Relative Clause Hypothesis, it has been put forward by Bravmann (1961)[60], which claims that the particle intervening between the two nouns and undergoing deletion is a relative marker on the basis of comparative evidence, ranging from Akkadian to Arabic:

(178) Noun (+ Relative Marker) + Genitive Phrase

All these analyses raise the same question: why is the particle intervening between the two nouns dropped, whatever its nature? Putting aside Arab Grammarians' and Bravmann's proposals, which are vague in explaining the particle-deletion as the consequence of a generic Least

[60] Bravmann's proposal was hinted at by previous scholars, like Friedrich Eduard König at the beginnings of the XX century, and revived in the sixties by Ewald Wagner: see Pennacchietti (1968, pp. 55-58 and references therein) for details.

Effort Principle, even Ouhalla's motivated account does not explain the behavior of particles of the Free State like the Classical Arabic *min*. This latter, in fact, is not marked by number / gender but is nevertheless phonologically realized, contrary to the expectation.
In the remaining respects, the different versions of the Transformational Hypothesis are empirically grounded in diachrony and not mutually exclusive. If Arab Grammarians' preposition of the Free State, in fact, is a (former) noun à la Ouhalla, the complement of specification is reinterpreted as an apposition, a construction that is both pervasive in Arabic (cp. Arab Grammarians' nominal interpretation of the adjective: Section 3.2.2.1) and strictly connected, according to the Western Grammar, with the non-restrictive relative clause (NONRC henceforth). But this is precisely Bravmann's characterization of the complement of specification. To sum up:

(179) Noun [+ $\text{Noun}_{>Preposition}$ + Genitive Phrase]$_{Apposition=NONRC}$

It should be added that in diachrony the modified noun of the Free State is documented both in the indefinite state (noun plus -*n*), as assumed by Arab Grammarians and in the definite state (*n*-less noun), as proposed by Ouhalla and in fact Bravmann (1961, p. 387) admits both the states as the original form of the noun at issue.
In particular, the definite state goes back at least to Akkadian - which strongly suggests that the article preceding the noun in the definite state is not necessarily dropped, but is lacking at all (see (7) in Section 1.1.3).
The Drift Hypothesis reverses the directionality of the process that connects the Free State with the Construct State:

(180) Construct State → operation X → Free State

and therefore assumes that the Free State develops out the Construct State as a consequence of a more general diachronic shift of Arabic from a synthetic type to an analytic type, the driving force of this process being the need of simplification.
This position, advocated among many others by Eksell (1984) and Blau (2002), is flawed at least in two points, highlighted by Owens (2006). On one side, the Free State is not necessarily simpler than the Construct State, because the former construction is marked for gender and number features, absent in the latter and, on the other side, the Free State seems to be at least as old as the Construct State (cp. also Section 1.1.3).
Thus, on closer look, the shift argued for by Eksell (1984) and Blau (2002) seems to be modelled after the linguistic history of Indo-european languages, rather than on the Arabic (and Semitic) data - another instance of logical matrix.
On these grounds, the Drift Hypothesis is very likely to be discarded, in favor of the Transformational Hypothesis.
Notice, however, that this latter, to be proved valid, requires a more accurate description of its fundamental ingredient, ie the (non-restrictive) relative clause - modifier.
It would be also desirable that such an approach be not incompatible with Pennacchietti (1968)'s Copula Hypothesis, given that this latter is able to encompass also Hamitic languages (see (7 - 10) in Section 1.1.3), although Pennacchietti (1968, pp. 55-58) takes issue with the partisans of the Relative Clause Hypothesis and conceives his Copula Hypothesis as alternative to it.

3.3.7 Relative Clause

Three main interpretations of the Arabic relative clause have been put forward in the literature:

- Definiteness Hypothesis
- Restrictiveness Hypothesis
- Copula Hypothesis

The Definiteness Hypothesis and the reasons for rejecting it have been discussed at length in Section 3.2.2.3.
The Restrictiveness Hypothesis has been recently proposed by Gensler (2004), which contends that the relative marker (*llaḏī, illi*) originally signals only a non-restrictive relative clause.
Gensler (2004)'s main argument is the statistical evidence he provides from Old Arabic (Koran). What weakens this claim, however, is that the statistical data are not supported by purely linguistic arguments: Gensler, in fact, provides no clear definition of the categorial status of the relative marker *llaḏī, illi* (pronoun, conjunction, complementizer, to mention the main proposals made in the literature), nor does he develop any syntactic diagnostics for defining it.
Finally, the Copula Hypothesis, due to Pennacchietti (1968), interprets the relative marker as a copula, rather than a pronoun or as a complementizer on the basis of comparative evidence (see (7 - 10) in Section 1.1.3).
The advantage of this proposal lies in its empirical coverage, the sample of languages in which the relative marker / copula is attested including not only Semitic but also Hamitic languages (see also Section 3.3.6).
Nonetheless, the same criticism raised with regard to Gensler's approach applies here. One expects, in fact, that the copular nature of the relative marker *llaḏī, illi* be tested against a syntactic argument, but Pennacchietti does not take up this issue.
On the whole, it seems that the Definiteness Hypothesis is to be abandoned and that the Restrictiveness Hypothesis and / or the Copula Hypothesis could be adopted instead, provided that some syntactic diagnostics for determining the categorial status of the relative marker is given.

3.4 Summary

This chapter has offered a critical review of the main data and hypotheses concerning the Arabic NP, as discussed in the received view and in less traditional threads of research.
A closer and integrated investigation of already known data has resulted in rejecting some commonly entertained hypotheses about the modification-structure (e.g. Feature Copying Hypothesis, Definiteness Hypothesis, (In)definiteness Spreading Hypothesis, Adjective Hypothesis, Drift Hypothesis) in favor of the Specificity Hypothesis and of the Relative Clause Hypothesis. Interestingly, these latter are not mutually exclusive and provide an unified account of the Arabic modifiers, which can be *all* thought of as *specific* and *relative clauses*.
In particular, the nominal appositional nature of the '(extended) adjective' and of the morphemic material introducing the Free State (and, at some level of representation, the Construct State) as well as the contexts of occurrence of the Old Arabic relative clause point to a *non-restrictive* characterization of the Relative Clause Hypothesis (Restrictiveness Hypothesis).

Notice also that the Relative Clause Hypothesis thus formulated absorbs the Apposition Hypothesis and the Transformational Hypothesis worked out, respectively, with respect to the 'adjective' and the 'Construct State' (= a subtype of Free State).
This interpretive trend, however, raises - at least - three questions:

- Can some diagnostics be developed to assign the 'adjectival article' *l-* and the relative marker *llaḏī, illi* a more precise categorial status?
- Which is the exact syntactic position of the subject covert pronoun posited by Arab Grammarians in the 'adjective'-structure?

Nevertheless, the Specificity Hypothesis and the Restrictiveness Hypothesis / Relative Clause Hypothesis are not the only way to provide an unified account of the Arabic NP. Pennacchietti (1968)'s Copula Hypothesis, in fact, albeit flawed in some respects, is able to encompass not only Old Arabic, Classical Arabic as well as the modern Colloquial Varieties (South-Arabian included) but also Hamitic.
It would be therefore desirable that the Copula Hypothesis, given its descriptive power, be not incompatible with the Specificity Hypothesis and the Relative Clause Hypothesis.
Turning to the adnominal markers, and starting with Case, it has been plausibly shown that the Copula Hypothesis and the Language Contact Hypothesis are superior to the others: accordingly, Case can be better understood as a copula borrowed from the pre-Semitic / Hamitic parastrate.
The validity of such an account, however, rests on its ability to answer the following question:

- Which is the exact categorial status of the copula, given that it cannot be taken to be a pronoun?
- How can the allomorphy NOM / ACC, OBL be explained?

As for number and gender, they are plausibly accounted for along the lines of the Unified Hypothesis.
This approach has it that a pluralizing morpheme *-w-* can be not only infixed in the BROKEN PL but also that it is to be identified with the long vowel typical of the SOUND PL, irrespective of its phonological realization. When the *-w-* is not detectable, in fact, it has undergone some process of glide contraction (Agglutination Hypothesis).
It also holds that the SG marker *-at* is to be interpreted as a Classifier.
On these grounds, all of BROKEN PL, SOUND PL, SG derive from the input form COLL.
But some aspects of the Unified Hypothesis still call for an explanation:

- Where are the expected numeral / quantifier in the three term system (160), (161) (see Section 1.1.3)?
- Which is the singulative counterpart of the SOUND PL-ending *-na*?
- How can the Depluralization Hypothesis (145) be integrated within both the Agglutination Hypothesis and the Classifier Hypothesis?

Finally, the adnominal markers *l-, -n* are plausibly to be regarded as elements that originally do not signal definiteness / indefiniteness.

Instead, the marker *-n* seems to act either as a copula (Copula Hypothesis) or as a sign of nominality (Demarcative Hypothesis). The validity of these hypotheses rest on their compatibility:

- Why does the noun associated with the copula *-n* stand alone under certain circumstances (Demarcative Hypothesis), if the copula by definition requires a modifier - predicate (Copula Hypothesis)?

On the other hand, the marker *l-* is arguably an element originally intervening between a noun and its adjectival modifier, along the lines of Wackernagel's law (Clitic Hypothesis / Copula Hypothesis). Etymologically, it is connected with the particle *la* documented in Amorite and in Arabic (which, therefore, inherits *la* from the Amorite parastrate).
This scenario, however, raises the following question:

- Why and how is *l-* subsequently used outside the modification-structure and found at the left of the noun?

An overall overview of the adnominal markers, therefore, point to an analysis in terms of the Copula Hypothesis. On one side, in fact, this proposal has been put forward on independent grounds for Case, *-n* and *l-*; on the other side, the morpheme *-w-* seems to play a crucial role in the expression of number and gender, but at the same time, it is strictly connected with Case, and hence, indirectly, with the Copula Hypothesis.
The picture that emerges from this chapter is that a motivated account of the Arabic NP should take into account four hypotheses, namely

- Restrictiveness Hypothesis / Relative Clause Hypothesis
- Copula Hypothesis
- Classifier Hypothesis
- Agglutination Hypothesis

as well as to answer to the questions they raise and to make more precise the relationships between them.

Part II

A broader perspective

Chapter 4

The minimalist framework: (Dynamic) Antisymmetry

4.1 Introduction

This Chapter lays out the formal framework that will serve in the next Chapters as the basis for the analysis of the empirical data. This is the approach to syntax currently referred to as MINIMALIST PROGRAM, in its essence the most recent incarnation of the generative (/ transformational) grammar (Boeckx, 2006).
Such a development of the paradigm is an attempt to embed the syntactic theory in a cognitive and biological setting (Boeckx, 2006, Ch. 2) which, among others, has the advantage of deriving the otherwise stipulative properties of syntax, as it will be seen in Section 4.5.
In particular, the very notion of program enclosed in the term 'Minimalist Program' is to be understood in the sense ascribed to it by the philosopher and historian of science Imre Lakatos: it is a scientific mode of investigation following broad guidelines. This flexibility leaves room for the formulation of a great many theories. Accordingly, in the Minimalist Program (MP henceforth), one finds several and sometimes overlapping theories revolving around two central concerns (Boeckx, 2006, Ch. 1):

- which are the most primitive elements ('atoms') of syntax?
- why do atoms of syntax undergo displacement (cp. fn. 57)?

Some research threads of MP answer the former question reviving the so-called ITEM-AND-ARRANGEMENT MODEL, which was worked out originally to account for the agglutinative languages (Thornton, 2005, Ch. 6). This idea has recently informed the work of the Italian 'cartographers' (see Cinque & Rizzi 2008 and references therein) but, as Thornton (2005, *Ibidem*) and Baker (2002) show, it was already pursued by Baker (1988) and goes back to the very beginnings of generative grammar (Chomsky, 1957).
Other threads answer the latter question relating it to the fact that human speech has a temporal aspect, the Saussurean LINEARITY: words ('atoms') follow each other and it is exactly this property that triggers displacement. This claim, first made in Moro (2000) is designated as DYNAMIC ANTISYMMETRY.

This is but a sample of the many proposals put forward in the minimalist literature and is by no means exhaustive.
The just mentioned hypotheses, however, stand out as particular appealing because of a subtle link between them: both of them are deeply rooted in one of the most influential instantiations of MP, namely Kayne (1994)'s ANTISYMMETRY. Therefore, a strict application of the Kaynean antisymmetry implies combination of two apparently unrelated hypotheses and ensures consistency between them: the item-and-arrangement model (à la Chomsky 1957, Baker 1988, Cinque & Rizzi 2008) and Moro (2000)'s Dynamic Antisymmetry.
Before concluding this Section, a caveat is in order for those who are familiar with versions of generative grammar previous to MP: in this framework, Deep Structure and Surface Structure are dispensed with. This topic will be discussed in greater detail in due course - for the time being, suffice it to say that this happens because MP is a partial return to the technical apparatus of (Chomsky, 1957)'s *Syntactic Structures*, where these two theoretical constructs had not yet been introduced (Boskovic & Lasnik, 2007, p. 6).
In a likewise fashion, those familiar with the so-called PRINCIPLES-AND-PARAMETERS MODEL should be aware that it will not be dealt with in the present work[1]. The item-and-arrangement model, in fact, does not necessarily hinge on this model as it shown by the fact that Chomsky (1957), the theoretical core of MP, makes use of the former but not of the latter. Moreover, the Principles-and-Parameters model, while compatible with Dynamic Antisymmetry in its main lines, is not crucial to it (Moro, 2000, Ch. 4).
The remainder of this Chapter aims at offering an overview of the just outlined minimalist scenario, focusing on its main ingredients: *Syntactic Structures* and (Dynamic) Antisymmetry. It will be shown in the next Chapters that this version of MP is empirically grounded in the Arabic data.

4.2 Syntax, lexicon and the item-and-arrangement model

The present work is primarily concerned with the (MORPHO-)SYNTAX of the Arabic NP.
This term is to be understood here as both

- the way how (at least two) morphemes combine to yield a word.
- the way how (at least two) words combine to yield a larger complex (a PHRASE), and

In other words, the difference between a MORPHEME and a WORD is not relevant for this work. It is customarily assumed, in fact, that a morpheme is defined on the basis of its meaning: it is the minimal unit conveying a piece of information. On the other hand, the definition of a word is *regardless* of meaning, rather relying upon its form: it is the minimal unit that can be surrounded by breath-pauses in the stream of speech (ISOLABILITY) and that cannot be broken up by any intervening material, e.g. another word (UNINTERRUPTABILITY) - contrast *dogs* with **dog- three -s*[2].
Therefore, it is in principle possible that a minimal meaningful unit has a form such that it

[1] But see Section 8.10.2

[2] This is a working definition of morpheme and word, which summarizes Lyons (1968, pp. 180-208) and Thornton (2005, Ch. 1) and is not meant to be exhaustive: a satisfactory characterization of these entities, in fact, is still lacking and matter of debate: see the above mentioned works for more discussion and references.

stands in isolation and resists insertion of material in its middle (cp. English *in*). In this case, what is a morpheme (from the standpoint of meaning) is also a word (from the standpoint of form): since Bloomfield (1933) onwards, the morpheme functioning as a word is referred to as FREE MORPHEME.

When the morpheme does not have this ability (e.g. the PL-ending *-s* in English: **He saw three -s*), it is forced to combine with one or more morpheme(s) to yield a word (*dog + s > dogs*). In this case, it is said to be a BOUND MORPHEME.

(Morpho-)syntax in the present work is precisely the way how *-s* combines with *dog* to yield *dogs* (morpheme + morpheme), but also the way *full-blooded* combines with *dogs* to yield *full-blooded dogs* (word + word).

To summarize:

(181) cough + ing → (BOUND) MORPHEME + (BOUND) MORPHEME: WORD
cough → (FREE) MORPHEME: WORD

A less familiar example is the pair SG / COLL *be'u:da / be'u:d* ('gnat / a swarm of gnats') in Omani Arabic, reported by Greenberg (1990b, p. 179):

(182) be'u:d + a → (BOUND) MORPHEME + (BOUND) MORPHEME: WORD
be'u:d → (FREE) MORPHEME: WORD

This approach, worked out by Bloomfield (1933, Ch. XIII), is usually designated as ITEM-AND-ARRANGEMENT MODEL, where ITEM stands for morpheme and ARRANGEMENT for combination (Lyons, 1968; Thornton, 2005, and references therein).

This model revives the traditional opposition between the 'material' subject to combination, the so-called LEXICON, and the operation of combination in and by itself, referred to as 'GRAMMAR': accordingly, morphemes / items can be also designates as LEXICAL ITEMS, while (morpho-)syntax / arrangement is the more recent equivalent of 'grammar' (Lyons, 1968).

In the traditional accounts, in fact, the term syntax is restricted to the combination of words, combination of morphemes being referred to as MORPHOLOGY.

On this view, morphology would be further split into two interacting oppositions: SUBSTANTIVE vs. FUNCTIONAL and DERIVATIONAL vs. INFLECTIONAL.

On one side, substantive morphemes denote in and by themselves the entities ('substances') spoken about, while functional ones provide more precise information about them. Thus, in the word *be'u:d-a*, the substantive morpheme *be'u:d* refers to a kind of insect that the speaker has in mind, while the functional morpheme *-a* precises that the insect at issue is only one.

But a functional morpheme appears to be further distinguishable into DERIVATIONAL and INFLECTIONAL. Derivation is combining a substantive morpheme with a functional morpheme to express a new notion: e.g. the noun *Kashmir*, referring to a *country*, if combined with the morpheme of nationality *-ian* yields *Kashmirian*, which denotes its *inhabitant*. Inflection is combining a substantive morpheme with a functional morpheme to express some kind of quality, e.g. plurality (*dog + s* to yield *dogs*) - hence, Arabic adnominal markers of case, gender, number, definiteness in the received view are seen as inflectional morphemes (the quality denoted by case being subjecthood etc...).

The item-and-arrangement model developed in Chomsky (1957, p. 104) keeps to the traditional distinction between SUBSTANTIVE and FUNCTIONAL morphemes but does not consider it as restricted to morphology, since both the kinds of morpheme according to him do occur as words,

ie in the traditional syntax: cp. the Arabic and English 'substantive' COLL *be'u:d*, *fish* and the English 'functional' preposition *in* as well as its Arabic equivalent *fī*.
Furthermore, the item-and-arrangement model à la Chomsky (1957, p. 105) does not seem to go beyond the traditional distinction between SUBSTANTIVE and FUNCTIONAL morphemes[3].
This is arguably due, first and foremost, to the fact that both derivational and inflectional morphemes share a salient property opposing them as a whole to the substantive morphemes (Lyons, 1968, pp. 435-437). The latter, in fact, are an open set, since the entities of the universe of discourse they refer to are virtually infinite (not only a *dog*, but also, in modern times, a *robot*), while the former are a *closed* set, that is one of fixed and small membership. For example, the derivational morpheme of nationality in English has typically a half dozen values (*-ish / -an / -ian / -ese / -i*: cp. *English, American, Kashmirian, Lebanese, Bengali*), the inflectional morpheme of gender in German has only three values (M / F / N).
To this, one may add that it is not always easy to draw a clear-cut distinction between derivation and inflection, given that in Arabic and sometimes in English the *derivational* morpheme of nationality works as the singulative of a COLL, ie as an *inflectional* morpheme of number, a fact recognized by both native speakers of English (Greenberg, 1990b, p. 183) and Arabic (Fleisch, 1961, vol. II, p. 306):

(183) Irish → Irish-man
COLL → COLL-NATIONALITY / SG

(184) Yūnān → Yūnān-iyy 'Greeks → Greek'
COLL → COLL-NATIONALITY / SG

To sum up, the item-and-arrangement model revives the traditional opposition lexicon vs. grammar, refining these notions.
As for the lexicon, it is made up of lexical items, which sometimes appear in isolation (words) and can be further distinguished into substantive (= open-class) morphemes and functional (= close-class) morphemes: consequently, '(lexical) item' will be used henceforth as a cover term for both morphemes and words.
As for the grammar (renamed '(morpho-)syntax'), it combines the primitive elements provided by the lexicon - technically, syntax IS FED by the lexicon.

4.3 Form and meaning

The definition of morpheme given on p. 152 can be made more precise by observing that the piece of information (technically, FEATURE or SEMEME) conveyed by it is actually assigned a sound string (the MORPH):

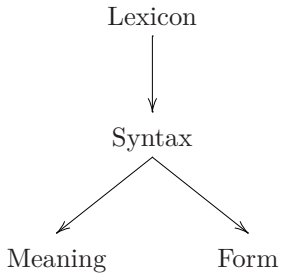
(185) be'u:d-a = MORPH_1 - MORPH_2
gnat-SG = FEATURE_1 - FEATURE_2

As Thornton (2005, pp. 165-167) remarks, implicit in the item-and-arrangement model exemplified in (185) is the assumption that the traditional concept that a given word has a meaning and a sound is interpreted in COMPOSITIONAL terms.

[3] 'Open-endedness' and 'grammatically functioning' categories in Chomsky (1957, pp. 104-105)'s terminology. See also Thornton (2005, pp. 96-100) for critical discussion

Under this perspective, meaning and form are the result of COMPOSITION (= combination) of more primitive ingredients, morph and features, respectively, a property indicated in the literature as DISCRETENESS (Moro, 2008, Ch. 1)[4].
Accordingly, a one-to-one correspondence can be set between a morph and a feature (ONE FORM - ONE MEANING PRINCIPLE). Such a state of affairs can be illustrated as follows:

(186)



There are some reasons, however, to believe that the compositional hypothesis should be given a weak formulation, since it may occur that meaning can be decomposed into simpler ingredients, but form cannot, to the effect that a one-to-one correspondence between morphs and features is not found.
This is exemplified by the following piece of data, reported in Thornton (2005, p. 170): in Italian the word *ladro* 'thief' is overwhelmingly more used than its rare synonym *rubatore*, although the former is far less compliant with the item-and-arrangement model (185, 186) than the latter:

(187) ladr-o = MORPH$_1$ - MORPH$_2$
steal.A-SG.M = FEATURE$_1$ - FEATURE$_2$ - FEATURE$_3$ - FEATURE$_4$

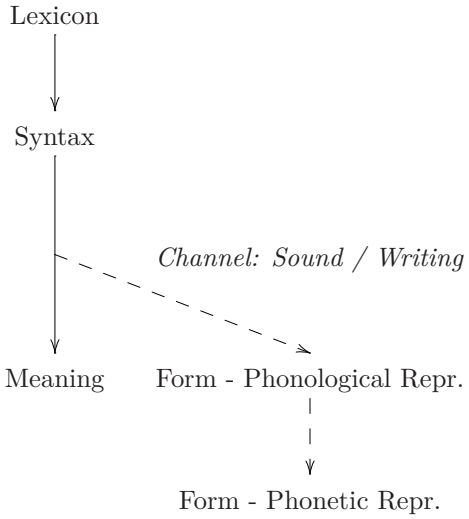
(188) rub-at-or-e = MORPH$_1$ - MORPH$_2$ - MORPH$_3$ - MORPH$_4$
steal-A-M-SG = FEATURE$_1$ - FEATURE$_2$ - FEATURE$_3$ - FEATURE$_4$

The explanation of this anomaly plausibly is that discreteness is a crucial property of meaning but not of form: other properties are more relevant in this latter, having to do with pronunciation (or 'SPELL-OUT'). Simplifying Marotta (1999)'s account, in Italian the OPTIMAL LENGTH of a word is two / three syllables and therefore the bysillabic *ladro* is preferred to the quadrisyllabic *rubatore* on a phonological rather than syntactic basis (Thornton, 2005, p. 170).
A similar case is in English the irregular past form *took*, occurring instead of the more transparent *'taked'* (Chomsky, 1957, pp. 33ff.).
To account for this state of affairs, Chomsky (1957) and Chomsky & Halle (1968) introduces the notion of MORPHOPHONEMIC RULES: in this view, form is a set of rules made up of two subsets. The former subset associates to each feature a morph, e.g. *taked*, which yields a PHONOLOGICAL REPRESENTATION: this level still manifests discreteness and actually is the item-and-arrangement model proper Thornton (2005, Ch. 6). The latter subset of rules 'schizophrenically' destroys the work of the previous subset affecting the phonological representation with

[4]Or MORPHOSEMANTIC TRANSPARENCY, in Thornton (2005)'s terminology. This view seems to go back to the first modern modelization of the opposition form - meaning: Saussure (1916, p. 156), in fact, before introducing it states that both meaning and form can be thought of as 'une série de subdivisions contiguës dessinées à la fois sur le plan indéfini des idées confuses (A) et sur celui non moins indéterminé des sons (B)'('a series of adjoining subdivisions simultaneously imprinted both on the plane of vague, amorphous thought (A), and on the equally featureless plane of sound (B)' in Harris's (1986, p. 110) translation).
For the non-compositional meaning (idioms etc...), see Section 4.5.2.

some sound change, e.g. *took* etc..., which results in a PHONETIC REPRESENTATION.
Simply put, form is more distant from syntax than meaning, and additional evidence in favor of this claim is that form, as realized in the human language (acoustic or written material) is by no means necessary to express (whatever) meaning: as Saussure (1916, p. 103) and Bloomfield (1933, p. 28) point out, visual signals like ships' flags, and electrochemical waves of the nervous system equally well serve the same purpose. Pushing this reasoning to the limit, one may even agree with Chomsky (1995, p. 221) that in the best case syntax would put together features to express meaning without no need of form: 'If humans could communicate by telepathy, there would be no need for a phonological component, at least for the purposes of communication'.
On these grounds, (186) is replaced by (189):

(189)



The different communication channel used to assign a form to the meaning crucially impacts on both of them: in utterance (or writing) it is *sound* and hence words are presented and captured *one after the other in time*, while ships' flags rely upon vision, to the effect that they are presented and grasped *simultaneously*. This property from Saussure (1916, p. 103) has been referred to as LINEARITY (see also Moro 2008, Ch. 1).
On the whole, a given SYNTACTIC STRUCTURE (combination of two or more items) passes through the steps depicted in (189) in order to be fully realized. This process has been designated from Chomsky (1957, Ch. 4) onwards as DERIVATION, an usage of this term not to be confused with that found in the traditional grammar (see Section 4.1).

4.4 More on syntax: the generative and transformational components

Bloomfield (1933, p. 195) observes that the phrase resulting from the combination of two items has two interesting properties, dubbed in subsequent works as PHRASE STRUCTURE RULES:

(I) ENDOCENTRICITY: in some intuitive sense, an item (call it the HEAD) is not on the same footing of another (call it the DEPENDENT).

(II) ability to be organized in RANKS: the phrase combines with another item to form a larger phrase

Endocentricity (I) can be tested trying to replace the entire phrase with one of the items it is made up of (SUBSTITUTION TEST): the item functioning as an Head can replace the phrase while that functioning as a Dependent cannot, as exemplified in (190).
The ability to be organized in ranks (II) is illustrated in (191):

(190) English (Bloomfield, 1933, p. 195)

[poor John]	→	John: OK / *poor: NO
PHRASE	Substitution	HEAD DEPENDENT

(191) English (Bloomfield, 1933, p. 195)

a.

fresh	+ milk	→	[fresh milk]
DEPENDENT	+ HEAD	Combination	PHRASE

b.

very	+ [fresh milk]	→	[very [fresh milk]]
DEPENDENT	+ HEAD	Combination	PHRASE

Chomsky (1970) makes more precise Bloomfield's notion of endocentricity, observing that a Dependent can both precede or follow its Head:

(192) English (Chomsky, 1970, p. 199)

[John's proof of the theorem]	→	proof: OK / *John's, of the theorem: NO
PHRASE	Substitution	HEAD DEPENDENT

The item-and-arrangement model à la Bloomfield in characterizing syntax in terms of phrase-structure rules, however, has an intrinsic limit: it does not account for the native speaker's perception that a *same* item can be found in *different* positions (Chomsky 1957, Ch. 8, Thornton 2005, Ch. 6 and fn. 57).
A case in point is the English morpheme 3SG *-(e)s*, if the declarative sentence *John eats an apple* is contrasted with its interrogative counterpart *does John eat an apple?*.
Another telling example (already known to traditional grammar) is the word *John* understood as a (logical) subject, if the active sentence *John eats an apple* is contrasted with its passive counterpart *An apple is eaten by John*.
To explain this phenomenon, Chomsky (1957, p. 46) ascribes an additional property to syntax:

(III) TRANSFORMATION: ability to 'rearrange' items

To recap, in a revised version of the item-and-arrangement model, syntax ('grammar') is better seen as made up of a phrase-structure component, characterized by endocentricity and ranks, as well as of a transformational component, responsible for rearrangement of items.
In subsequent developments of Chomsky (1957)'s work, the former component is referred to as GENERATIVE component, endocentricity as HIERARCHY, ranks as RECURSION, while, as far as the transformational component is concerned, rearrangement is designated either as DISPLACEMENT or, more commonly, as MOVEMENT (Moro, 2008, Ch. 1).
Recall also from Section 4.3 and (189) that the syntactic structure produced by these two components is *soundless*, and only later is it assigned a 'sound' via the morphophonemic rules.
Before concluding this Section, something else should be said about how the generative and transformational components are thought of within the generative tradition.
As pointed out in Section 4.1, MP is a return to a *status quo ante* of the theory, and therefore

implements these two notions along the lines of Chomsky (1957) rather than within the influential framework of the GOVERNMENT AND BINDING THEORY. This theoretical choice impacts on the characterization of both the generative and transformational components.
To put it more concretely, MP and Government and Binding Theory share the idea that syntax is the way how items *combine* and *move* but sensibly differ in the account they put forward for these phenomena. In Government and Binding Theory, in fact, combination and movement of items are

- all-at-once
- scaled in time

operations: *first*, *all* the items are combined in the generative component, *then*, *all* the instances of displacement take place in the transformational component. The generative component (or 'phrase structure rules') and the transformational components are kept distinct and when thus characterized are referred to as DEEP STRUCTURE and SURFACE STRUCTURE, respectively. MP develops Chomsky (1957) and argues instead that combination and movement of items - the generative and transformational components - are *not* separate: rather, they interact, applying

- step by step
- 'on line' (simultaneously)

For example, Government and Binding Theory assigns the sentence:

(193) Which story does John believe that Mary knows?

the following Deep Structure:

(194) John + does + believe + that + Mary + knows + which + story?

(cp. its emphatic counterpart: *John does believe that Mary knows this story*). What Deep Structure is doing is just putting together the items of the interrogative sentence according to the same word order SVO also found in its affirmative and emphatic counterpart.
After that combination of the items has taken place in the Deep Structure, they are 'sent' to the Surface Structure, which scrambles ('rearranges') them yielding the word order typical of the interrogative clause, exemplified in (193).
On the other hand, MP tries to handle smaller syntactic structures, e.g. a clause instead of a sentence. For example, suppose that the syntactic structure of the embedded clause of the sentence (193) has been formed:

(195) that + Mary + knows + which + story?

(cp. *(John believes) that Mary knows this story*). MP assumes that the items *which + story* are displaced at the beginning of the embedded clause as much as it happens in its interrogative counterpart:

(196) Which story does Mary know?

This displacement results in a new syntactic structure of the embedded clause, something like:

(197) which + story + that + Mary + knows?

Thus, even if the items of the matrix clause of (193), namely:

(198) John + does + believe

have no longer been put together with the embedded clause, displacement has already taken place. In other words, combination of items is only partially realized, but this does not prevent movement from applying. Movement too is partial, because it has not yet affected the item *does* (contrast (198)with (193)). Hence, under this approach, the two operations take place step by step and simultaneously.
This scenario also implies that in order to obtain (193), some other combination and movement of items is needed. Accordingly, the derivation proceeds as follows: the items of the matrix clause are combined with the embedded clase, giving rise to

(199) John + does + believe + which + story + that + Mary + knows?

At this point, the item *does* is displaced before the item *John*:

(200) does + John + believe + which + story + that + Mary + knows?

The syntactic structure illustrated in (200), in fact, is highly reminiscent of an interrogative English sentence in which no interrogative element occurs:

(201) does John believe him?

Displacement of *does*, therefore, is independent on the interrogative element *which story*, and hence the former is not displaced along with the latter (contrary to what was assumed to happen in the Surface Structure). Finally, *which story* of (200) is displaced before *does* along similar lines.
In what follows, when speaking about combination and / or movement of items, these phenomena are to be understood along the lines of (*Syntactic Structures* and) MP, not of Government and Binding Theory.

4.5 The Minimalist Program: the biological foundations of language

MP looks at the linguistic phenomena discussed so far from a broader perspective: the study of the biological mechanisms underlying human mind. On this view, language would be the result of the interplay among some MODULES and LEXICA (Fodor, 1983; Chomsky et al., 2002, 2005; Scheer, 2007).
These notions (and in particular 'lexicon') are *not* to be understood in a linguistic sense: modules and lexica are biological systems designed for a given function - for example, they are also involved in the human sight.
More precisely, a module or COMPUTATIONAL SYSTEM is the entity that takes a given input, performs some kind of operation over it ('PROCESSES it') and returns an output.
The lexicon is the entity that provides (FEEDS) the module with the pieces of information serving as its input. In particular, the module has been characterized as DOMAIN-SPECIFIC and ENCAPSULATED.
This means that a given module A has a proprietary lexicon A: accordingly, it only 'sees'

(READS, INTERPRETS) the input A provided by the lexicon A and performs its operations only over this input A (DOMAIN-SPECIFICITY). In addition to that, the module only performs its operations over information that has been present in the input, a property known as ENCAPSULATION: no information *can be added or eliminated* in the course of the computation. In the (pre-)minimalist literature encapsulation is generally designated as STRUCTURE PRESERVATION PRINCIPLE or, more recently, as INCLUSIVENESS CONDITION (see e.g. Haegeman 1991, Ch. 3 and Chomsky 1995, Ch. 4).

To put it shortly, a module can receive inputs from only one dedicated lexicon.

In this respect, a module (say, A) sensibly differs from its lexicon (A), which can receive inputs from many different modules (say, B, C, D). Recall, however, that the lexicon A must also be endowed with an input A, that is the only kind of input that the module A 'sees'.

Another difference between a module and a lexicon lies in their tasks: a lexicon does not carry out any computation, that is does not carry out any operation over any input to yield an output. What the lexicon (e.g. A) does is no more than *associating* its own input (A) with one of the just mentioned external non-specific inputs (say, B).

Remarkably, a beneficial side-effect of this association is that the non-specific external input B, once linked with the domain-specific input A, becomes part and parcel of this latter, to the effect that these two inputs give rise to a kind of 'enlarged' input A:

> "for a given module, however, the origin and the language of the stimulus that has provoked the injection of some element is neutralised: it only sees the output side of its lexicon."
> (Scheer, 2007, p. 11)

The outcome of the association taking place in the lexicon - CROSS-MODAL LINKAGE, in Fodor (1983, p. 132, fn. 13)'s terminology - is that the module A can process a *more complex* load of information, not limited to its proprietary lexicon: actually, it can process information of whatever other module.

A case in point is phonology: if taken to be a module (see Scheer 2007 and references therein), it has the ability to carry out some operation (e.g. Vowel-deletion) over its own input. Scheer (2007, and references therein) brings evidence that such an input (call it 'p') is what is traditionally referred to as consonant or vocal. Of course, consonants and vocals have different acoustic realizations (*p, t, k, a, i, u ...*): according to Scheer (2007, p. 11), this is the piece of information originating from a different module, the acoustic-phonetic one. The phonetic content of consonants and vowels can be thus conceived as the external non-specific input (call it 'a') of the lexicon of phonology. The Fig. 4.1 summarizes and exemplifies the discussion so far.

A more accurate examination of the association taking place in the lexicon of phonology brings out that not all the information originally conveyed by the external input gets associated with the module-specific input. Actually, the phonology-specific input 'consonant' receives only part of the phonetic information encoded in the external input, say, *k*: in (some dialects of) English the pronunciation of *kit* and *skill* are slightly different, because the former is aspirated, the latter is not - nevertheless both of them are conceptualized (and written) as '*k*'. In traditional phonology, the aspirated *k* of *kit* (as found in pronunciation) is said to be a PHONE, its idealized counterpart (as found in writing) a PHONEME.

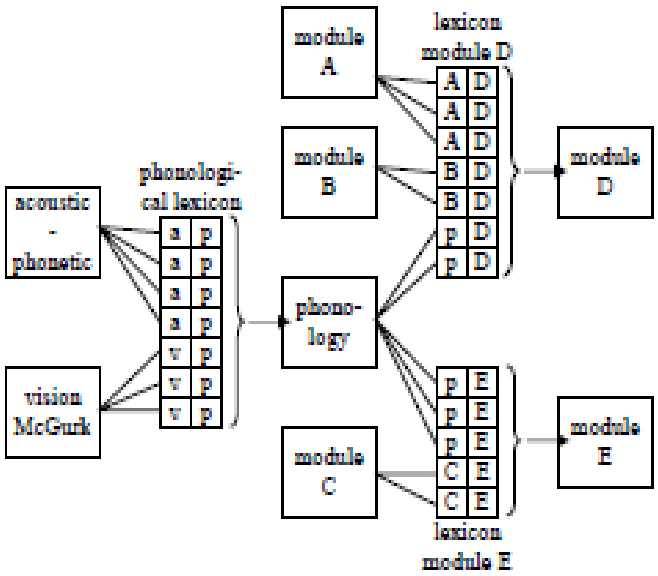


Figure 4.1: Lexica and modules (Scheer, 2007, p. 11)

Under the current approach, this difference can be explained as a partial loss of information occurring when an external input (e.g. *k*) gets associated with the domain-specific input (e.g. the consonant), a phenomenon called INFORMATION BOTTLENECK (Scheer, 2007, p. 9 and references therein).

The just cited example thus unveils another property of lexica: only part of the external input associates with the domain-specific input (PARTIALITY).

The biological and cognitive scenario sketched out in this Section is generally referred to as MODULARITY HYPOTHESIS. Adopting this approach, two properties are distinctive of modules: domain-specificity and encapsulation. On the other hand, the proper of a lexicon is association and partiality.

4.5.1 Syntax as a human faculty and recursion

Capitalizing on the Modularity Hypothesis, MP takes language to be a faculty of human mind very much as sight: accordingly, it comprises at its very core a lexicon, ie the Lexicon of traditional grammar, and a module, ie Syntax. In this light, Syntax might plausibly carry out some kind of operation over (pieces of information of) the Lexicon.

To define the operation in question, it seems convenient to remark that Syntax is a key-ingredient of human language only subsequent to some evolutionary process that exapted Syntax away from its original adaptive function (Chomsky et al., 2002).

There are good reasons to think (Chomsky et al., 2002) that this original adaptive function was 'spatial navigation and foraging, where problems of optimal search are relevant'. In particular, it seems that these activities involved a very simple strategy of space mapping, meant to help the brain / mind to recall locations searched (feeding sites).

Under the Modularity Hypothesis, the function of spatial navigation and foraging is served by

a dedicated module, and the space mapping strategy that drives these actions is the operation performed by such a module (Chomsky et al., 2002, p. 1578). Accordingly, this module will be referred to here as 'module for spatial navigation and foraging (MSNF)'.
In the spirit of Collett & S. (2000, p. 257), the space mapping strategy at issue can be represented concretely as one that sets a conceptual centre and divides the space into circles constructed around it, in a way highly reminiscent of the CONTOUR LINES METHOD in present-day topography, as illustrated in Fig. 4.2 (see Chomsky et al. 2002 for a less informal description).
According to Chomsky et al. (2002, 2005), the main advantage of this strategy lies in that

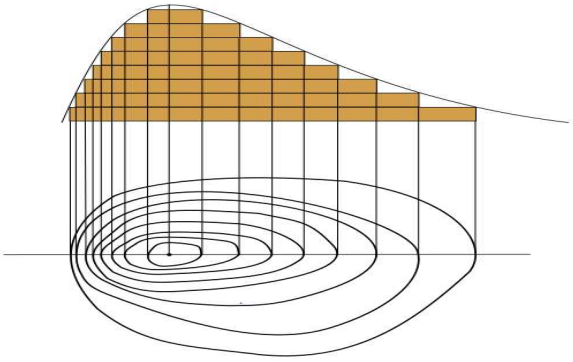


Figure 4.2: Contour lines method in topography (Source: Wikipedia, s.v. *Contour line*)

every circle constructed around the centre works as the SUCCESSOR FUNCTION in mathematics: this latter is a function that simply adds 1 to its input (x + 1) but, in so doing, takes a *finite* set of discrete elements (x, 1) to yield a potentially *infinite* array of discrete expressions (x + 1 = y, y + 1 = z, z + 1 = w etc...). This property of the successor function is RECURSION.
On these grounds, two feeding sites abstracted into the concentric circles of Fig. 4.2 can be conceptualized as follows:

(202) [[X]Y]...
X, Y = location (feeding site)
[ ] = spatial region (physical piece of space)

More accurately, the (concepts of) locations / feeding sites are the inputs feeding MSNF.
The MSNF is still found in primates, birds and insects where, interestingly, it seems to be subject to *internal* constraints imposed by organic limitations such as locomotion and memory costs, the so-called CONDITIONS OF OPTIMAL DESIGN (Chomsky et al., 2002, 2005).
For example primates, when reaching a location in which food is not found, look for another feeding site continuing in a new direction and do not return on the covered path, thus avoiding to reach an already exploited feeding site (Garber & Jelinek, 2006, p. 302 and references therein). In other words, the circles / feeding sites preceding the circle in which the primate is located are no longer relevant and are considered as an undifferentiated spatial region. This is also evidenced by the fact that honeybees and desert ants return directly home from a feeding site, even if it is located several kilometres from the hive, skipping intermediate locations (Collett et al., 2006). This behavior is technically defined as NO BACKTRACKING (Chomsky et al. 2002, p. 1578,Chomsky et al. 2005, p. 187) and is illustrated in (203):

(203) [[X Y]Z]...

X, Y, Z = location (feeding site)

[ ] = spatial region (physical piece of space)

Furthermore, the same locomotion and memory costs lead primates to 'minimizing distances travelled between food patches' (Garber, 1987, p. 349), so that between two feeding sites, the closest one is chosen. This is a kind of 'minimal search' condition (Chomsky et al., 2002, p. 1578) that can be represented as follows:

(204) [[[X] Y] Z] ...
YES *NO*

A comparison between the space mapping strategy of MSNF depicted in (202) and the ranks observed by Bloomfield in Syntax (e.g. *milk* > *fresh milk* > *very fresh milk*), here repeated as (205):

(205) [[fresh] milk]...

very [fresh milk]]...

milk, fresh, very = items

[ ] = syntactic structure

reveals that Syntax conceptualizes the syntactic structure very much as MSNF conceptualizes space. Notice in particular that in (205) *fresh milk* is an undifferentiated whole relative to *very*, as much as in (203) the locations X, Y are relative to Z. In fact, the expression *fresh milk* functions as a *single* Head (see Section 4.4) and the locations X, Y are conceptualized as a *single* spatial region.

In mathematics, this property is captured by rewriting a successor function like [(1 + 1) + 1] as (2 + 1).

It ensues that both space and the syntactic structure are constructed around a conceptual centre, which allows the *infinite* conceptualization of *discrete* locations and items (Recursion). The comparison between (202) and (205) arguably demonstrates that human Syntax is no more than MSNF found also in other primates, birds, insects etc...: a *module* performing the operation of recursion.

Human Syntax, however, differs from MSNF in one crucial respect, namely the *lexicon* that feeds it: not locations but what is traditionally referred to as *words*, ie items (Chomsky et al., 2002, p. 1578)[5].

Some versions of MP, Antisymmetry and Cartography (see Cinque & Rizzi 2008, p. 50, Julien 2002, p. 12 but see also Chomsky 1995, p. 33ff.) make more precise this Fodorian scenario and therefore give the items feeding the module Syntax a characterization along the lines of the previous Section.

In consequence of this, each item is taken to be an external input *associated* with a module-specific input, the so-called (MORPHO)SYNTACTIC CATEGORY, corresponding to the 'PART OF DISCOURSE' of the traditional grammar. Thus, the item *of* is associated with the part of discourse 'preposition':

[5]And numbers too: as discussed above, when these latter feed MSNF, the successor function obtains. For example, the integer list is constructed on the basis of recursion: 1 + 1 = 2, 2 + 1 = 3, 3 + 1 = 4 etc... (Chomsky et al., 2002, p. 1577).

(206) LEXICON = EXTERNAL INPUT$_z$ $\Leftrightarrow$ [MODULE-SPECIFIC INPUT$_i$] > [MODULE$_i$]
Lexicon = item $\Leftrightarrow$ [Synt. Category$_i$] > [Syntax$_i$]

These research threads of MP, in fact, bring evidence in favor of the claim that a very close match must hold between the item and the syntactic category: as Julien (2002, p. 165, 301) puts it 'the terminal nodes are paired with appropriate lexical items', even in the problematic case of idioms: 'the elements that an idiom consists of are matched individually with appropriate syntactic terminals'.
In the (pre-)minimalist literature, the just described association between an external input (the item of the item-and-arrangement model) and a Syntax-specific input is stated as the following principle (Haegeman, 1991, p. 55):

Minimalist definition 4.1 (Projection Principle)**.**
Lexical information is syntactically represented.

A more general formulation of the Projection Principle, which highlights that this architectural setting of Syntax is actually derived in biological terms, would be:

(207) EXTERNAL INPUT : MODULE-SPECIFIC INPUT
: = association

Once that the input of Syntax has been construed, Lexicon feeds Syntax with it and Syntax performs the operation of recursion over it.
The simplest manifestation of recursion is combination of two items associated with their syntactic categories (e.g. *fresh*$_{ADJ}$ + *milk*$_N$ > *[fresh milk]*$_N$), which outcomes in a syntactic structure: in modular terms, this is the output of the module Syntax. Given that each item is actually a pair morph - feature (see Section 4.3), the output of Syntax can be represented as follows:

(208) a. [MORPH + MORPH] Sound
b. [FEATURE, SYNT. CATEGORY + FEATURE, SYNT. CATEGORY] Meaning

Biologically-oriented considerations, moreover, lead to expect that items, once recursively assembled along the lines of (208), are affected by the 'minimal search' condition (204), which in the minimalist literature is either known as RELATIVIZED MINIMALITY after Rizzi (1990), or as LOCALITY (see e.g. Moro 2008, Ch. 1). Actually, this expectation is met, if one considers contrasts like:

(209) English (Haegeman, 1991, p. 401)
?? Which problem do you wonder how John could solve t$_{whichproblem}$ t$_{how}$?

(210) English (Haegeman, 1991, p. 401)
* How do you wonder which problem John could solve t$_{whichproblem}$ t$_{how}$?

(209, 210) show that between the two *wh*-constituents candidate for occupying the sentence-initial position typical of English interrogatives via displacement, only that one closest to such a position is actually displaced: (209) albeit clumsy, is grammatical, (210) is not.
Locality (204), when applying to the particular case of displacement, as in (209, 210), is formulated as follows (Chomsky, 1995, p. 311):

Minimalist definition 4.2 (Minimal Link Condition)**.**

K attracts α only if there is no β, β closer to K than α, such that K attracts β.

The picture that emerges from the discussion so far is that two key-properties of Syntax discussed in Sections 4.3, 4.4 are motivated in biological terms: Syntax is endowed with recursion *qua* MSNF and with discreteness *qua* module fed by discrete inputs of a lexicon.
Each input of the lexicon is twofold, being actually made up of a module-specific input, the syntactic category, and of an external input, ie the item of the item-and-arrangement model. While the module Syntax performing recursion is also found in other living species (MSNF), the lexicon thus characterized is uniquely human. The external input of the module Syntax, the item, is a pair morph - feature.
Two additional properties of the module Syntax *qua* MSNF has showed up: no backtracking and locality, which are imposed by organic limitations *internal* to this module (e.g. energy and / or memory costs). More generally, this kind of biological constraint is referred to as Condition of Optimal Design.
On the other hand, two remaining properties of the module Syntax, namely hierarchy and displacement (see Section 4.4) do not seem to belong to MSNF (Chomsky, 2008; Moro, 2000) and therefore their existence is to be motivated independently on Syntax (and its Lexicon). This issue is dealt with in the next Section.

4.5.2 The Performance systems and displacement

MSNF is not the only module primates, birds and insects are equipped with: it has been long been observed in them, in fact, a module responsible for articulating and recognizing sounds and one involved in interpreting the world and in interacting with it (Chomsky et al., 2002, and references therein).
They roughly correspond to the traditional components of phonology and semantics and are designated in MP as A(RTICULATORY)-P(ERCEPTUAL SYSTEM) and as C(CONCEPTUAL)-I(NTENTIONAL SYSTEM) respectively, but they can also be subsumed under the generic label of PERFORMANCE SYSTEMS.
Once again, human beings plausibly differ from other primates, birds and insects not in the module per se but in the lexicon feeding it: in the former ones, but not in the latter, the modules A-P and C-I can be fed by the output of Syntax (exemplified in (208)). Or, to put it differently, MSNF and the performance systems are interlocked in the human species, while they are incommunicado in the other primates, birds and insects (Chomsky et al., 2002, p. 1578).
Chomsky et al. (2002, 2005) contend that the change from separate to connected modules is an evolutionary shift driven by communicative needs. Such a shift exapted the modules away from their special function (e.g. recursion, in case of Syntax), creating an interacting system - the faculty of language in the broad sense (FLB): 'communication must be one of the primary selective forces that influenced the evolution of FLB' (Chomsky et al., 2005, p. 186).
In greater detail, as discussed in Section 4.5 and illustrated in Fig. 4.1, the output of Syntax is able to work as input of A-I or C-I only insofar as it has the ability to be *associated* with a module-specific input of their proprietary lexicon (call it A-P-specific input and C-I-specific input respectively).

Given that the A-P / C-I lexica make the output of Syntax to associate with an A-P / C-I-specific input, these lexica are also said INTERFACE LEVELS.

As for the exact nature of the A-P / C-I-specific input, it seems to be an additional piece of information for the output of Syntax.

Considering C-I first (and heavily simplifying a matter which is much more complex), the meaning of an English idiom like *to take pictures* clearly is not the sum of its parts *to take* and *pictures* as taken from the Lexicon and combined in Syntax, since it stands for *to photograph*: 'a strong argument' according to Chomsky (1995, p. 208), to posit idiomatic meaning (ie NON-COMPOSITIONALITY) at C-I.

Under the Modularity Hypothesis, the non-compositional meaning of the idiom *to take pictures* is an additional piece of information associated with its syntactic / compositional meaning and therefore the non-compositional meaning is the C-I-specific input, the compositional one the external input (the output of Syntax).

Marantz (1997) generalizes Chomsky's claim to the effect that C-I assigns non-compositional meaning to *all* output of Syntax: e.g. not only idioms, but also abstract nouns such as *transmission*. Thus *transmission* standing for *gearbox* is a non-compositional nuance of this word that is stored, according to Marantz, in the lexicon of C-I. Marantz describes the non-compositional meaning as part and parcel of the speaker's knowledge of the world and then calls the C-I-specific input ENCYCLOPEDIA[6].

Turning to A-P, the A-P-specific input according to Scheer (2007, p. 1) can be identified with Chomsky (1957)'s / Chomsky & Halle (1968)'s phonetic representation (e.g. *took*: see Section 4.3). On this scenario, the module A-P has the ability to replace the phonological representation *tak + ed* with the phonetic representation *took* only subsequent to the association of inputs occurring in its dedicated lexicon. The lexicon of A-P receives the phonological representation *tak + ed* from Syntax (external input) and makes it associate with the A-P-specific input *took*.

To summarize the discussion so far, the output of Syntax can work as the (external) input of A-P / C-I provided it is associated with the A-P / C-I-specific input: non-compositional meaning or a phonetic representation, respectively.

Matters however are more complex than they at first appear: the careful reader will have in fact noticed that what is actually associated with the A-P-specific input is only the sound side of the output of Syntax - say, (208): its phonological representation (208a). By the same token, the C-I-specific input is associated only with the meaning side of the output of Syntax: the compositional meaning and the syntactic category (208b).

Simply put, the output of Syntax when having to do with the lexica of A-P / C-I (interface levels) is split in two. MP takes the correlation between the splitting of the output of Syntax and the involvement of the interface levels to be not accidental.

The former is interpreted in fact as an *external* constraint imposed by organic limitations of the latter: this is the inability for a lexicon to preserve all the information conveyed by the external input, a limitation referred to as Information Bottleneck / Partiality in Section 4.5.

[6]Notice that Marantz (1997, p. 203) points out that this generalization holds irrespective of whether the approach he proposes, the so-called DISTRIBUTED MORPHOLOGY, is followed or not. This happens because his Distributed Morphology basically adopts MP's architectural settings, only departing from it in that morphs are not found in the Lexicon of Syntax, but in the lexicon of A-P (LATE INSERTION). This issue does not affect the nature of C-I.

Under this perspective, association taking place in the proprietary lexicon of A-P loses part of the information encoded in the external input - the output of Syntax. The lost information is that related with the meaning. Vice versa, association taking place in the proprietary lexicon of C-I loses part of the information encoded in the external input (the output of Syntax), but in this case the lost information is that related with the sound.
It could be easiliy noted that association appears to privilege that part of information of the external input which is more relevant for the performance system: the sound-side of the output of Syntax in case of A-P, its meaning-side in case of C-I. The rationale for this behavior is functional: although the meaning-side of the output of Syntax can in principle associate with the A-P-specific input and the other way around (see Section 4.5), there is no doubt that in this case the result of the computation of the performance systems does not fit the purpose for which the former independent recursive MSNF, A-P and C-I evolved into an interconnected system: communication (see the beginning of this Section).
If, in fact, the compositional meaning of the output of Syntax, say (208b), were used as the external input of A-P, it could be no longer available as the external input of C-I and hence the syntactic structure could not be organized into a meaningful expression. By the same token, if the phonological representation, say (208a) of the output of Syntax were used as the external input of C-I, it could be no longer available as the external input of A-P and hence the syntactic structure could not be assigned a pronunciation.
But the interaction between Syntax and the performance systems (FLB) would thus yield a syntactic structure that would not serve the purpose of communication their interaction has been devised for (technically, 'the derivation CRASHES'). It follows that splitting of the output of Syntax as described above (the phonological representation of the output of Syntax 'sent' to A-P, its compositional meaning to C-I) is the only way for FLB to yield a syntactic structure tailored to the communicative needs (technically, to allow 'the derivation CONVERGE')[7]. For this reason, the *external* constraint of Information Bottleneck thus characterized is designated in the minimalist literature as CONDITION OF FULL INTERPRETATION.
Alternatively, since this constraint forces the lexicon of A-P / C-I to 'strip away' from the output of Syntax the information useless for the concerned performance system, the Condition of Full interpretation is also called BARE OUTPUT CONDITION.
Satisfaction of this condition results in a kind of complementary distribution for the split output of Syntax. A simplified representation of this state of affairs, which abstracts away from the syntactic categories is the following:

(211) Lexicon of A-P: $[\text{MORPH} + \text{MORPH}]_{SOUND}$: $\text{PHONETIC REPRESENTATION}_{SOUND}$

(212) Lexicon of C-I: $[\text{FEATURE} + \text{FEATURE}]_{MEANING}$: $\text{ENCYCLOPEDIA}_{MEANING}$

[7] Chomsky, Hauser and Fitch in fact state that:

> "It is possible, as we discuss below, that key computational capacities evolved for reasons other than communication but, after they proved to have utility in communication, were altered because of constraints imposed at both the periphery (e.g., what we can hear and say or see and sign, the rapidity with which the auditory cortex can process rapid temporal and spectral changes) and more central levels (e.g., conceptual and cognitive structures, pragmatics, memory limitations)." (Chomsky et al., 2002, pp. 1569-1570)

The sound-side of the output of Syntax is thus associated with the A-P-specific input, an association dubbed in the minimalist literature P(HONOLOGICAL) F(ORM), while its meaning-side with the C-I-specific input, an association designated as L(OGICAL) F(ORM).
For the sake of completeness, it is worth noting in this connection that, as Chomsky (2008, p. 158, fn. 11) remarks, 'some confusion has been caused by recent use of the term "LF" to refer to the C-I interface itself, thus departing from its definition... as the output of (overt and covert) syntactic operations and the input to rules of semantic interpretation that map LF to C-I'.
To summarize:

(213) LEXICON OF A-P = EXTERNAL INPUT $\Leftrightarrow$ [MODULE-SPEC. INPUT$_i$] > [MODULE$_i$]
Sound = morph + morph $\Leftrightarrow$ [A-P-spec. input$_i$] > [A-P$_i$]

(214) LEXICON OF C-I = EXTERNAL INPUT $\Leftrightarrow$ [MODULE-SPEC. INPUT$_i$] > [MODULE$_i$]
Meaning = feature + feature $\Leftrightarrow$ [C-I-spec. input$_i$] > [C-I$_i$]

(215) LEXICON OF A-I / C-I = OUTPUT OF SYNTAX $\Leftrightarrow$ [MODULE-SPEC. INPUT$_i$]
interface level = phonological / logical form

It has been noticed in the minimalist literature, however, that the Principle of Full Interpretation is not totally satisfied by putting morphs and features in complementary distribution, for at least two reasons.
First, it seems that there are some functional features (e.g. Case) that because of their close-endedness can turn out to have no or very poor meaning: in this case, very much as the meaningless morphs, they are useless for C-I (see immediately above). Then, if C-I processed these features, it would yield a syntactic structure not suitable for communication, and the derivation would crash. Features of this kind are generally referred to as UNINTERPRETABLE.
The same difficulty arises for morphs: they originate in Syntax, which put them together *prior* to Spell-Out. It ensues that they are not ordered one after the other, because such an ordering is a consequence of the time axis on which they are located *subsequent* to Spell-Out - this is the Saussurean property of linearity examined in Section 4.3. But *unordered* morphs are useless for A-P, which is the module responsible for their pronunciation and hence, as Scheer (2004) points out, carries out its operations only over *ordered* morphs, so that these latter can be properly uttered on the time axis: this is why Scheer (2004, pp. 235ff.) describes such operations as LATERAL[8].
Nevertheless, if no morph is suitable for A-P, no Spell-Out of the syntactic structure takes place and the derivation crashes.
The just outlined scenario can be illustrated as follows:

(216) [MORPH + MORPH$_;$] // PHONETIC REPRESENTATION$_>$
; = non-linear
> = linear
// = no Full Interpretation (no association and crash)

[8]For example, the standard account of the English PL morpheme *-s* (Chomsky & Halle, 1968) has it that it is a voiced consonant (/-z/) undergoing devoicing (/-s/) when preceded by an unvoiced consonant: *t* of *bat-* forces the phonological representation *bat + z* to be phonetically realized as *bat + s*. Plainly, if the module Phonology carried out the operation of devoicing over an unordered input (either *bat + z* or *z + bat*), devoicing could not successfully apply because the doamin of application of this operation is not clear.

(217) [FEATURE$_{UN}$ + FEATURE$_{MEANING}$] // ENCYCLOPEDIA$_{MEANING}$
UN = Uninterpretable feature
// = no Full Interpretation (no association and crash)

The dominant research trend in the MP concentrates on the problem observed at the lexicon of C-I (217) and maintains that it is an IMPERFECTION OF LANGUAGE. It is also assumed that Syntax capitalizes on the property of *locality* to get around this problem. Remarkably, a non-trivial by-product of this problem-solving strategy is displacement. In other words, the rationale for displacement is Full Interpretation *at C-I*.
Chomsky has given this idea three different implementations, whose details are not relevant for the present work (see Moro 2000, Ch. 1 and Baker 2002 for critical review and references). What matters here is, as Moro remarks, that the basic line of reasoning underlying Chomsky's idea has remained constant throughout its different formulations:

> "This hypothesis is crucially based on two assumptions: first, obviously, there are uninterpretable features; second, a feature can be deleted only in a proper local configuration with another feature of the same type. [...] The mechanism is essentially this: any (item containing an) uninterpretable feature that is introduced into derivation must be "paired" before reaching LF with (another item containing) the same type of feature in a local relation. This is what deletes the uninterpretable features".
> (Moro, 2000, pp. 11-12)

This account of the imperfection of language and of displacement is usually referred to as FEATURE CHECKING and can be schematized as follows:

(218) [$\underline{\text{feature}_{UN}}$ [feature$_{IN}$ ... [t$_{UN}$ etc...
$\underline{\text{X}}$ = deleted feature
IN = interpretable feature
... = distance

An alternative proposal, the so-called DYNAMIC ANTISYMMETRY put forward by Moro (2000), concentrates on the problem observed at the lexicon of A-P (216) and calls into the picture the property of *hierarchy* to get around it: in this proposal too, the rationale for displacement is Full Interpretation (and locality is still involved in it: see Moro 2000, Appendix and Moro 2004). Nevertheless the performance system is no longer C-I, but *A-P*.
Moro's approach is a radical departure from the mainstream trend, given that displacement is divorced by C-I: the uninterpretable features, in fact, are no longer regarded as its trigger (although their presence can be easily integrated within the Dynamic Antisymmetry framework: see Moro 2004).
On theoretical grounds, Moro (2000)'s Dynamic Antisymmetry seems to be superior to Feature Checking because it is able to derive two other properties of Syntax and of A-P which in the dominant research trend are unrelated:

- endocentricity ('hierarchy, phrase structure')
- invisibility of the trace at A-P

Recall from Section 4.4 that endocentricity is the native speaker's intuition that an item, the Head, is not on the same footing of another, the Dependent. Recall also from fn. 57 that the trace represents the position of a constituent before displacement, notationed t but not actually heard).
Thus, from the vantage point of Dynamic Antisymmetry, *phrase structure and displacement are not independent* (Moro, 2000) and the zero-realization is not an *ad hoc* postulate.
Two are the guiding ideas of this approach. The former is due to Kayne (1994)'s Antisymmetry and has it that the Bloomfieldian endocentricity / hierarchy observable in Syntax is *not* a primitive property, but something modelled after Saussurean linearity at A-P. Under this perspective, endocentricity / hierarchy in Syntax obtains iff and only if two items are put together in a special way, one that 'mimics' A-P.
The key-notion to bridge the gap between Syntax and A-P is LINEAR ORDERING, to be understood *in the set-theory sense, not in the Saussurean one.* In particular, what makes it possible to bridge the gap between Syntax and A-P is a property of the set-theoretical notion of linear ordering: ANTISYMMETRY (hence the name of the Kaynean approach).
As a first approximation, Kayne's approach can be described as follows (from Kayne 1994, p. 5)

Minimalist definition 4.3 (Antisymmetry)**.**

There should be a very close match between the linear ordering relation on the set of terminals and some comparable relation on nonterminals

where 'TERMINALS' are the syntactic categories and 'NONTERMINALS' are, as a first approximation, the morphs.
The latter idea, proposed by Moro (2000) is that endocentricity / hierarchy obtains not only by (a special) combination of two items ('PHRASE STRUCTURE BUILDING') but also through their displacement - hence the designation 'dynamic ':

Minimalist definition 4.4 (Dynamic Antisymmetry)**.**

Movement is driven by the search for antisymmetry.

(Dynamic) Antisymmetry is discussed in greater detail in the next Sections.

4.5.3 Phrase structure: Antisymmetry and Cartography

Under the (Dynamic) Antisymmetry approach, endocentricity / hierarchy is the syntactic counterpart of phonological linearity. The following question therefore arises: how can a syntactic phenomenon be likened to a phonological one? Prima facie, in fact, they look very different.
On one hand, it is an empirical hypothesis in MP that Syntax has a *vertical* dimension and that syntactic relations are built *bottom-up* (Radford 2004, Ch. 3, Scheer (2004, Ch. 7)).
'Relation' is a notion to be understood along the lines of a (simplified) set-theory: intuitively, a relation holds between two objects belonging to the same set. As discussed at length in the previous Sections, the simplest relation possible is the combination of two items, an operation called in MP MERGE (Chomsky 1995, Ch. 4, Moro 2000, Ch. 31-34): e.g. *John's* and *proof* can be put together via Merge to yield a NP (*proof of the theorem*).
It has also been pointed out that an item is in reality a pair morph(s) - feature(s) associated with a syntactic category and that this latter is the module-specific input over which Syntax

carries out the operation of recursive Merge. In what follows, accordingly, only the syntactic category is taken into account when Syntax is dealt with. In this light, the N *proof* and the PP *John's* can be regarded as two syntactic categories X, Y, irrespective of their meaning and it is X and Y that enter into relation as a consequence of Merge.
This *syntactic relation* can be geometrically represented by having X and Y share an intersection point (the NODE Z), as depicted in Fig. 4.3.a, and moreover it can be taken to be:

- *un*ordered
- *not* hierarchical

The former property has been already discussed at the end of the previous Section. As for the latter property, without entering into technicalities (see Moro 2000, p. 22, 31-40 and Chomsky 2008, p. 139 for details), it straightforwardly follows from characterizing Syntax as a module, as done in Section 4.5.1: being either an Head or a Dependent is not a piece of information encoded in the external input of Lexicon (= lexicon of Syntax), nor is it encoded in its internal input. In addition, assigning the status of Head / Dependent is not the operation carried out by narrow Syntax, ie MSNF. The syntactic relation thus characterized is depicted in Fig. 4.3.a,b.
On the other hand, A-P is sharply different from Syntax, since it has an *horizontal* dimension, being put on the time axis (cp. the Saussurean linearity and Scheer's laterality discussed in the previous Section). A case in point is the sound-side of the two items *John's, proof*. Setting aside the number of morphs involved in these two items, not relevant here, it is important to note that such morphs are organized into two major complexes (the word *proof* and the phrase *John's*) corresponding to the syntactic categories N, PP or, more generally, X, Y. Accordingly, these two complexes can be referred to as x, y.
But x, y by virtue of their belonging to the same sound string, enter into a *phonological relation* which is:

- ordered
- *not* hierarchical

The former property has been already dealt with at the end of the previous Section, while the latter can be easily detected by submitting to the substitution test illustrated in Section 4.3 two expressions void of *linguistic* meaning (e.g. an interjection or a cry) and hence likenable to two morphs: neither an Head nor a Dependent can be singled out from them (see Scheer 2004 for more technical discussion).
The phonological relation between x and y can be geometrically represented by having them share an intersection point, that is juxtaposing them: *Johnsproof* (Moro, 2008, p. 58). Abstracting away from the intersection point, the same relation can be represented along the lines of Fig. 4.3.c.
Fig. 4.3 summarizes the discussion so far and makes adamantly clear that X and Y are 'on the same footing' (not hierarchical) and so do x and y.
The careful reader will have noticed that casting the syntactic categories X, Y (Fig. 4.3.b) and the (complexes of) morphs x, y (Fig. 4.3.c) in set-theoretic terms answers the question raised above: set-theory, in fact, reveals that both of syntactic categories and morphs, in spite of their different nature, can be compared relative to the *property of being objects belonging to a set*. In

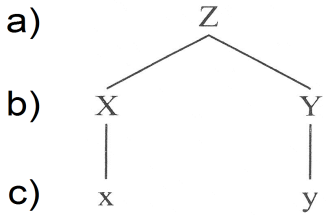


Figure 4.3: The most basic instance of Merge (Moro, 2000, p. 22)

the wake of Kayne (1994), syntactic categories and (complexes of) morphs will be henceforward designated as TERMINALS and NONTERMINALS respectively.
Sticking to the set-theory, it is convenient to add to the picture the notion of LINEAR ORDERING, a less generic kind of relation that has three defining properties:

- it is transitive
- it is total
- it is antisymmetric

Before exemplifying this notion, it should be noted with Moro (2000, p. 116, fn. 5) that the expression *linear ordering* could lead scholars accustomed to the Saussurean notion of *linear(ity)* into error, because in the set-theory 'the term *linear* is not inherently related to the fact that words are physically put in line (in sequence) with one another' on the time axis.
For example, a linear ordering could hold between two objects B, C possessed of size, which are visually *simultaneous*: Fig. 4.4 is a handy illustration of the properties of linear ordering.
The relation between B and C is transitive because C is bigger than B, B is bigger than A,

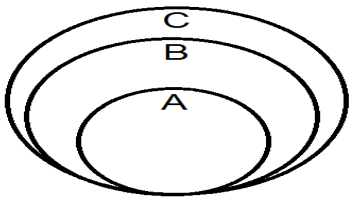


Figure 4.4: Linear ordering (set-theory)

and hence not only B is bigger than A, but also C is bigger than A.
It is total iff and only if it covers *all* the objects of the set, which is actually the case: 'being bigger than' is a relation encompassing both the objects B, C of the relevant set.
It is antisymmetric in the sense that C cannot be bigger than B and B bigger than C in the same set.
Crucial to Kayne (1994)'s proposal is the claim that the difference between A-P and Syntax *boils down to the presence vs. absence of the (set-theoretical) linear ordering*: the relation holding between x, y on (the time axis of) A-P is a linear ordering, that between X, Y is *not*.
As a non-trivial corollary, Syntax is realigned with A-P *if linear ordering is created*.
Considering A-P first, the relevant relation of linear ordering is what is traditionally called

precedence, which has clearly to do with the horizontal dimension of this module.
In the string *(John's) proof of the theorem*, *of the theorem* precedes *proof*, and *proof* precedes *John's*, but *of the theorem* too precedes *John's* (transitivity).
The very notion of precedence implies both *proof, of the theorem* (totality).
But, if *proof* precedes *of the theorem*, it is obvious that *of the theorem* cannot precede *proof* (antisymmetry).
The same holds if the traditional notion of *subsequence* is considered: in both the cases, A-P therefore manifests (set-theoretical) linear ordering (as well as Saussurean linearity).
Turning to Syntax, the relevant relation to look at is *height*, which has clearly to do with the vertical dimension of this module. Technically speaking, it is referred to as DOMINANCE and defined as follows (Haegeman, 1991, p. 85):

Minimalist definition 4.5 (Dominance)**.**

Node A dominates node B if and only if A is higher up in the tree than B and if you can trace a line from A to B going only downwards

The relation of height thus formulated, however, is too weak, because it covers only part of the 'syntactic' set depicted in Fig. 4.3.a,b, namely one out of two objects, and their intersection point Z: a line can be traced downwards *either* from Z to X *or* from Z to Y.
In other words, DOMINANCE is not transitive, total, antisymmetric, properties that by definition imply at least *two* objects in the relevant set.
Dominance is then recast to include a second object in the set, and then the relation known in the (pre-)minimalist literature as C-COMMAND obtains - this is described *under the points (a), (b)* in the following (Kayne 1994, p. 4, Moro 2000, p. 21) - for the time being skip the adverb *asymmetrically* and the point (c,d):

Minimalist definition 4.6 (Asymmetric c-command)**.**

A asymmetrically c-commands B if and only if
a. A does not dominate B
b. the first node dominating A also dominates B
c. A c-commands B and B does not c-command A
d. A and B are categories

Now, the requirement of totality is met, since the relation of double dominance (b) 'encoded' in the node Z of Fig. 4.3.a,b involves both X and Y.
However, this is not still enough for satisfying transitivity, since a third object is required.
To get things even worse, antisymmetry is violated because X c-commands Y *and*, symmetrically, Y c-commands X: adopting a more familiar visual representation, it is as if X were bigger than Y and Y bigger than X. The scenario observable in the 'syntactic' set depicted in Fig. 4.3.a,b is therefore at odds with the property of antisymmetry of linear ordering and rather is defined by the property of SYMMETRY. This is why Moro (2000, p. 22) calls this configuration POINT OF SYMMETRY (POS).
Consequently, the 'syntactic' set depicted in Fig. 4.3.a,b is rebuilt, through simple insertion (technically PHRASE STRUCTURE BUILDING or BASE GENERATION) of a nonterminal, as illustrated in Fig. 4.5.
In this case, totality still occurs, since the relation of double dominance (b) 'encoded' in the

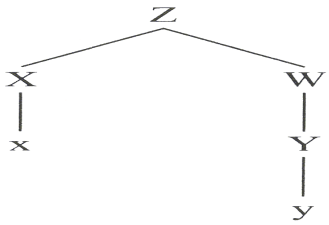
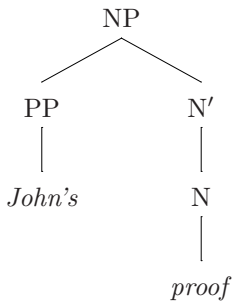


Figure 4.5: Addition of nonterminals to the POS depicted in Fig. 4.3.a,b (Moro, 2000, p. 21)

node Z involves both X and W.
It is also transitive, because not only Z dominates X and W, but also W dominates Y, so that Z dominates Y.
Yet, c-command is not antisymmetric, since X c-commands W *and* W c-commands X: to ensure antisymmetry, a sort of change of perspective is needed, to the effect that the relation of c-command to consider in Fig. 4.5 holds between X and *Y*, *not* X and *W*. This has as outcome, in fact, that X c-commands Y, but Y does *not* c-command X because the first node dominating Y, ie W, does *not* dominate X.
In other words, antisymmetry requires as a postulate negligibility of the (symmetric) relation of c-command between X and W, which leads to rephrase the definition of c-command in order to include such a condition: accordingly, the point (c) is added to and c-command is rewritten as Asymmetric c-command.
The picture that emerges from the discussion so far is that both the 'syntactic set' of nonterminals X, Y and the corresponding 'phonological set' of terminals x, y possess linear ordering. Such a property is intrinsic to the terminals while it is ensured in the nonterminals iff and only if the syntactic structure depicted in Fig. 4.5 occurs.
To make the discussion more concrete, X, Y can be exemplified by the nonterminals PP, N, and x,y by the terminals *John's, proof*, as depicted in (219):

(219)



The notation N', P' employed for (219) will be clarified shortly. Notice also that the mirror image of Fig. 4.5, exemplified in Fig. 4.6 is barred, because it implies a directionality right to left in the spell-out of the terminals (**proof John's*), which actually is not the case (Moro, 2000, pp. 21-22).
However, this is not the whole story: given that Syntax is by its own nature recursive, the operation of phrase structure building does not give up and further piece(s) of syntactic structure can be merged (= added) to the phrase structure in Fig. 4.5.
In principle, recursion of MSNF can derive several syntactic structures from that depicted in Fig. 4.5 but Bare Output Conditions allow just one of them, illustrated in Fig. 4.7. Among the

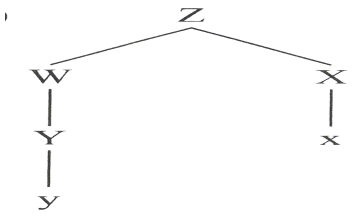


Figure 4.6: Mirror image of Fig. 4.5 (Moro, 2000, p. 21)

possible configurations at issue, in fact, this is the only one satisfying all the requirements of linear ordering: transitivity, totality and antisymmetry (see Kayne 1994, Ch. 2 and Moro 2000, Ch. 2 for an overview of all the illegitimate syntactic structures).

As shown in Fig. 4.7, the legitimate syntactic structure displays a peculiarity: its topmost

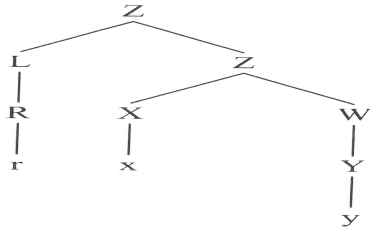


Figure 4.7: Addition of nonterminals to the syntactic structure depicted in Fig. 4.5 (Moro, 2000, p. 24)

nonterminal is a syntactic category *identical with* and *taking the place of* the topmost syntactic category of the structure previous to its insertion, ie Z in Fig. 4.5 (Kayne 1994, p. 16, Moro 2000, p. 24).

Two aspects of this syntactic structure then need clarification: why it possesses linear ordering and why it 'clones' the topmost syntactic category of its previous derivational step.

As for the linear ordering of the syntactic structure in Fig. 4.7, Kayne (1994, p. 16-18) builds his argument as follows.

The nonterminal L enters into many syntactic relations, namely with X, W, Y. Actually, it asymmetrically c-commands X, W, Y (totality, antisymmetry).

Transitivity too is found: the higher Z dominates L and the lower Z, the lower Z dominates X, W, Y and hence the higher Z 'transitively' dominates X, W, Y.

Moreover, X enters into a syntactic relation with Y, asymmetrically c-commanding Y (totality, antisymmetry).

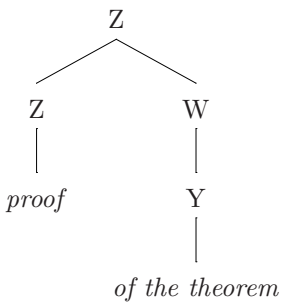
As for transitivity, it holds as well: the lower Z dominates X and W, W dominates Y and hence the lower Z 'transitively' dominates Y.

A closer look at Fig. 4.7, however, raises a non-trivial problem: the postulate (c) of the definition of Asymmetric c-command prevents the lower Z from *symmetrically* c-commanding L, but not from entering into a syntactic relation with the only nonterminal left so far out of the picture, namely R. This relation holds because it possesses all the properties of linear ordering: the lower Z *asymmetrically* c-commands R (totality, antisymmetry) and the former dominates the latter by virtue of transitivity (the higher Z dominates the lower Z and L, L dominates R

and hence the higher Z dominates R).
But this is not desirable, given that this is precisely the syntactic scenario depicted in Fig. 4.6 (Kayne 1994, pp. 10-11, 16, Moro 2000, pp. 23-24), which implies a directionality right to left in the Spell-out of terminals (see immediately above).
The problem is circumvented if the nonterminal higher than the (lower) Z, if any, is merged into the syntactic structure depicted in Fig. 4.5 as a 'clone' of this nonterminal. Simplifying, this has the effect of eliminating the undesired Asymmetric c-command from (the lower) Z to R since (the new and higher) Z enters into a new relation with R, namely dominance, which is in compliance with the directionality of the Spell-out of terminals[9].
Such a state of affairs demonstrates that the syntactic structure in Fig. 4.7 possesses linear ordering. It also explains why the syntactic structure depicted in Fig. 4.7 'clones' the topmost syntactic category found in its previous derivational step (Fig. 4.5): this is the only way to render the whole syntactic structure depicted in Fig. 4.7 endowed with linear ordering.
To put it differently, the definition of Asymmetric c-command requires as a postulate negligibility of the syntactic category replaced by its higher counterpart, which leads to include into this definition the condition (d).
Far from being an ad hoc solution, 'cloning' of the topmost syntactic category observed in the syntactic strcuture of Fig. 4.7 is the syntactic representation of the phenomenon of SUBSTITUTION first observed by Bloomfield (see Section 4.4) - actually, 'cloning' of the topmost syntactic category derives endocentricity / hierarchy.
Recall in fact from Section 4.4 that the Bloomfieldian endocentric analysis of pairs like *proof* and *John's proof* is precisely interpreting the complex *John's proof* as *the same category* N characterizing *proof* because the complex *John's proof* has the ability to replace *proof* which is no doubt a N.
In this light, in the syntactic structure of Fig. 4.7 the lower Z 'cloned' into the higher Z is a Head (e.g. *proof*) and the piece of structure that precedes the lower and the higher Z is the Dependent (*John's*).
Remarkably, the parallelism between the Kaynean Antisymmetry and the Bloomfieldian endocentricity / hierarchy forces to interpret the expression *proof of the theorem* in the same way: after all, the substitution test reveals that *of the theorem* is a Dependent as much as *John's* is. The only difference lies in that *of the theorem* follows its Head.
But this is tantamount to saying that in the syntactic structure illustrated in Fig. 4.5 the piece of structure following X is the Dependent and that X is the Head. As a corollary, if X is an Head, it undergoes 'cloning' into the topmost syntactic category of Fig. 4.5, ie Z: X is then better seen as a lower nonterminal Z, cloned into a higher nonterminal Z. This is illustrated in (220), which rewrites the relevant portion of Fig. 4.7:

[9]Notice in passing that insertion of the new and higher Z into the syntactic structure depicted in Fig. 4.5 gives rise to a syntactic structure that is legitimate in all the other respects, as discussed immediately above.

(220)



At this point, one may object that this theoretical move is justified within the Bloomfieldian framework, but not in Kayne's, because there is no compelling theory-internal (say, 'geometric') reason to reinterpret X as Z.

This objection loses its bite when one looks at Fig. 4.7 more carefully.

Suppose for a while that the lower Z in it is not the 'clone' of X and that the two nonterminals are independent each other. Recall also that the lower Z in Fig. 4.7 is no longer a syntactic category, because it has been replaced in this function by the higher Z. Accordingly, X symmetrically c-commands L and asymmetrically c-commands R: adopting a more familiar visual representation, it is as if the lower Z would no longer exist and therefore X were immediately dominated by the higher Z.

As already shown above, both these relations are not legitimate and therefore the condition (c) and (d) of Asymmetric c-command are introduced in order to rule them out. Nevertheless, in Fig. 4.7 the condition (c) has the ability of inhibiting X from symmetrically c-commanding L but the condition (d) has no bearing on X.

This happens because the condition (d) states that the Asymmetric c-command from a given nonterminal to another nonterminal is possible iff and only if both of them are syntactic categories. Now, this is actually the case, X and R being syntactic categories, so that this kind of Asymmetric c-command cannot be ruled out and makes the whole syntactic structure illegitimate.

The problem evaporates as soon as one no longer thinks of X as a syntactic category: but the only way to implement this idea in the Kaynean theory is merging into the syntactic structure an higher 'clone' of a given syntactic category, which replaces this latter in its function. Accordingly, if X is regarded as a nonterminal Z 'cloned' into the node Z dominating it and W, Y, the condition (d) has the ability of ruling out the Asymmetric c-command from X to R because it states that X and R be syntactic categories.

This condition, in fact, is no longer met, since X is no longer a syntactic category, having 'cloned' into Z.

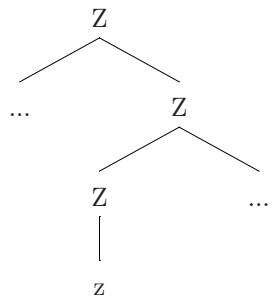
It follows that the whole syntactic structure depicted in Fig. 4.7 is legitimate iff and only if X is reinterpreted as Z. The same holds for the syntactic structure depicted in Fig. 4.5, which is equivalent to that in Fig. 4.7 insofar as the Head is concerned.

As a corollary, Kayne's theory is realigned with the Bloomfieldian endocentric analysis. To put it differently:

(221) endocentricity / locality is not an intrinsic property of Syntax, but a property imposed on Syntax by A-P

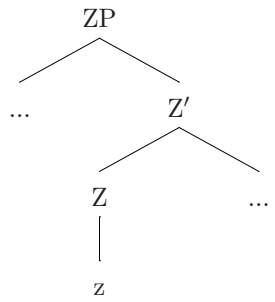
Staying with endocentricity / locality a little longer, it is worth noting that in Fig. 4.7 the syntactic category Z has been 'cloned' twice, to the effect that three instances of Z are found, as illustrated in (222):

(222)



In order to distinguish among them, the lowest Z is generally notationed as Z, the intermediate Z as Z' (pronounced 'Z bar') and the highest Z as ZP, which stands for 'phrase of Z', to the effect that the standard representation of (222) is (223)[10]:

(223)



In particular, the designation 'phrase of Z' for the highest manifestation of Z in the syntactic structure in Fig. 4.7 and (223) is based on the theory-internal claim that what is traditionally referred to as 'phrase' (ie a combination of items larger than a word: see Section 4.2) cannot be assigned a syntactic structure bigger than that depicted in Fig. 4.7 and (223).
Theoretical evidence for this hypothesis is that recursive addition of nonterminals to the syntactic structure in Fig. 4.7 yields a syntactic structure not possessed of linear ordering.
It is worth noting in this connection that even resorting to the strategy of 'cloning' the topmost syntactic category of the previous derivational step is not able to endow the new syntactic structure with the required linear ordering.
In fact, if this strategy is resorted to, there obtains the syntactic structure illustrated in Fig. 4.8, which is flawed in one crucial respect: it is a POS (Moro 2000, pp. 22, 25 and see also Kayne 1994, pp. 22-27 for more arguments on the illegitimacy of the syntactic structure in question).

Actually, contrasting Fig. 4.7 and Fig. 4.8 shows that the condition (c) of Asymmetric c-command ('A c-commands B and B does not c-command A') is met in both of the syntactic structures under discussion.
In Fig. 4.8 the intermediate Z cannot symmetrically c-command F very much as in Fig. 4.7 the lower Z cannot symmetrically c-command L.
In a likewise fashion, in Fig. 4.8 L cannot symmetrically c-command F very much as in Fig. 4.7 X cannot symmetrically c-command L.

[10]Other notations are also possible. They include Z^0 instead of Z and $\bar{Z}$ instead of Z'.

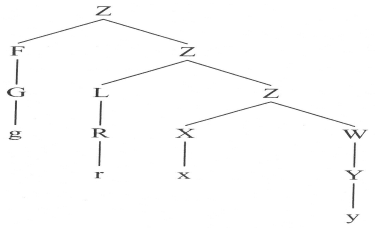
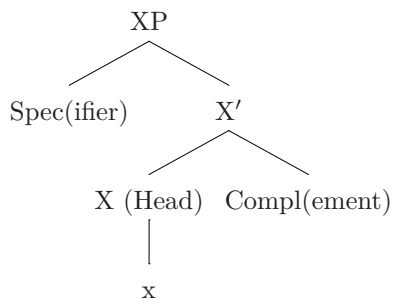


Figure 4.8: Addition of nonterminals to the syntactic structure depicted in Fig. 4.7 (Moro, 2000, p. 25)

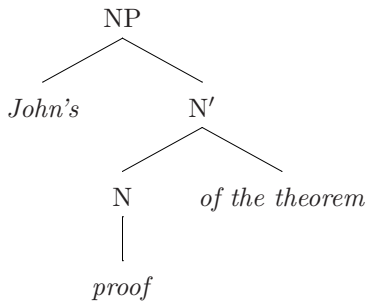
Matters, however, get more complicated when the condition (d) of Asymmetric c-command is taken into account ('A and B are categories').
True, in Fig. 4.8 the intermediate Z cannot asymmetrically c-command G very much as in Fig. 4.7 the lower Z cannot asymmetrically c-command R, because neither of the intermediate Z or the lower Z is a category.
Nonetheless, in precisely the same way as in Fig. 4.7 X is potentially able to asymmetrically c-command R, thus in Fig. 4.8 L is potentially able to asymmetrically c-command G.
At this point, one may expect that the condition (d) of Asymmetric c-command applies, in order to inhibit X and L from asymmetrically c-commanding R and G respectively, on the ground that X and L are not categories.
But this holds valid for X only: as seen above X is actually an Head Z that, having been replaced in its function by an higher 'clone' Z, is no longer able to work as syntactic category. On the other hand, L *is* a syntactic category, actually the syntactic category associated with the Dependent (e.g. the PP *John's*).
Thus the nonterminal L, part and parcel of the nonterminals added to the syntactic structure in Fig. 4.7 to yield its counterpart in Fig. 4.8, asymmetrically c-commands G and the condition (d) has not the ability of preventing L from doing so.
This results in the same scenario depicted in Fig. 4.6, which is not compatible with Spell-out of terminals at A-P and is therefore illegitimate. To get things even worse, in Fig. 4.8 not only L asymmetrically c-commands G, but also F asymmetrically c-commands R, which creates a POS.
For these reasons, the syntactic structure depicted in Fig. 4.7 and in (223) is the maximum syntactic structure that can be assigned to a phrase. In fact, Fig. 4.7 is the only configuration allowed by Asymmetric c-command as defined in this Section, which in turn is the only way to force the nonterminals (Syntax) to realign with the terminals (A-P) in terms of the set-theoretic notion of linear ordering.
Such a syntactic structure in the (pre-)minimalist literature is generally referred to as X-BAR FORMAT: the symbol X is preferred over Z to refer to any generic Head, while the term *bar* highlights that the syntactic constituents involved in endocentrictiy / hierarchy are not only the Head (X) and the phrase (XP), but also the intermediate X' (actually pronounced 'X bar').
Notice also that after Chomsky (1970), the Dependent preceding the Head is called SPEC(IFIER) and that following the Head COMPL(EMENT) and then these terms are to be understood in purely syntactic and geometric terms, along the lines discussed so far, not in the familiar (semantic) sense:

(224)



Or, more concretely:

(225)



To sum up, looking at both the nonterminals entering into the X-bar format (224) and the terminals through the lens of the set theory has shown that both of them can be regarded as sets (call the former 'd(A)' and the latter 'T') which have a common a nature, ie a linear ordering. In Kayne (1994, p. 6)'s definition:

Minimalist definition 4.7 (Linear Correspondence Axiom (LCA))**.**
d(A) is a linear ordering of T.

Since linear ordering first and foremost characterizes A-P (*qua* implied by linearity), endocentricity / hierarchy is the only way possible for Syntax to get the same property and be thus compatible with A-P. It is important to keep in mind in this connection that endocentricity / hierarchy has a precise formal content. Linear ordering, in fact, forces the simple relation holding between two items belonging to the same 'syntactic set' (less informally, between two nonterminals of d(A)) to become a more complex geometric relation: Asymmetric c-command. Asymmetric c-command, in its turn, allows two kinds of syntactic structures only: either the one in Fig. 4.5 or the one in Fig. 4.7, this latter being usually referred to as X-bar format.
To get a more complete understanding of Asymmetric c-command, it seems convenient to look at it also from a distributional standpoint: are the syntactic structures depicted in Fig. 4.5 and 4.7 mutually interchangeable (FREE VARIATION) or are they mutually exclusive (COMPLEMENTARY DISTRIBUTION)? Occurrence of more than one Dependent in a given syntactic structures (say, *all, Mahler's* in *all Mahler's symphonies*) arguably offers evidence that the latter solution is to be opted for, as suggested in Cinque (2006, p. 164, fn. 36). The argument is built along the following line of reasoning.
It follows as a corollary from Asymmetric c-command that a Dependent preceding its Head on the time axis (say, *all* in *all Mahler's symphonies*) cannot be merged into the syntactic structure as a Spec of the same phrase hosting the Head (*symphonies*), if this latter is already preceded by another Dependent (*Mahler's*): this would yield the illegitimate syntactic structure depicted

in Fig. 4.8. Accordingly, recursive merger into the syntactic structure of a Dependent other than the first one restarts the process of phrase structure building from its very beginning.
Asymmetric c-command also implies that the new piece of syntactic structure is added to the left of the pre-existing syntactic structure: new nonterminals are allowed to derive the syntactic structure in Fig. 4.7 from that in Fig. 4.5 iff and only if they are merged to the left of this latter. This twofold behavior of Asymmetric c-command evidences that *all the dependents are merged to the left of their Head, the first as its Specifier, the others as phrases higher than the phrase hosting the Head*. This is much more so, if one considers that two phrases merged at the right of XP violate antisymmetry since they c-command each other Kayne (1994, p. 30, 136, fn. 28).
A non-trivial consequence of the antisymmetric approach to Dependents is that a Dependent following the Head (say, the Complement PP *of the theorem* in *John's proof of the theorem*) occurs in this position only subsequent to some displacement, its base generation position being to the left of the Head (Cinque 1999, Cinque 2006, Ch. 5, 6). For example, one may assume that the Head *proof* has been displaced, so that (225) has the following derivational history:

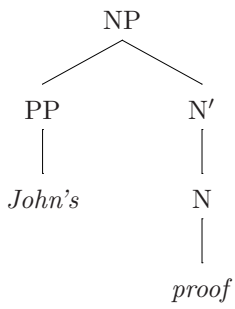
(226) [John's [proof [of the theorem [t_{proof} ...
[Spec [Head [Compl ...

(226) in particular shows that the Compl hosting the Dependent *of the theorem* following the Head *proof*, is itself made up of a Specifier and of a Head. In other words, the Compl of the X-bar format depicted in (224) and in Fig. 4.7 is a cover term for any complex Spec - Head that follows another complex Spec - Head:

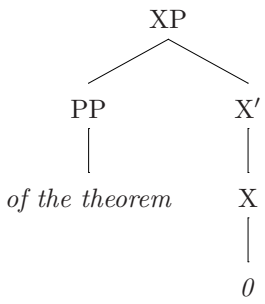
(227) [John's [proof + [of the theorem [t^{proof} ...
[Spec [Head + [Spec [Head ...

The claim that the X-bar format is based on the more primitive structure Spec - Head comes with no surprise for at least two reasons. First, the complex Spec - Head is actually the simplest syntactic structure compatible with A-P that Syntax can generate (see Fig. 4.5 and (219). Second, the X-bar format is no more than a recursive instantiation of the complex Spec - Head and recursion is precisely the task carried out by Syntax (see Section 4.5.1).
This characterization of the X-bar formatis even desirable because it provides an unified account of Dependents: not only *John's* but also *of the theorem* is now interpreted as a Specifier. They only differ for the Head they are combined with: the first dependent co-occurs with a substantive item, all the dependents other than the first co-occur with an item that, whatever is nature, is *not* a substantive item (this issue will be tackled again later).
Furthermore, bringing into comparison the complex Spec - Head with the syntactic structure depicted in Fig. 4.5 reveals that the former can be also identified with the latter:

(228)



(229)



The complex Spec - Head and the twofold instantiation of this complex, generally designated as X-bar format, are thus in complementary distribution in the sense that a pure complex Spec - Head (Fig. 4.5) manifests itself when the first Dependent is combined with a substantive item, while the X-bar format is found when a Dependent other than the first is combined with an item that is not substantive (Cinque, 2006, p. 164, fn. 36).

More generally, Asymmetric c-command derives as corollaries all the properties of the X-bar format which were stipulative in the previous literature (see Cinque 1996, pp. 449ff. for a summary, also quoted in Moro 2000, p. 26).

In particular, it explains *how* movement takes place - irrespective of *what* triggers it.

First, movement is upwards RAISING, not downwards LOWERING. A raised Head or Spec asymmetrically c-commands its trace, which matches the left-to-right directionality of Spell-Out of the corresponding terminals. A lowered Head or Spec would yield the configuration depicted in Fig. 4.6, where the Head or Spec *is c-commanded by* its trace, to the effect that directionality of Spell-Out of its corresponding terminals turns out to be right-to-left, an undesired result (see above). Accordingly, movement becomes synonimous with raising.

Second, Spec-to-Spec movement is allowed, Spec-to-Head is not: Kayne (1994, p. 18) in fact observes that this latter would yield the configuration in Fig. 4.6 [11], which results in the undesired right-to-left directionality of Spell-Out of the corresponding terminals.

Third, Head-to-Head movement is allowed, Head-to-Spec is not: within the same XP, this once again would yield the illegitimate configuration in Fig. 4.6. As for the Head-to-Spec movement outside XP, Cinque (1996, p. 449, fn. 6) remarks that it would violate the Minimal Link Condition (see Section 4.5.1), since in the syntactic structure a given Spec, YP target of the movement (LANDING SITE) is more distant from the Head X to be raised than its Head Y.

Fourth, a Compl, being structurally similar with a Specifier (both of them are Dependents), is treated as a Specifier as far as movement is concerned: Compl-to-Spec movement is thus allowed.

Last but not least, and without entering into details, the observation that Asymmetric c-command prohibits merger of more than one Specifier (see Fig. 4.8) into the complex depicted in Fig. 4.5 can be stated more generally: no more than one nonterminal can be merged into such a complex. Recalling from the discussion immediately above that nonterminals are Specifiers and Heads, *no more than one Head can be added to the Head already found in this complex.*

The antisymmetric approach thus derives (Baker, 1988)'s Mirror Principle, reproduced in the following under Belletti (1990, p. 27)'s formulation:

Minimalist definition 4.8 (Mirror Principle)**.**

[11] Kayne proposes another explanation, which is not reproduced here for the sake of simplicity.

In a given word the respective order of affixes that may be present reflects the syntactic derivation of the word, ie, the order in which the affixation has taken place through syntactic movement.

Belletti (1990, p. 27) exemplifies this statement as follows:

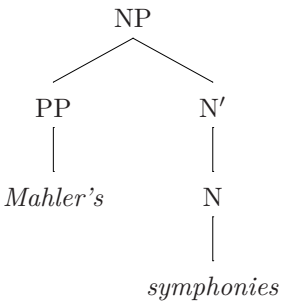
(230) ...[X [Spec [Y... → ...[Y+X [Spec [t...

To recapitulate, Antisymmetry takes movement to be raising of nonterminals only, deriving its following properties:

- Spec-to-Spec movement and Compl-to-Spec movement (or NONHEAD MOVEMENT)
- Head-to-Head movement (or HEAD MOVEMENT)

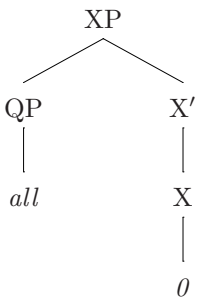
Other kinds of movement (e.g. Head-to-Spec movement) are allowed in other versions of MP (see e.g. Matushansky 2006 and Pesetsky & Torrego 2004 after her) but barred in the Kaynean framework.
With this in mind, it is possible to return to an issue left open above: the exact nature of the Head combined with the Dependent other than the first one. Baker (1988) also refers to the Mirror Principle exemplified in (230) as INCORPORATION.
Recall in fact that under the antisymmetric approach, the first Dependent (e.g. *Mahler's* in *all Mahler's symphonies*) is a Specifier combined with a *substantive* item acting as a Head, which characterizes this kind of phrase as a Spec - Head complex, as illustrated in (231) below (as well as in (228) above):

(231)



On the other hand, any Dependent other than the first (e.g. *all* in *all Mahler's symphonies*) is a Specifier combined with a different item acting as a Head, whose nature warrants further discussion, as illustrated in (232) below (as well as in (229) above):

(232)



This kind of phrase is a Spec - Head complex always associated with a lower complex Spec - Head: the complex Spec - Head *all + 0* actually is followed by the complex Spec - Head

Mahler's symphonies - this latter being thus the 'Compl' of *all + 0*. Recall that this configuration is usually referred to as X-bar format (cp. (224)).
The assumption discussed above that the two just mentioned kind of phrases are in complementary distribution and that the former instantiates a substantive item leads to think that the latter instantiates what in traditional grammar is regarded as the opposite of a substantive item, namely a functional item (see Section 4.2). But, at the same time, this move entails a radical departure from traditional grammar. If, in fact, the Head combined with a Dependent other than the first is functional, that Dependent too will be functional: accordingly, what is traditionally considered to be a substantive (= open) category (e.g. the adjective *all*) is now understood as a functional (= close) category.
Cinque (1994, 1999) develops this line of reasoning and argues that this is actually the case in many non-Indoeuropean languages: here for example adjectives and adverbs, which in the traditional descriptions of Indo-european languages are considered open classes, rather behave as close classes (Cinque, 2006, pp. 4-5 and references therein).
Accordingly, Cinque (1994, 1999, 2006) provides broad typological evidence to unify the notion of Dependent and of functional ('close') category - a position hinted at in Kayne (1994, p. 30). Cinque's proposal also brings crosslinguistical evidence that associating a functional Head with the Specifier hosting the Dependent is empirically grounded: while in Indoeuropean languages a Dependent is realized as a word (ie stands in isolation), in many non-Indoeuropean languages the same Dependent tends to be a morpheme attached to another stem. This difference can be accounted for interpreting the word as a Specifier and the morpheme as a Head belonging to the same (functional) phrase.
A particularly telling example is the English adverb *completely* (a word), if contrasted with its counterpart in Fula Fulfulde[12] - in this language, the same meaning is conveyed by a morpheme attached to the verb, as exemplified in (233):

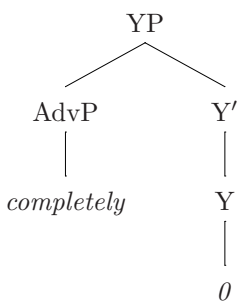
(233) Fula Fulfulde (Cinque, 1999, p. 100)

Kuugal timm-id-i
work finish-COMPLETIVE-TENSE

'the job is completely finished'

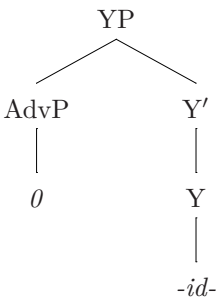
According to Cinque, both of *completely* and *-id-* carry out the same completive function and accordingly are to be assigned the same syntactic structure (say, the phrase YP), but the former sits in Spec, YP, the latter in its Head:

(234)



[12] A Niger-Congo language, mainly spoken in Senegal, Gambia, Guinea, Cameroon and Sudan.

(235)



In addition, Cinque (1994, 1999) capitalizes on the (pre-)minimalist finding that in the syntactic structure of the clause the functional material is systematically ordered as follows Chomsky (1995, p. 55 and references therein):

(236) CP > IP > VP
CP = that... IP = (do)es... VP = eat...

and put forwards that in a likewise fashion, the word order of *functional* Dependents (e.g. adjectives, adverbs) is rigidly fixed *in Syntax* to the left of their *substantive* head (N, V) along the lines of (238).
For example, at this level numerals always precede (= are higher in the structure than) adjectives: on this view, word orders other than (238) documented crosslinguistically are to be accounted for as having undergone movement.
A case in point is Moroccan Arabic, where the numeral can follow the adjective, as shown by (64) discussed in Ch. 3 and here repeated as (237):

(237) Moroccan Arabic (Ihsane, 2003, p. 269)

lə-wlad z-zwiin-in l-xams-a hadu
the-boys the-handsome-PL the-five-*at* these

'these five handsome boys'

Such a state of affairs is interpreted in the cartographic framework as a result of the raising of the adjective over the numeral (see e.g. Shlonsky 2004).
This research program, strictly interlocked with the antisymmetric approach (see Cinque & Rizzi 2008 for details and references) can be summarized as follows, insofar as the (Arabic) NP is concerned (see Cinque 2003, Shlonsky 2004):

Minimalist definition 4.9 (Cartography)**.**

(238) $Dem > Num > A > N$

where:

(239) Dem, Num, A = $[_{XP}$ Spec $[_{X'}$ Head] YP]
N = $[_{XP}$ Spec $[_{X'}$ Head]

(238), in particular, refines the account of the complementary distribution sketched so far:

(240) The substantive item (N) is assigned the following syntactic structure: complex Spec - Head

(241) The functional item (Dem, Num, Adj) is assigned the following syntactic structure: X-bar format

(240,240) translates in syntactic and antisymmmetric terms the traditional notion that a substantive item can occur alone (cp. English *dog*), while a functional item cannot (*dog-s* vs. **-s*): accordingly, in Syntax *dog* may give rise to a Spec - Head complex, *-s* to a Spec - Head complex *always* combined with another Spec - Head complex, that of *dog*: the so-called X-bar format.
The just described functional / close items displaying the X-bar format are often dubbed FUNCTIONAL PROJECTION.
Having explored how Kayne (1994) and Cinque (1994, 1999, 2006) resort to Antisymmetry to derive (the geometry of) phrase structure, it is now possible to see how Moro uses the same notion to derive movement.

4.5.4 Movement: Dynamic Antisymmetry

It has been seen in the previous Section that what Moro calls point of symmetry (POS) is the most basic structure that Syntax yields and that this configuration, albeit perfectly legitimate in proper Syntax, becomes illegitimate when Syntax 'gets in touch' with A-P (as a consequence of an evolutionary change).
Therefore, a POS has to be somehow manipulated to be compatible with A-P: according to Kayne (1994), manipulation is phrase structure building.
Moro (2000) puts forwards that another way of manipulating a POS in order to make it compatible with A-P is movement. Accordingly, the definition of Dynamic Antisymmetry given at the end of Section 4.5.2 can be more accurately restated as follows (from Moro 2004, p. 396):

Minimalist definition 4.10 (Dynamic Antisymmetry)**.**
Movement intervenes if Merge generates a point of symmetry.

Technically, what movement does is eliminating the POS by eliminating one of the two items it is made of. The basic idea behind this proposal is that the problem of a POS is that it is an output of Syntax, where the item A and B are not ordered. Hence A-P is not able to 'decide' whether the item A precedes the item B or vice versa (see the beginning of previous Section). If one of the two items is deleted, this problem no longer exists (an alternative solution is making the item A and B ordered in Syntax by adding more structure, which gives rise to the complex Spec - Head and to the X-bar format: see previous Section).
Recalling that in Syntax a POS bears two kinds of information (syntactic / nonterminal and phonological / terminal: see the previous Section), each of the two items is better seen as a pair nonterminal - terminal: X, x and Y, y. Call this pair a POLE: movement thus eliminates the POS by eliminating one of the two poles it is made of.
More precisely, what movement actually eliminates is not the whole pole, but only what can be seen by A-P and disturbs it, namely the terminal (e.g. x), *not* the nonterminal (e.g. X). This scenario is summarized in Fig. 4.9 (the example reported in this figure will be clarified shortly).
To make the discussion more concrete, consider the NP *John* and the NP *the cause of the riot*, ie the pairs X, x and Y, y. According to Moro (2000, Ch. 3) these two pairs constitute a POS because they are created equal (they have in fact the same categorial status NP) and one of them is hence to be deleted, but what is deleted is either the terminal *John* (x) or the terminal

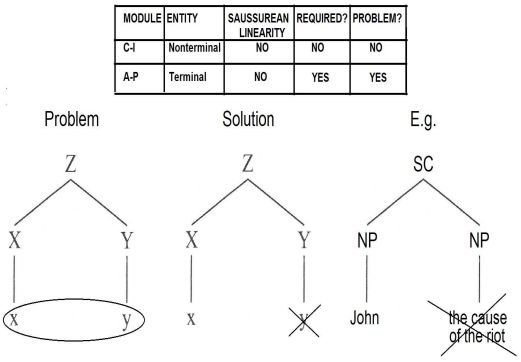


Figure 4.9: Neutralization of a POS

the cause of the riot (y), which by their own nature interact with A-P, *not* the nonterminals NP, NP (X, Y) which have nothing to do with A-P.
Eliminating a pole of the POS therefore consists of turning it into a nonterminal associated with a zero terminal - the so-called 'trace'.
If, however, eliminating phonological information from (a portion of) a POS is fine with the module A-P, this raises a non-trivial problem for the module Syntax: its property of Encapsulation / Structure Preservation Principle, in fact, imposes that no information be eliminated in the course of computation (see beginning of Section 4.5).
In order to overcome this *impasse*, the pole going to be affected by phonological deletion, be it either X,x or Y,y, is copied into another syntactic position before that this operation takes place. Plainly, the position at issue in the best case is part and parcel of a configuration compatible with A-P, that is not a POS itself[13]: this is the empty Spec of a higher functional projection.
On this vantage point, the widespread (pre-)minimalist postulate that a moved element leaves a trace, and that the former, but not the latter must have a phonological realization, is straightforwardly accounted for by Dynamic Antisymmetry. As Moro puts it:

> "...deletion of phonological features is not to be regarded as a side phenomenon with respect to movement as an independent phenomenon. Rather, it constitutes an essential part of movement: movement is nothing but deletion of the phonological features of either element constituting a point of symmetry and copy of the same element in a suitable LCA compatible position, pace locality conditions. This also gives us a principled reason as to why the topmost "copy" is pronounced: the "original" element in the low position cannot be pronounced because the LCA prevents the structure to be linearized. In fact, this is why movement takes place.".
> (Moro, 2004, pp. 396-397)

The most representative and well-studied example of POS undergoing movement occurs in primary predication.
Consider, in fact, a copular sentence like *John <u>is</u> the cause of the riot*). It is assigned in the (pre-)minimalist literature the basic structure (see Moro 2000, Ch. 3 for references and details):

(242) [[NP] [NP]]
[[John] [the cause of the riot]]

[13]This latter case is discussed at length in Section 5.3.3 of the next Chapter.

where no *be*-verb is posited, on the basis of both its secondary predication counterpart (*I consider [Maigret very intelligent]*) and crosslinguistic evidence (cp. CA *Zaydun ḥasan*, lit. 'Zayd handsome', ie 'Zayd is handsome').

The basic structure in (242), *John - the cause of the riot*, is generally referred to as SMALL CLAUSE (SC) and, in Moro's view, is a POS (see immediately above and Figg. 4.9 and 4.10).

Hence either the subject *John* or the predicate *the cause of the riot*, being a pole of a POS moves

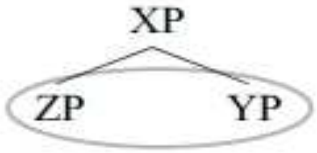
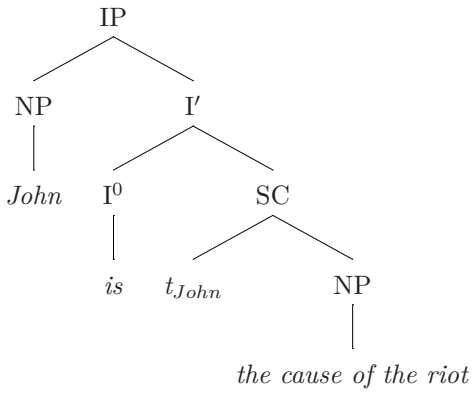


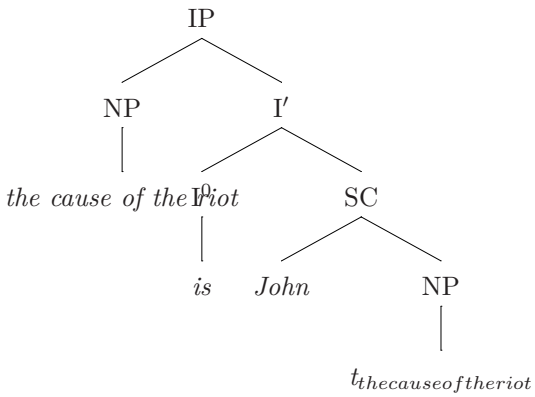
Figure 4.10: The POS XP - XP (Moro, 2000, p. 32)

to the higher Spec available. Since primary predication by definition conveys information about tense (e.g. *is*), this latter is a functional projection which constitutes an appropriate landing site for either the subject *John* or the predicate *the cause of the riot*:

(243) Canonical Predication



(244) Inverted Predication



As shown in (243), in Moro's view the familiar kind of copular sentence (CANONICAL PREDICATION) is no more than a strategy to break the symmetry of the small clause through movement

of the subject.
On the other hand, the copular sentence featuring a fronted predicate (INVERTED PREDICATION) depicted in (244)[14] is taken by Moro to be the same strategy that in this case deploys movement of the predicate (Moro, 2000, Ch. 3, Appendix): in this light, the copula is no more than a 'hub' facilitating the movement needed (Moro, 2000, p. 120, fn. 20). Notice that *movement of the predicate* corresponds to the more traditional notion of PREDICATE INVERSION.
As for secondary predication, of particular interest here is that Moro (2000, 2004) proposes that the POS depicted in Fig. 4.11 occurs in this domain.
In this POS, the items 'on the same footing' are two Heads and thus it can be also designated

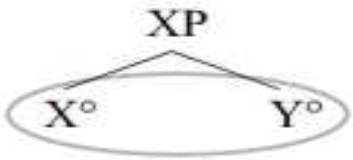


Figure 4.11: The POS Head - Head construction (Moro, 2000, p. 32)

as HEAD - HEAD CONSTRUCTION (Moro, 2000, pp. 84-85).
Moreover, Moro (2000, pp. 49ff.) contends that it is this kind of POS that explains the so-called *was-für* split construction in Germanic languages, where the *wh*-element is raised alone, leaving in situ other material associated with it. (245) exemplifies this phenomenon in Dutch: while in English the items *what novels* cannot be separated, their Dutch counterparts, namely *wat...romans* are detached one from the other.

(245) Dutch (Moro, 2000, p. 50)

wat heeft hij voor romans geschreven?
what has he for novels written

'what novels has he written?'

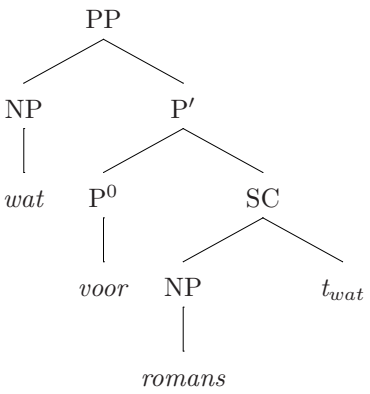
According to Moro, the derivation of (245) runs as follows (recall from Section 4.4 that in MP derivations are step-by-step processes). The NP *romans* and the NP *wat* are base-generated as a small clause:

(246) [[NP] [NP]]
[[romans] [wat]]

This gives rise to a POS, which is neutralized moving *wat* to Spec, PP (P the preposition *voor*), as illustrated in (247):

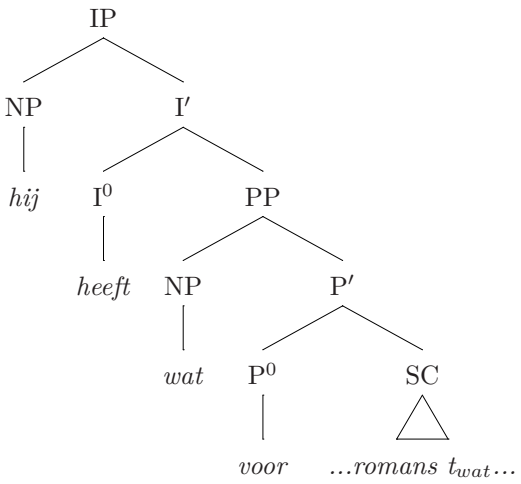
[14]Fronting of the Predicate over the Subject is designated in literature also as INVERSE PREDICATION or REVERSE PREDICATION and these terms are (almost) synonimical, regardless of the domain this phenomenon is observed in: either primary or secondary predication (e.g. the modification structure, where the Subject is the modified noun and the Predicate its modifier: see the next Chapter). Recently, however, den Dikken (2006) uses the terms Inverted / Inverse Predication and Reverse Predication with two different meanings. In his theoretical framework, Inverted / Inverse Predication indicates the configuration Predicate > Subject deriving from some displacement, whereas Reverse Predication indicates the same configuration base-generated as such. See Section 5.3.2 for more discussion..

(247)



The derivation proceeds and therefore the auxiliary *heeft* and then the subject *hij* are merged into the syntactic structure - as a result, not only *wat* (subsequent to movement) sits in Spec, PP but is also adjacent to the auxiliary *heeft*, because this latter is base-generated immediately above PP. This is depicted in (248):

(248)



At this point, Moro (2000, p. 55) observes that *wat*, very much as clitics (Italian *lo*, French *le* etc...) syntactically has an ambiguous status.

On one side, *wat* and clitics behave as phrases (in technical sense, ie X-bar configurations) because they do not tolerate addition of Dependents (see the discussion surrounding Fig. 4.7 and Fig. 4.8).

On the other side, they behave as Heads: these latter, in fact, in the vast majority if not all of the examples reported by Cinque (1999, Ch. 3) are monosyllabic (contrast e.g. English *completely* (234) with Fula Fulfulde *-id-* in (235)) and so are clitics and *wat*.

This is why clitics in Kayne (1994)'s standard antisymmetric approach have an interesting behavior: they function both as phrases and as Heads (Kayne, 1994, pp. 17-21 and references therein).

Moro pursues this reasoning to its extreme limits: if *wat* functions (also) as an Head, is 'on the same footing' than the Head *heeft* and therefore the configuration in which *heeft* and *wat* are adjacent, depicted in Fig. 4.11, can be likened to the Head - Head construction in Fig. 4.12.

In other words, *heeft* and *wat* create a POS, which is neutralized triggering movement of *wat* to the left of *heeft*: this yields the word order *wat heeft (hij voor romans geschreven)* (245).

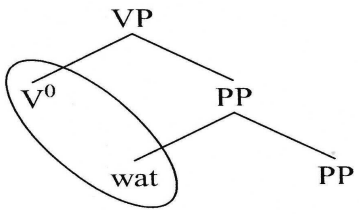
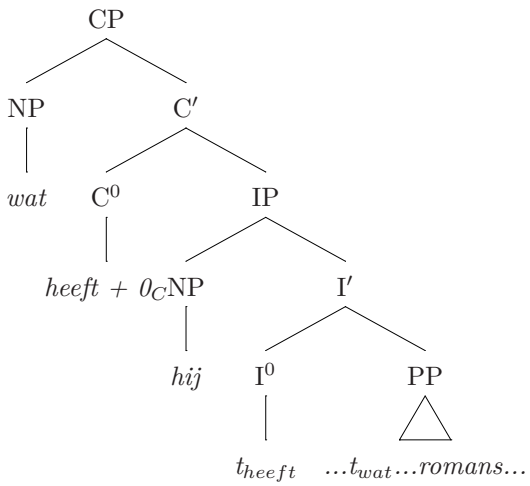


Figure 4.12: The POS Head - Head (Moro, 2000, p. 32)

In greater detail, the derivation takes a further step, and the functional projection CP is merged into the syntactic structure, immediately above IP (cp. (236)). *Wat*, *qua* Specifier, raises to the first Specifier immediately above IP, namely Spec, CP, in order to neutralize the POS, while *heeft*, *qua* Head raises to the Head C. Setting aside for the time being both the exact nature of CP and the rationale for movement of *heeft* (which will be clarified in due course), the just described derivational step can be represented along the following lines:

(249)



Before concluding this Section, it is worth noting that Moro's (Dynamic) Antisymmetry is not at odds with the original version of Antisymmetry à la Kayne (1994).

First, the two potential sources of POS discussed here, ie the ambiguous nature of clitics and the small clause are given a theoretical status also in the Kaynean Antisymmetry (Kayne, 1994, Ch. 3, 7). Second, already in Kayne's view *movement is very loosely related with C-I.*

While in MP it is standardly assumed that movement can also take place at C-I (COVERT MOVEMENT), Kayne (1998) and Cinque (2006, Ch. 6) show that movement taking place in Syntax (OVERT MOVEMENT), if combined with the antisymmetric phrase structure described in the previous Section is able to account for the phenomena usually explained in terms of covert movement, making it possible to dispense with this latter.

Finally, Moro's Dynamic Antisymmetry is not only a theory about the module Syntax (syntactic movement is triggered by the need for Syntax to be compatible with phonological linearity of A-P) but is also a theory about predication.

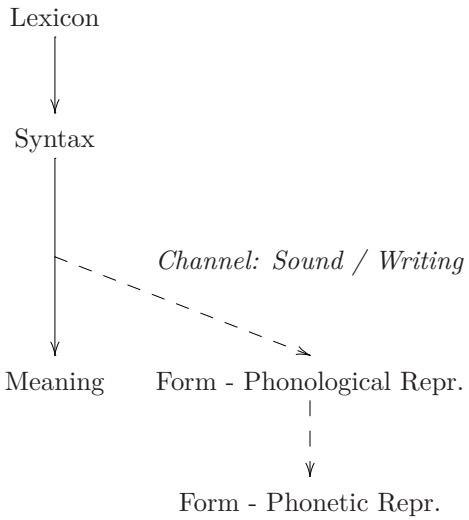
Upon closer scrutiny, in fact, a predicational profile can be assigned not only to the small clause *John + the cause of the riot* (242), but also to the complex *romans + wat* Moro (2000, pp. 51ff.). *Wat*, in fact, acts as a modifier of *roman* 'novel' as if it were an adjective (say, 'nice'),

and modifiers can be taken to be predicates because they ascribe a quality to the noun they modify - hence conceivable as the subject (see Section 5.1 for more details on this issue).
This allows, at least in principle, to call Dynamic Antisymmetry into the picture whenever a (syntactic) predication theory is taken into account, as it will be seen in the next Chapter.

4.6 Summary

The theoretical framework of the present dissertation is (Dynamic) Antisymmetry. This is a branch of MP, the most recent incarnation of the Chomksian linguistic paradigm.
MP assumes the following architecture of grammar:

(250)



and founds it in biological and evolutionary terms. Syntax is a key-ingredient of human language only subsequent to some evolutionary process that exapted Syntax away from its original adaptive function (recursion). This evolutionary shift was driven by communicative needs and turned Syntax and A-P from separate to interacting modules.
Once Syntax and A-P were no longer incommunicado (as it still happens in other primates etc...), some cognitive strategies were needed to ensure exchange of information between these so different systems.
Basically, Syntax and A-P differ in that items at A-P are ordered one after the other because their pronunciation puts them on the time axis (Saussurean linearity), which is clearly not the case for Syntax. When in Syntax, in fact, items are not pronounced and therefore not located on the time axis - hence are not ordered one after the other.
It is an empirical hypothesis that a strategy used to put Syntax 'in touch with' A-P has to do with phrase structure and is Bloomfieldian endocentricity / hierarchy. Under this perspective, endocentricity / hierarchy observable in Syntax is *not* a primitive property, but something modelled after Saussurean linearity at A-P and obtains iff and only if two items are put together in a special way, one that 'mimics' Saussurean linearity at A-P.
This special combination of items is either the complex Spec - Head or the X-bar format.
These two syntactic structures are in complementary distribution as follows:

(251) The substantive item (e.g. N) is assigned the following syntactic structure: complex Spec - Head

(252) The functional item (e.g. Dem, Num, Adj) is assigned the following syntactic structure: X-bar format

This hypothesis was first put forward by Kayne (1994) and is designated as Antisymmetry. Further developments include Cartography (Cinque & Rizzi, 2008, and references therein), which lend crosslinguistic evidence in favor of (251,252).
The Kaynean Antisymmetry derives the following properties of the phrase structure:

- Phrases having a syntactic structure other than (251,252) are barred
- Functional projections (252) are merged to the left of the substantive projection (251)
- Movement is only upwards (raising)
- Only Spec-to-Spec, Compl-to-Spec and Head-to-Head movements are allowed
- A moved Head attaches to the left of another Head

More recently, Moro (2000) proposes that an alternative strategy used to put Syntax 'in touch with' A-P has to do with movement. Under this approach, Bloomfieldian endocentricity / hierarchy obtains not only by (a special) combination of two items but also through their displacement - hence the designation 'Dynamic Antisymmetry'.
The basic idea behind this proposal is that the problem of combination of two items is that it is an output of Syntax, where the item A and B are not ordered (see above). Hence A-P is not able to 'decide' whether the item A precedes the item B or vice versa. If one of the two items is deleted, this problem no longer exists.
(Unordered) combination of two items is designated as Point of Symmetry (POS). Movement thus eliminates the POS by eliminating one of the two items it is made of. The item is eliminated from the POS by displacing it to another portion of the syntactic structure.

Chapter 5

Predication in the NP: a crosslinguistic survey

5.1 A syntactic and crosslinguistic theory of predication

An accurate description of what predication is falls far beyond the scope of this work. Its main intention, in fact, is to provide a predicational rationale for the behavior of the adnominal markers and of the modifiers in the Arabic NP (see Ch. 1).

Given that the core of the proposal is likening the Arabic NP to the Numeral Classifier contructions found in the Southeast Asian languages *on syntactic grounds*, it seems sufficient here to lay down a *syntactic* account of predication formulated on the basis of the analysis not only of Indoeuropean and Semitic languages but also of the above mentioned Numeral Classifier languages of South East Asia.

Den Dikken's THEORY OF RELATORS AND LINKERS is a good candidate for this purpose because, empirically, it covers the three-fold data sample needed in this work (see e.g. den Dikken 2006, den Dikken & Singhapreecha 2004 on Indoeuropean and Numeral Classifier languages and den Dikken 2007 on Semitic).

Den Dikken's theory has also a theoretical advantage: it is based on the antisymmetric (and cartographic) scenario depicted in Section 4.5.3 (see e.g. den Dikken 2006, p. 233 as well as pp. 17, 40 59 and den Dikken & Singhapreecha 2004).

It should be also noted in this connection that den Dikken (2006) albeit explicitly rejecting Moro's POS and therefore opting for a feature-checking account of movement, nonetheless divorces movement from C-I in some cases. In dealing with particular instances of Predicate Inversion (involving *wh*-in-situ, quantification and Antecedent-Contained Deletion, which will not be further discussed here), he either proposes to dispense with the covert movement analysis previously worked out in literature (den Dikken, 2006, pp. 136-142) or mentions alternative accounts which do not resort to it at all (den Dikken, 2006, pp. 130-31, 131-35).

Den Dikken maintains that his account is based on a *syntactic* notion of PREDICATE, essentially equivalent to Aristotle's *kategoroúmenon*, ie the syntactic constituent denoting the property ascribed to another syntactic constituent (den Dikken, 2006, p. 8). By the same token, Den Dikken takes the term SUBJECT to generally refer to the syntactic constituent to which this

property is ascribed. Consequently, this notion does not necessarily correspond to the EXTERNAL ARGUMENT (or AGENT), nor is it necessarily identified with the so-called GRAMMATICAL SUBJECT, thought of as the constituent fulfilling the subject position of the sentence.
On these grounds, both a noun combined with a modifier and a topic combined with a comment are regarded by den Dikken (2006) as subject - predicate constructions (PREDICATION STRUCTURE):

(253) English (den Dikken, 2006, pp. 9-11, 25)
The Earth^{SUBJECT} is round^{PREDICATE} → COPULAR SENTENCE
A beatiful^{PREDICATE} flower^{SUBJECT} → NOUN + MODIFIER
Brian^{SUBJECT}, Imogen really adores^{PREDICATE} → TOPIC + COMMENT

Den Dikken (pc) also takes a determiner combined with a noun to be another instance of predication structure, along the lines of Campbell (1996) and Giusti (2002) (see also Ouhalla 2000 for similar arguments for Arabic):

(254) This^{SUBJECT} book^{PREDICATE} → DETERMINER + NOUN

In spite of its appeal, den Dikken (2006)'s has significant shortcomings. First and foremost, as Tanase Dogaru (2008, p. 302) remarks, den Dikken (2006) sometimes implements his ideas about the NP more through evidence from primary predication than through evidence from the NP in and by itself. A case in point is den Dikken & Singhapreecha (2004)'s and den Dikken (2006)'s account of the Numeral Classifier Constructions, whose syntactic structure is left vague (see fn. 11). Second, in some cases he posits in the syntactic structure some items whose presence is not justified on semantic grounds (see fn. 3). Third, movement applies for at least two different reasons, which threatens the internal coherency of the theory (see Section 5.2.5). Finally, den Dikken (2006)'s analysis relies upon some syntactic structures that are not totally in line with the antisymmetric framework supposed to derive them, as it will be seen in Section 5.3.1.
The present Chapter proposes a substantial revision of den Dikken (2006)'s minimalist theory of predication, which aims at fixing the just mentioned problems. In particular, it provides empirical and theoretical evidence in favor of a more economical version of den Dikken (2006)'s theory. This impacts on both the generative and transformational components of syntax (phrase structure and movement). As for the phrase structure, it will be argued that

- all the manifestations of predication in the NP involve a predicate that can be likened to a clause

This is the topic of Section 5.2. As for movement, it will be argued that:

- a strict application of Kaynean Antisymmetry to the manifestations of predication in the NP implies motivating movement in terms of Moro (2000)'s Dynamic Antisymmetry

This is the topic of Section 5.3. In line with den Dikken (2006), the empirical coverage of data corroborating the present view on predication of the NP is threefold: Semitic, Indoeuropean and Numeral Classifier languages of South East Asia.
Data from Indoeuropean and Numeral Classifier languages are set forth and discussed in this Chapter, while data from Semitic languages are the subject of the next two Chapters. Semitic

and especially Arabic, in fact, deserve a separate treatment because of the interpretive problems their NP raise (see Ch. 2 and Ch. 3). It will be shown in the next two Chapters that the syntactic theory of predication developed in this Chapter sheds new light on such problems.

5.2 Functional items and the theory of predication

The syntactic constituents dealt with in the previous Section are the SUBSTANTIVE parts of predication structure[1] and even from the brief account just sketched out it can be easily seen that Den Dikken's approach, in this respect, owes much to more familiar linguistic theories.
What is new in den Dikken (2006)'s proposal, however, is its focus on those elements of predication structure that have no substantive meaning. Since these *functional* elements will play a key-role in (re)interpreting the Arabic adnominal markers, as well as their nature and tasks, it seems convenient to discuss them in greater detail in this Chapter.
Den Dikken argues that such elements are substantially of two kinds, namely the RELATOR and the LINKER: the former carries out the task of establishing a (Canonical) Predication relationship, the latter that of manipulating it giving rise to the so-called Inverted Predication[2].
In more familiar terms, the Relator corresponds to the copular element found in the Canonical Predication, the Linker to that found in the Inverted Predication. For example, in Dutch the particle *als* acts as a Relator, the particle *van* as a Linker, and in its English equivalent, the particle *like* is a Relator, the particle *of* is a Linker:

(255) Dutch (den Dikken, 2006, p. 174)

een	vent^{SUBJECT}	als^{RELATOR}	een	beer^{PREDICATE}
a	bloke	like	a	bear

'a bloke like a bear'

(256) Dutch (den Dikken, 2006, p. 174)

een	beer^{PREDICATE}	van^{LINKER}	een	vent^{SUBJECT}
a	bear	of	a	bloke

'a bear of a bloke'

According to den Dikken (2006), however, Canonical Predication and Predicate Inversion are not created equal, because semantically the latter seems to imply the former. He captures this insight in syntactic (and transformational) terms interpreting Inverted Predication as an instance of Canonical Predication which is associated with a Linker and then undergoes preposing ('movement') of both the Relator and the Predicate, the preposed Relator combining with the Linker. Simply put, Predicate Inversion (256) is Canonical Predication (255) plus movement.
This is the core of den Dikken (2006)'s idea that, as a first approximation, can be represented as follows:

(257) Subject + Relator + Predicate → Predicate + Relator + Linker + Subject

[1] But recall from Section 4.5.3 that under Cinque's approach, the predicate - adjective is not necessarily to be considered as such (Cinque, 1994, 1999, 2006).

[2] Not to be confused with Reverse Predication: see fn. 14 in the previous Chapter. Reverse Predication will be discussed in greater detail in Section 5.3.2.

Thus, what seems to be a simple Linker at phonological level (say, *van, of* of (256)) is actually a more complex copula, consisting not only of a Linker, but also of a null Relator (*0 + van, 0 + of*).
Another detail to keep in mind as a guideline throughout the following discussion is that, according to den Dikken (2006), the Predicate of the predication structure is typically introduced by a null Head. Now, given that the Relator itself *qua* functional item is a Head (see end of Section 4.5.3), the Canonical Predication structure is characterized by adjacency of the Relator Head and of the null Head introducing the Predicate:

(258) ... Relator Head + $[_{\text{Predicate}}$ Null Head ...

Last but not least, *any* functional Head in principle can function either as a Relator or as a Linker, provided it is able to mediate a predication relationship between a Subject and its Predicate. Consequently, the (Canonical) Predication relationship is thought of by den Dikken (2006) as a given functional projection in X-bar format (see end of Section 4.5.3 and (241)):

(259) Predication = X-bar format
Semantics = Syntax

where the Spec functions as a Subject, the Head as a Relator, the Complement (containing a Null Head) as a Predicate:

(260) $[_{\text{XP}}$ Spec $[_{\text{X}'}$ Head $[_{\text{YP}}$ Compl]]] = XP
Subject Relator Predicate = functional projection

Then, the functional projection represented in (260) is generically referred to as a projection RP headed by the Relator R, but, as den Dikken (2006, p. 29) remarks, this is merely a notational artifact, not a real and dedicated functional projection. Actually, in the best case, it seems to be appropriate to assign RP a precise identity.
In a likewise fashion, the Inverted Predication is conceived by den Dikken (2006) as follows:

(261) $[_{\text{XP}}$ Spec $[_{\text{X}'}$ Head+Head $[_{\text{YP}}$ Compl]]] = XP
Predicate Relator+Linker Subject = functional projection

In this case too, for the sake of commodity, the Head performing the function of a Linker can be generically referred to as F, and the corresponding functional projection as FP, but the same remarks made before remain valid.
These are the central tenets of den Dikken (2006)'s theory, which will be exemplified and motivated in the next Sections.

5.2.1 The Relator

Den Dikken's account raises the question of why in Dutch and English the Canonical Predication structure underlying *een beer van een vent / a bear of a bloke* (256) undergoes Predicate Inversion, while the Canonical Predication surfacing as *een vent als een beer / a bloke like a bear* (256) does not.
On den Dikken (2006)'s view, the answer lies in that, all the other things equal, in the Canonical Predication that will turn into an Inverted Predication (256) the Relator Head is *not* coindexed with the null Head introducing the Predicate, while in Canonical Predication that will remain as such (255) it is.

In what follows, it will be first discussed *how* coindexing between two Heads arises and then *why* its absence blocks movement while its presence triggers it.
To begin with, it is important to remark that the the null Head in question, on closer examination, might be plausibly identified with a *null pronoun-like element* ('pro')[3].
A second thing to note is that the distinctive *property of coindexing found in als / like (255) correlates with another property of their, namely that in diachrony they are particles grammaticalized* out of an adverb and an adjective, respectively. Den Dikken in fact points out that Dutch *als* is a combination of the quantifier *al-* (able to function as Relator) and of the adverb *-s* (cp. their English equivalents *all, so*), while English *like* is etymologically connected with the adjective *like* still found in the relic forms *like minds, a like sum* (den Dikken, 2006, p. 291, fn. 19).
It could be added that the adverb *so* still in present-day English can function as a (dummy) predicate:

(262) English (Chomsky, 1957, p. 67)
Bill is my friend and so is John

and that in literary English *like*, when not grammaticalized, can also work as a noun (cp. the article preceding it) in the relative clause *NP the like of which is NP*:

(263) English (D. A. Grandy, *Everyday Quantum Reality*, p. 19)
Bohr is proposing that quantum physics involves more than cognitive mastery of rules and principles, the like of which is a soft pillow for drifting off into ideological slumber

What *so* and *the like of which* have in common in constructions like (262, 263) is that both of them fulfill the topmost position of the clause they belong to. Crosslinguistically, in this portion of the (subordinate and main) clause there is also found a marker deployed to signal that a new clause is beginning, the so-called COMPLEMENTIZER, which is generally a grammaticalized pronominal form.
Two cases in point are the English relative clause *that I saw* introduced by the pronominal form *that* in the sentence *the man that I saw* and the CA particle *'inna*, of demonstrative origin (Fleisch, 1961, vol. II, p. 62), in sentences like *'inna Zaydan munṭaliq*, ' *'inna* Zayd (is) leaving'.
In the (pre-)minimalist literature, this functional marker is taken to be the Head (C, the standard abbreviation for Complementizer) of a dedicated projection (CP) and *wh*-phrases like *the like of which* are taken to sit in Spec, CP on the basis of the following (diachronic) evidence:

(264) Middle English (Radford, 2004, p. 229)
a. ... in every peril which that is to drede ...
b. ... in every peril $[_{\mathrm{CP}}$ which $[_{\mathrm{C'}}$ that [is to drede ...

This makes it possible to take relative clauses where the C *that* is apparently lacking to feature a null counterpart of this pronominal form, ie pro. Thus, the modern English equivalent of (264) would be:

[3]Den Dikken has it that this null Head is a different kind of abstract element, namely the feature *LIKE*, already encoded in the Dutch and English morphemes *als / like* (den Dikken, 2006). A double occurrence of the same piece of meaning, however, sounds a bit. Apart from this, identifying the null Head with pro seems more appropriate, because it makes it possible to realign this instance of Canonical Predication with a more general pro-pattern also found in Predicate Inversion, as it will be seen in the next Sections.

(265) Modern English

a. ... in every danger which is ...

b. ... in every danger $[_{CP}$ which $[_{C'}$ pro [is ...

This holds valid also for the relative clause *NP the like of which is NP* and, in addition, for *so*, because it has the same distribution of *the like of which*. Therefore, not only *who* of (264, 265), but also *the like of which* and *so* sit in the Spec of the projection headed by C / pro:

(266) ... $[_{CP}$ the like of which / so $[_{C'}$ pro ...

Another property shared by *the like of which, so* is that they both enter into a predication relation with the NP following them, *the like of which* as Subject, and *so* as Predicate (cp. *so is John* above).

It is worth staying with English *the like of which* construction a little bit more to consider the behavior of the phrase *of which*, which can become covert because it is phonologically deleted. Technically, the relative clause in which deletion of a relative marker has taken place is designated as a REDUCED RELATIVE CLAUSE. This phenomenon is illustrated by the following contrast:

(267) a. English (J. R. Lawler, *Rimbaud's Theatre of the Self*, p. 163)
This is prose of a special kind, the like of which is not found elsewhere in the Illuminations

b. English (J. Pinkerton, *The Curious and Remarkable Voyages and Travels of Marco Polo*, p. 116)
Other mountains also in this province yield stones called Lapis Lazuli, whereof the best azure is made, the like is not found in the world

In both these examples, *the like of which* introduces a verbal predicate (VP), but (267b) differs from (267a) in that the relative pronoun *of which* has a covert realization, which can be notationed as *Op*:

(268) the like + of which + VP (267a)
the like + Op + VP (267b)

This scenario brings out a general alternation *the like of which VP - the like VP* that finds a striking parallel in the pair *NP the like of which is NP* (263) - *NP like NP* (255): this allows to interpret *NP like NP* of (255) as a construction made up of a NP (say, *a bloke* in (255)) followed by the Reduced Relative Clause *like NP* (e.g. *like a bear* in (255)). In this light, *a bloke like a bear* in (255) is a paraphrase of *a bloke, the like of which is a bear*.

But den Dikken (2006, p. 291, fn. 19) reports an interesting piece of data about the construction *NP like NP* (255) that calls for a refinement of this description. He points out that in the construction under discussion *like* is the outcome of the fusion of two particles: a quantifier, plausibly performing a copular function (ie acting as a Relator), and *like*. The quantifier is no longer detectable in English *like*, but it is in its Dutch counterpart *ge-lijk*, where *-lijk*, the cognate of English *like*, is preceded by a quantificational element *ge-*. This realigns English *like* - actually, something like 0_{ge} *+ like* - with Dutch *als*, itself made of the quantifier *al(l)* and of *s(o)* (see above). This is summarized in (269) (for the position Spec, CP of both *like* and *so*, see (266)):

(269) Relator: (ge) + Spec, CP: like → like (English)
Relator: al + Spec, CP: s(o) → als (Dutch)

Phonological fusion between the Relator *0$_{ge}$ / al-* and *like / -s* according to Den Dikken results in a reinterpretation of the original *meaning* of these functional items, to the effect that *like / so* no longer function as a noun and an adverb respectively, but as Relators: they take over the function of *ge / al*.

This phenomenon is widely documented in the typological literature, where it is designated as REANALYSIS. In particular, Hopper & Closs Traugott (2003, p. 51) follow Langacker (1977, p. 58) in defining it as follows: 'change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation'. Hopper & Closs Traugott (2003, pp. 40-41) also contend that the most frequent instance of reanalisys is fusion, which consists of 'combining of two or more words into one, usually with consequences for semantics, morphology, and phonology (and sometimes also syntax) from the perspective of the new whole word and the former individual parts'. Simply put, reanalysis is some reinterpretation (e.g. of the syntactic structure) and has to do with fusion of two items into one - what is traditionally called UNIVERBATION or WORD FORMATION.

On these grounds, the grammaticalization process observed by den Dikken (2006) can be seen as as a typical instance of *reanalysis*. Consider *like* first.

In this case, plausibly reanalysis is triggered by the covert realization of the relative marker: the complex Subject *the like Op* of the Reduced Relative Clause, is reinterpreted as a simple Subject *Op*, while the noun preceding it, ie *like* is reinterpreted as a Relator and therefore fused with *0$_{ge}$*:

(270) 0_{ge} + <u>like</u> Op →
0_{ge} <u>like</u> + Op

The same representation, laid out in minimalist terms, is:

(271) $[_{R'}\ 0_{ge}$ $[_{CP}$ <u>like</u> + Op $[_{C'}$ pro →
$[_{R'}\ 0_{ge}$ + <u>like</u> $[_{CP}$ + Op $[_{C'}$ pro

More precisely, the item *like* resulting from the reanalysis is a Head conveying the (functional) meaning of comparison (den Dikken, 2006, p. 172), to the effect that the real nature of RP is a dedicated functional projection Deg(ree)P(hrase). The comparative item *like* introduces a Reduced Relative Clause, which in its essence is a CP (see above):

(272) $[_{RP}$ Spec $[_{R'}$ Head $[_{YP}$ Compl]]] = $[_{DegP}$ Spec $[_{Deg'}$ like $[_{CP}$ Compl]]]
Subject Relator Predicate = Subject Comparison Predicate

As clearly shown by (270), the property of *locality* (see end of Section 4.5.1) plays a key-role in the reanalysis, given that the Relator / Head and Spec, CP are *adjacent*.

Notice also that in (270) *like*, being originally a Subject sitting in Spec, CP share some features with its Head C (e.g. agreement): feature sharing, technically COINDEXING, in MP in fact is taken to be the result of the fact that two nonterminals occur within the complex Spec - Head discussed in the previous Chapter. This phenomenon is designated as SPEC-HEAD RELATION, whose definition (adapted from Chomsky 1995, p. 86) is a follows[4]:

[4]The term SPEC-HEAD AGREEMENT is also used in the literature, but this is avoided here because it creates confusion with the traditional notion of AGREEMENT. In MP the traditional notion agreement is no more than

Minimalist definition 5.1 (Spec-Head Relation)**.**

Spec-Head relation provides X with (a) feature(s) of Spec, XP and the other way around.

Specifically, that Spec, CP and C have some features in common (ARE COINDEXED) is evidenced by the fact that in French the complementizer *que* (cp. *Je dis que...* 'I say that'), a Head, is sensitive to the covert relative pronoun Op sitting in Spec, CP to the effect that when this latter is Subject, *que* turns into *qui*. For example, in the sentence *C'est Luc qui vient* 'It's L. who's coming' the beginning of the relative clause can be analyzed along the following lines (Rowlett, 2007, p. 226 and references therein):

(273) ... $[_{\text{CP}}$ $\text{Op}_{SUBJECT}$ $[_{\text{C}'}$ qu-$\text{i}_{SUBJECT}$...

Thus, *like* prior to reanalysis is coindexed with pro located in the Head C because it sits in Spec, CP (Spec - Head relation) and therefore *like* licenses pro via a Spec - Head relation. After reanalysis, *like*, albeit reinterpreted as a Head / Relator, preserves its coindexing with pro located in the Head C:

(274) $[_{\text{R}'}\ 0_{ge}$ $[_{\text{CP}}\ \underline{\text{like}}_i + \text{Op}_i\ [_{\text{C}'}\ \text{pro}_i \rightarrow$
$[_{\text{R}'}\ 0_{ge} + \underline{\text{like}}_i\ [_{\text{CP}}$ $\text{Op}_i\ [_{\text{C}'}\ \text{pro}_i$

This is much more so because *like*, by virtue of its *overt* realization, is the only item coindexed with pro that has the ability to license it: true, Op sits in Spec, CP and therefore potentially licenses the pro occupying C by entering into a Spec - Head relation with it, but de facto it is not able to do so because it is covert as much as pro is.

In a likewise fashion, it seems plausible to assume that the grammaticalization à la Den Dikken of Dutch *-s* into a Relator is a consequence of reanalysis, because of the strict structural similarity between Dutch *-s* and English *like* (269): in this light, Canonical Predication structures like *een vent als een beer* (255) develop out of a different meaning.

Since the English (former) noun *like*, before fusing with 0_{ge} is an elliptic form of *the like of which is*, Dutch *-s*, prior to its fusion with *al*, might be plausibly accounted for along similar lines. The dummy-predicate function of the English equivalent of Dutch *-s*, namely *so*, in sentences like *Bill is my friend and so is John* (262) strongly suggests that Dutch *een vent als een beer* (255) is an elliptic form of a syntactic structure that can be represented in quasi-English as follows: *(Bill is) a bloke and so is a bear*. Recall that the dummy-predicate *so* in this type of sentences is an instance of Inverted Predicate.

If this reasoning is on the right track, what triggers reanalysis is plausibly the ambiguous interpretation of the NP following *so*, e.g. *John* in *(Bill is my friend) and so is John* of (262) and *a bear* of the quasi-English paraphrase *(Bill is a bloke) and so is a bear*.

John and *bear*, being in sentence-final position can be seen not only as the postponed Subject of the Inverted Predicate *so*, but also as the Predicate of a Canonical Predication structure, especially if there is no overt element intervening between the Inverted Predicate *so* and the postponed Subject NP, which can be unambiguously understood as a Linker.

---

a subcase of Spec - Head Relation: a NP sitting in the Spec of a given phrase (say, *he*) shares its features of gender, number, case with the verb (say, *runs*), sitting in the Head of the same phrase. This is exemplifed in (1):

(1) ... $[_{\text{IP}}$ He_{SG} $[_{\text{I}'}$ (run)-s_{SG}...

.

As a consequence, a new null Subject Op introducing a reduced relative clause replaces *-s*, which is reanalyzed as a Relator and thus fused with the Relator *al-*. It ensues that the syntactic structure is no longer an elliptic form of a conjoined clause *'(Bill is) a bloke and so is a bear'*, but is reinterpreted as a Reduced Relative Clause, where Op acts as a covert Subject and *bear* as a Predicate: something like *a bloke, who is a bear*. The process of reanalysis can be illustrated as follows:

(275) al + $\underline{\text{s}}$ →
al $\underline{\text{s}}$ + Op

The same representation, laid out in minimalist terms, is:

(276) $[_{\text{R}'}$ al $[_{\text{CP}}$ $\underline{\text{s}}$ $[_{\text{C}'}$ pro →
$[_{\text{R}'}$ al + $\underline{\text{s}}$ $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro

In this case too, the item *als* resulting from the reanalysis is a Head conveying the (functional) meaning of comparison (den Dikken, 2006, p. 172), and therefore RP is better seen as a Deg(ree)P(hrase). The comparative item *als* introduces a Reduced Relative Clause, which in its essence is a CP (see above):

(277) $[_{\text{RP}}$ Spec $[_{\text{R}'}$ Head $[_{\text{YP}}$ Compl]]] = $[_{\text{DegP}}$ Spec $[_{\text{Deg}'}$ als $[_{\text{CP}}$ Compl]]]
Subject Relator Predicate = Subject Comparison Predicate

As depicted in (278), this results in coindexing between the Relator / Head *(al)s* and the Head pro located in Head C, along the same lines of (270): the Relator / Head *-s* prior to reanalysis sits in Spec,CP and is thus coindexed with pro sitting in C via the Spec - Head relation.

(278) $[_{\text{R}'}$ al $[_{\text{CP}}$ $\underline{\text{s}}_i$ $[_{\text{C}'}$ pro_i →
$[_{\text{R}'}$ al + $\underline{\text{s}}_i$ $[_{\text{CP}}$ Op_i $[_{\text{C}'}$ pro_i

This is much more so because *als*, by virtue of its *overt* realization, is the only item coindexed with pro that has the ability to license it: true, Op sits in Spec, CP and therefore potentially licenses the pro occupying C by entering into a Spec - Head relation with it, but de facto it is not able to do so because it is covert as much as pro is.
The just sketched scenario owes much to den Dikken (2006, pp. 177-181), who capitalizes on the (pretheoretical) facts that diachronically *like* and *als* are subject to reinterpretation and word formation and proposes to account these two items in a similar - though not identical - way. He characterizes reanalysis in (pre-)minimalist terms, adopting Baker (1988)'s approach, which takes this phenomenon to be a particular instance of Incorporation (see (230)): ABSTRACT INCORPORATION.
Without entering into details, Baker (1988, pp. 200-204) adds two more details to Langacker (1977)'s definition of reanalysis: the two items involved in it must be

- Heads
- coindexed for some feature

It is also worth observing in this connection that the analytical tool of Abstract Incorporation is compatible with Antisymmetry (see Kayne 1994, pp. 98, 157).
A case in point is the Italian causative contruction illustrated in (279). Here, two verbs behave as a single word, because they cannot be broken up by any intervening material, e.g. another word (Uninterruptability: see Section 4.2), as exemplified in (280) and discussed in Comrie

(1981, Ch. 8). The two verbs at issue are Heads, because a verb (e.g. *(he) jogs*) can replace a more complex expression (e.g. *(he) often jogs in the park*) and are coindexed for their syntactic feature V: they, in fact, belong to the same syntactic category 'verb'.

(279) Italian (Baker, 1988, p. 201)

Maria fa lavorare Giovanni
Mary makes work Giovanni

'Maria makes Giovanni work'

(280) Italian

*Maria fa Giovanni lavorare
Mary makes Giovanni work

'Maria makes Giovanni work'

The core of this process can be summarized along the lines of Baker (1988, p. 202):

Minimalist definition 5.2 (Abstract Incorporation)**.**

(281) $[_{\text{YP}} \ldots Y_i \ldots [_{\text{XP}} X_i \ldots]]$

den Dikken (2006) relies upon Baker (1988) and proposes that the Relator and pro (ie the Head Deg and the Head C) are coindexed for some feature subsequent to reanalysis:

(282) als_i / like_i + pro_i
fa_V + lavorare_V

Or, more accurately:

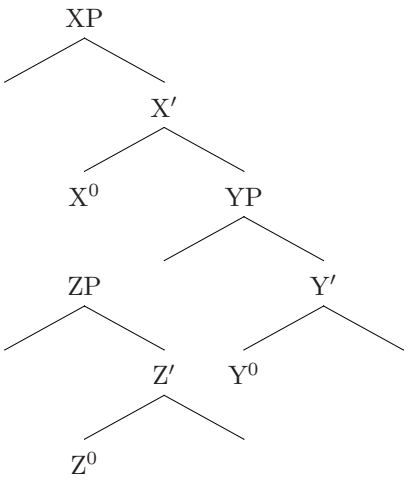
(283) $[_{\text{Deg}'}$ als / like_i ... $[_{\text{C}'}$ pro_i
$[_{\text{X}'}$ fa_V ... $[_{\text{Y}'}$ lavorare_V

In so doing, den Dikken (2006) implements the idea of reanalysis in a way different from that put forward here, where reanalysis and coindexing involve a Head and a Spec. At first sight, den Dikken (2006)'s approach seems to be preferable both on empirical and theoretical grounds: Abstract Incorporation between two Heads is a well documented phenomenon in the (pre-)minimalist literature, and no new explanation is needed to account for the reanalysis undergone by *like / als*. On the contrary, the account provided here complexifies the theory in that it allows an *ad hoc* process of reanalysis and coindexing between a Head and the Spec following it:

(284) a. Den Dikken's proposal: Head_i ... Head_i → cp. Abstract Incorporation
b. Current proposal: Head_i ... Spec_i → cp. ???

Recent research by Julien (2002) on Abstract Incorporation, however, shows that reanalysis between a Head and the Spec following it can be usefully re-thought through the lens of recursion: after all, the syntactic constituent sitting in the Spec of whatever phrase (say, YP) is itself a phrase (say, ZP) and, as such, contains a Head (e.g. Z). It is therefore possible, at least in principle, that the Head Z of the phrase ZP sitting in Spec, YP undergoes Abstract Incorporation with a higher Head X (Julien, 2002, p. 38). This is illustrated in the following:

(285) Abstract Incorporation



The bulk of evidence, according to Julien (2002, p. 38) for this instance of Abstract Incorporation lies in the Italian article (e.g. *lo*), when it is combined with the preposition *di*, the equivalent of English *of*. Italian, in fact, differs from English in that the preposition and the article, when put together, turn into a single word, as exemplified in (286):

(286) Italian (Julien, 2002, p. 38)

a. *di lo studente
of the student
'of the student'

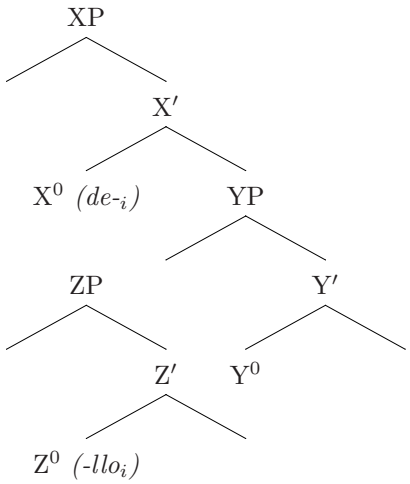
b. dello studente
of.the student
'of the student'

In this regard, Lyons (1999, p. 66ff.) remarks that the article crosslinguistically tends to create a single word if combined with a preceding preposition, noun or adjective[5]. This is not surprising in the light of the fact that all these parts of discourse have a nominal nature on a deeper level of analysis. Taking into account Arabic data, diachronic evidence for identifying the Arabic adjective with a noun has been provided in Ch. 3, while Bardeas (2009) and Ouhalla (2010) lend support to the hypothesis that the Arabic possessive preposition is actually a noun. It could be added that the possessive preposition is not the only one to have a nominal nature. Crosslinguistically it is common for nouns (and verbs) to grammaticalize as prepositions, a state of affairs that 'can either be described by saying that prepositions share certain properties with nouns or by saying that prepositions are a subclass of nouns' (Dryer, 2005c, p. 347).
In addition to the just mentioned parts of discourse, the demonstrative too can be regarded as a noun and the evidence corroborating this claim, once again, is not only from Arabic (see (61) in Ch. 3), but also crosslinguistic: 'there is good evidence that the adnominal demonstratives of many languages are independent pronouns that are only loosely adjoined to the noun in some kind of appositional structure' (Diessel, 2005, p. 175). Moreover, the demonstrative in many

[5]See e.g the examples from Amharic, a Semitic language, reported in (363) and (364) in Section 5.3 below. Notice that in (363) the Amharic article *u* is also found attached to the perfective verb preceding it: this behavior typologically is not unexpected (Lyons, 1999, p. 66ff.) and probably has to do with the fact that the perfective verb itself has a nominal origin in Semitic (Garbini & Durand, 1994, pp. 112-118).

languages (Indo-european included) evolves into the article (Dryer, 2005a)[6], to the effect that the article in the languages at issue can be taken to be a grammaticalized instance of noun.
Under Baker (1988)'s perspective, the combination between a preposition and an article is a natural candidate for Abstract Incorporation because they both are Heads and have in common one feature, namely the syntactic feature N.
These data therefore provide theoretical and empirical support in favor of Julien (2002)'s claim that Abstract Incorporation also affects a given Head and the Head embedded into the Spec immediately below it. Accordingly, the tree representation of (286) is (287):

(287) Abstract Incorporation



Julien (2002)'s account of Abstract Incorporation (reanalysis) can be extended to the Dutch and English data (255): *like / -s* head the phrase occupying the Spec immediately below the Head 0_{ge} / *al*, and *like / -s* and 0_{ge} / *al* are reinterpreted as coindexed (278) and turn into a single word.
Thus, as far as Abstract Incorporation is concerned, the proposal made here (284.b) and den Dikken (2006)'s one (284.a) are equivalent and there is no cogent reason to prefer one over the other. What makes the difference is their behavior relative to *locality*: in (255) the Head (e.g. *-s*) is closer to the Head *al* than pro:

(288) X Z Y
al s pro

Therefore, Minimal Link Condition (see Section 4.5.1) forces reanalysis to take place between the two closest items (X and Z), which rules out den Dikken (2006)'s proposal (284.a) and leads to adopt the alternative account worked out here (284.b).
To summarize the discussion so far, the Canonical Predication structure (255) is in its essence a NP / Subject modified by a Reduced Relative Clause / Predicate. The clause at issue is introduced by a null complementizer (pro), the Head C. It undergoes Abstract Incorporation (reanalysis) as follows:

(289) *Prior* to reanalysis, a given Head embedded into a Spec ('Head-within-Spec') is already into a Spec - Head relation and hence coindexed with the Head immediately *below* it (e.g. $\textit{-s}_i$ + $\textit{pro}_i$ in (255))

[6]Recall from Ch. 3 that this claim does not necessarily hold true for Arabic (see Ch. 3.

(290) If the Head-within-Spec is overt and the Head is covert (pro), the latter get licensed by the former by virtue of the Spec - Head relation (e.g. *-s$_i$ + pro$_i$* in (255))

(291) A given Head and the Head-within-Spec immediately below it can undergo reanalysis (e.g. *al + s > als + Op* in (255))

(292) This results in coindexing between the Head and the Head-within-Spec immediately below it, as well as in the covert realization (Op) of the Head-within-Spec (e.g. *al + s$_i$ > als$_i$ + Op$_i$* in (255))

(293) As a consequence, the covert Spec-within-Head (Op) is no longer able to license the covert Head (pro) immediately below it (e.g. *Op$_i$ + pro$_i$* in (255))

(294) Hence, the covert Head (pro) gets *indirectly* licensed by the covert Head located immediately above the Head-within-Spec (Op) through mediation of this latter, ie by virtue of the CHAIN created by the reanalysis (e.g. *als$_i$ + Op$_i$ + pro$_i$* in (255))

An abstract representation of this process is (295),(296):

(295) $[_{R'}$ X $[_{CP}$ Y$_i$ $[_{C'}$ pro$_i$...

(296) $[_{R'}$ X+Y$_i$ $[_{CP}$ Op$_i$ $[_{C'}$ pro$_i$...
Reanalysis

This account departs from den Dikken (2006)'s original formulation, according to which a *direct* coindexing and licensing between the Relator / Head and the null complementizer of the Predicate takes place: such a kind of licensing and coindexing, in fact, can be hardly deemed correct, because it is not local (288).
Abstract Incorporation (reanalysis) is a *local* operation which not only has the ability to license pro (indirectly, via a chain), but also:

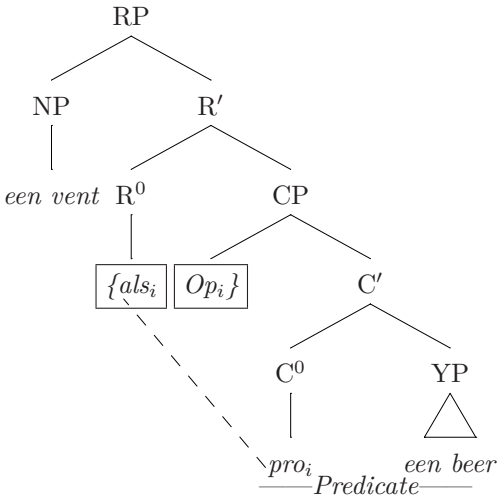
- *blocks* movement of the Predicate
- *g*ives rise to Word Formation ('fusion')

den Dikken (2006, p. 179) captures the descriptive correlation between Abstract Incorporation and movement blocking, but actually this correlation is likely to be more complex, since crosslinguistically it involves also Word Formation (Baker, 1988; Julien, 2002; Hopper & Closs Traugott, 2003).
This issue will be dealt with in greater detail at the end of Section 5.3.2.
Be that as it may, Abstract Incorporation results in the syntactic structure illustrated in (297) and, more abstractly, in (298)[7]:

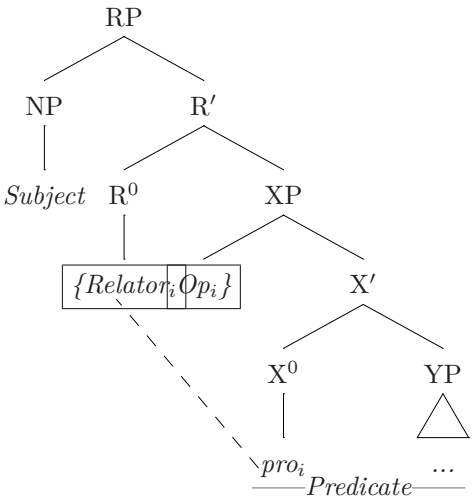
[7] In the wake of den Dikken & Singhapreecha (2004, p. 21, fn. 13) the syntactic tree in the main text presents a simplified scenario, where the reduced Relative Clause (CP) is the Compl of RP. Actually a more accurate antisymmetric representation of the syntactic structure under discussion has it that the Compl of RP is a functional projection (call it XP) and that CP sits in the Spec of this latter (Kayne, 1998; Cinque, 2003; Shlonsky, 2004), along the following lines:

(1) $[_{RP}$ een vent $[_{R'}$ als $[_{XP}$ $\boxed{[_{CP}$ Op $[_{C'}$ pro [een beer]]]} $[_{X'}$ ...

(297) Canonical Predication *cum* Abstract Incorporation (Movement blocked)



(298) Canonical Predication *cum* Abstract Incorporation (Movement blocked)



With all this in place, it is now possible to see *why* presence of coindexing blocks movement (while its lack triggers it). According to Rizzi (1986), coindexing is due to the manifestation of pro in the syntactic structure and might be plausibly explained as a repair - strategy that allows the hearer to understand *the syntactic position* and the *meaning* of a covert element in spite of the fact that it cannot be heard.

In its essence, the repair - strategy of coindexing consists of putting a covert element besides an overt counterpart having the same content, by combining them in Syntax. In this way, what is overt 'rescues' what is covert. Accordingly, not only the covert and overt elements must be coreferential, but also their relationship must be local, because their combination (Merge) for reasons of economy only applies locally (see Section 4.5.1).

'a bloke like a bear'

This scenario is a consequence of the corollaries of Kaynean Antisymmetry discussed at the end of Section 4.5.3. On the one hand, CP cannot be merged to the left of RP and fulfills this position only subsequent to movement. On the other hand, the only landing site available for movement of a whole phrase (e.g. CP) is a Specifier. To keep the syntactic structure transparent, however, the simplified scenario shown in the main text is used. This does not affect the point made here.

Such a strategy has been first observed in Italian by Rizzi (1986), which singles out the relationship holding between a pro acting as Subject and its verb in Italian, where the proper recognition of pro rests on two conditions :

- pro is associated with a verb carrying its same information about number and gender but explicitly, which makes it easier to correctly retrieve the information that pro conveys covertly (CONTENT-LICENSING)
- the information-retrieval is successful if the relationship is local (ie if pro is close to the verb) (FORMAL LICENSING)

This can be easily seen considering the Italian equivalent of *he runs* in fn. 4:

(299) ... $[_{\text{IP}}$ pro_{SG} $[_{\text{I}'}$ (corr)-e_{SG}...

In (255) exactly the same relationship holds, given that before the reanalysis the covert pro sitting in the Head C is adjacent to (= *local* to) Spec, CP occupied by a *coreferential* overt element (*like / -s*):

(300) ... $[_{\text{CP}}$ like (of which)$_i$ $[_{\text{C}'}$ pro_i...

Technically, the (null Head) pro is (content- and formally) licensed.
On the other hand, if in a given local configuration there is no overt element coreferential with the covert pro, the latter, according to den Dikken (2006), seeks another local configuration where an instance of the former is found and does so by moving to a new portion of the syntactic structure, as it will be seen in the next Section.

5.2.2 The Linker

The antisymmetric framework imposes that the covert pro searches the local configuration featuring its overt coreferential counterpart higher in the tree, because no syntactic structure can be merged in its lower portion. According to den Dikken (2006), the first local configuration available is some functional projection immediately above the Canonical Predication Structure ('RP'). Recall from the previous Section that RP has a NP / Subject as its Specifier and a CP, e.g. a Reduced Relative Clause as its Complement.
The projection immediately above RP (the so-called 'FP': see previous Section) is headed by a particle-like element (the Linker) that, whatever its exact nature, encodes some information also brought by pro. This is why pro raises to it.
Such a state of affairs can be exemplified by the Canonical Predication structure underlying the Inverted Predication *een beer van een vent / a bear of a bloke* (256). After merging of the Linker *van*, it can be represented as follows:

(301) ... $[_{\text{FP}}$ $[_{\text{F}'}$ Linker: van_i $[_{\text{RP}}$ een vent $[_{\text{R}'}$ Relator $[_{\text{RC}}$ Op pro_i een beer...

Semantically, (256) is equivalent to (255), since the former is commutable with the latter (den Dikken, 2006, p. 175) and therefore the RP of (256), depicted in (301) can be identified with DegP.
But syntactically the Linker (say, *van* of (301)), albeit coreferential with pro (*content*-licensing) is not able to *formally* license pro because their relationship is not local: this happens because

the NP / Subject *een vent* intervenes between them. As a first approximation, one may assume that to get around this problem, pro can move to the left of the Linker, ie to Spec, FP: this movement, in fact, would give rise to a Spec - Head relation between pro in Spec, FP and the Linker *van* in the Head F.

In other words, den Dikken (2006) adopts the Feature Checking Hypothesis (see Section 4.5.3), by virtue of which movement is driven by the fact that there are two identical features in the syntactic structure that need to be close each other for some reason. In this case, the two features need to be close each other in order to make a covert feature recognizeable via its overt counterpart.

In this perspective, the higher Head F *van* is the ATTRACTOR, while the lower pro to be raised, sitting in Spec, CP is the ATTRACTEE (see Section 4.5.1).

This idea, however, needs refinement. First, because Head-to-Spec movement of pro is not allowed in an antisymmetric framework (Section 4.5.3). Second, because the Minimal Link Condition (Section 4.5.1) forces the Head F *van* to attract to Spec,FP the constituent closest to it, which in this case is the NP *een vent*, *not* pro:

(302) ... $[_{FP}$ $[_{F'}$ Linker: van$_i$ $[_{RP}$ een vent $[_{R'}$ Relator $[_{RC}$ Op$_i$ een beer...

(303) Head / Linker: van$_i$ een vent ...pro ...
YES *NO*

As for the former problem, it can be easily solved by assuming that pro raises as a part of a larger phrase, namely the entire RC: Compl-to-Spec movement, in fact, is perfectly legitimate within an antisymmetric framework (see Section 4.5.3).

The latter problem, according to Chomsky (1995, p. 185) (and den Dikken 2006, pp. 115ff.) evaporates as soon as one takes into account the Attractor of pro, which is the Head R immediately above it:

(304) [Spec F [een vent R [...pro...]]]
ATTRACTOR #1 ATTRACTEE #1 ATTRACTOR #2 ATTRACTEE #2

The Head R, in fact, can move to the Head F, where it attaches to its left (see the Mirror Principle in Section 4.5.3). Now, the Head F is no longer obliged by the Minimal Link Condition to attract *een vent* to Spec, FP, because the Head F has become a more complex item, which contains also the Head R and this latter has still the ability to attract pro. To put it differently, Head movement has the ability to 'bypass' the Minimal Link Condition:

(305) [Spec R+F [een vent t [...pro...]]]
ATTR.OR #2 + ATTR.OR #1 ATTR.EE #1 ATTR.EE #2

This makes it possible to raise pro in order to render it local to the Linker *van* and therefore to formally license pro:

(306) [pro$_i$ R+F$_i$ [een vent t t]]
ATTR.EE #2 ATTR.OR #2 + ATTR.OR #1 ATTR.EE #1

In den Dikken (2006, pp. 98, 113)'s framework, this amounts to saying that Predicate Inversion

- is 'triggered by the need to license an empty predicate head'[8]

[8]Den Dikken takes this explanation 'to be sufficient discouragement for an attempt at deriving the rationale for Locative Inversion (and Predicate Inversion more generally) from information-structural notions such as focus and topic.' (den Dikken, 2006, p. 90). This issue will not be addressed here.

- requires the Relator (a Head) to move to the Linker (a Head)

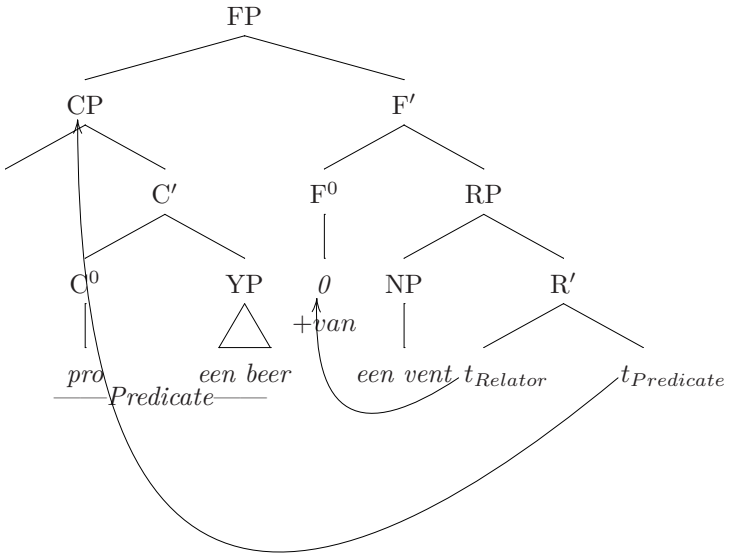
Technically, the just described interplay between Head Movement and Minimal Link Condition is stated by Chomsky (1995, p. 185) (and by den Dikken 2006 after him) as follows:

Minimalist definition 5.3 (Domain Extending Head Movement)**.**

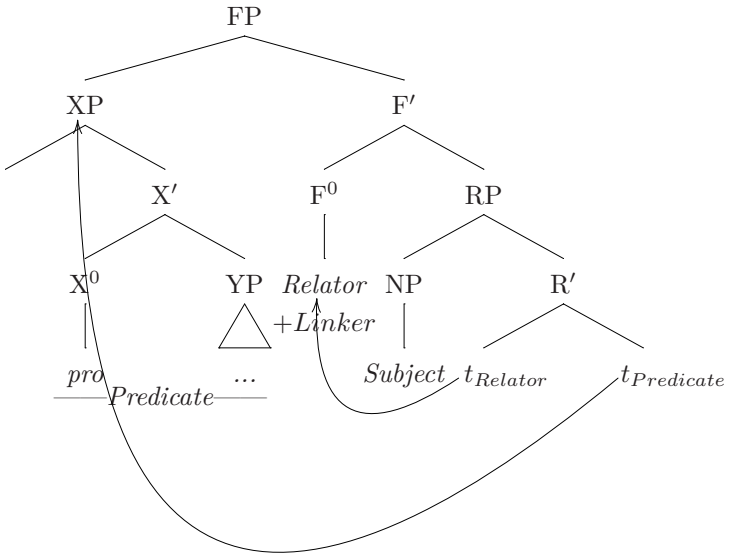
If Y adjoins [= head-moves] to X, forming the chain (Y, t) with the minimal domain Spec1, Spec2, ZP, then Spec1 and Spec2 are equidistant from ZP (or anything it contains), so that raising of (or from) ZP can cross Spec2 to Spec1.

On the whole, an account of Predicate Inversion à la Den Dikken assigns (256) the representation in (307) or, more abstractly, that in (308)[9]:

(307) Predicate Inversion (= 'scrambled' Canonical Predication)



(308) Predicate Inversion (= 'scrambled' Canonical Predication)



In particular, the parallelism between the cases of pro-licensing discussed by Rizzi (1986) and den Dikken (2006)'s account of Predicate Inversion is illustrated in (309)

[9]See fn. 7 on the departure point of this derivation (Canonical Predication).

(309) [Spec: pro$_i$ [Head: Relator + van$_i$]
[Spec: pro$_i$ [Head: corr + e$_i$]

However, as Tanase Dogaru (2008, p. 302) remarks, accounting Predicate Inversion in terms of pro-licensing is an idea that den Dikken (2006) implements more through indirect evidence from primary predication than through evidence from the NP per se.
Thus, it would be desirable to extend this analysis from the Dutch / English NP to a broader range of languages and / or functional projections.
Descriptively, the discussion so far has demonstrated that pro-licensing has to do with a Relator conveying the meaning of comparison (Deg) and with a modifier / Predicate consisting of a Reduced Relative Clause. In the best case it is thus expected that crosslinguistically the typology of Canonical Predication and Inverted Predication within the nominal domain features this kind of Relator and / or this kind of modifier / Predicate.

5.2.3 Predication and non-restrictiveness

The Burmese language appears to be a good starting point to corroborate the analysis outlined in the previous Section, since it instantiates reduced relative clauses which obligatorily undergo Predicate Inversion - that is, they are always *prenominal* relative clauses. (310) is illustrative of a Burmese Relative Clause (RC henceforth):

(310) Burmese (den Dikken & Singhapreecha, 2004, fn. 23, p. 36)

ca-naw weh thii sa-ouq
I buy *thii* book

'the book I bought'

Bearing in mind that under the current approach a reduced relative clause is characterized by a pro complementizer, the Burmese Reduced RC can be interpreted along the following lines:

(311) Burmese (den Dikken & Singhapreecha, 2004, fn. 23, p. 36)

[0 ca-naw weh]^{RC} 0^{RELATOR} thii sa-ouq t^{RELATOR} t^{RC}
pro I buy RELATOR LINKER book

Remarkably, the correlation found in the Burmese RC between prenominal position and lack of overt C is not language-specific, being rather a pattern widespread among world languages belonging to different families, as pointed out by Kern (2010, p. 16): Indo-european, Hamitic, Dravidian, Ural Altaic, Native American.
Kern (2010, p. 1), in fact, observes that 'in most RCpre constructions [= prenominal RCs], there is often no relativizer [= relative complementizer] at all'.
Kern (2010, p. 14) also remarks that in cases like Burmese the particle intervening between the (Inverted) Predicate / modifier and the Subject NP (e.g. *thii*) cannot be regarded as a complementizer, because this latter is also used to mark a sentential object of verbs of thinking and saying (cp. English *I say <u>that</u>...* and Section 5.2.1), which is not the case of *thii*.
Kern (2010)'s distributional considerations can be (re)formulated as follows:

Syntactic Diagnostics 5.1 (Marker of Predicate Inversion)**.**

A given functional item is a marker of Predicate Inversion iff and only if it can intervene between a Predicate and its Subject (in this order) and it does not introduce a sentential object of a verb of saying and thinking, a function typical of C.

Finally, Kern (2010) brings syntactic and crosslinguistic evidence that the prenominal RCs do not necessarily have a RESTRICTIVE reading[10], contrary to what it is usually assumed (cp. Den Dikken's English translation of (310)) and that their non-restrictive reading is also possible. In this light, (310) can be also rendered in English as '*a book, which I bought*' (provided that an appropriate context is found).

The discussion so far thus lends good support in favor of analyzing the prenominal position of the RCs as related with (actually, triggered by) the covert realization of their complementizers. Recall that traditionally 'covert realization of their complementizers' is designated as Reduced RC, while in minimalist terms it is the fact that C is pro (see Section 5.2.1).

Further evidence for this analysis comes from a complementary pattern, widely attested crosslinguistically. While, as just noticed, RCs whose C is pro ('reduced RCs') tend to precede the noun they refer to, RCs whose C is overt tend to follow it. Actually, not only RCs but *all* the subordinate clauses introduced by an overt C behave this way, as Cinque (2005, p. 53) concludes after examining a sample of 150 languages belonging to the same language families studied by Kern (2010).

Moreover, Cinque (2005, p. 53) defines the complementizer in the same terms of Kern (2010) as a particle having the ability to introduce both a RC and a sentential object of verbs of saying and thinking: 'In turn, the possibility for a clause to follow the V or the N seems to some extent related to the presence of initial complementizers'. This results in the typological scenario summarized in (312) and in Fig. 5.1.

(312) covert C (pro) ('Reduced RC') → RC + Noun (Kern, 2010)
overt C ('Full RC') → Noun + RC (Cinque, 2005)

The correlation between a *reduced* prenominal RC and a covert complementizer stated by Kern (2010) and that between a postnominal *full* RC and an overt complementizer stated by Cinque (2005) is thus derived in terms of Rizzi (1986)'s content-licensing of pro and its refinement worked out by den Dikken (2006): it is the need of content-licensing a covert C / pro that

[10] A NON-RESTRICTIVE RC is a kind of RC that adds non-necessary information about an already identified (and hence definite or specific) noun. It can be easily recognized in English by the presence of an intonational break (or, graphically, by a comma) which sets it off from the noun it modifies and it refers to. This latter is traditionally designated as ANTECEDENT. The (English) non-restrictive RC is exemplified in the following:

(1) English (Comrie, 1981, p. 138)
The man, who had arrived yesterday, left this morning

This intonational property is not found in the RESTRICTIVE RC, which differs from the non-restrictive RC also in that it serves to identify its antecedent and therefore conveys necessary information about it:

(2) English (Comrie, 1981, p. 138)
The man that I saw yesterday left this morning

Both the types of RC are clauses and as such are introduced by a CP (see Section 5.2.1), so that a complementizer can occur at their beginning. In the (pre-)minimalist literature, lack of a complementizer is taken to hold on A-P only, its semantic counterpart being at work in the syntactic structure and at C-I. Accordingly, in MP a complementizer always introduces a clause, even if it sometimes is not audible.

Both the types of RC must contain a pronoun coreferential with the antecedent, or ANAPHORIC PRONOUN, which is taken in the (pre-)minimalist literature to be located in Spec, CP and to have sometimes a covert realization. In this case, it is notationed as Op (see Section 5.2.1). This state of affairs sometimes creates a potential ambiguity in the interpretation of the item introducing the RC: is it a pronoun or a complementizer? If the matter is not clear (or not relevant), the more generic definition RELATIVE MARKER is adopted.

Table 5.1: Syntactic position of RCs and realization of their complementizer

Position	Type	Complementizer	Predicate	Source
Prenominal	Reduced	covert (pro)	Inverted	Kern (2010)
Postnominal	Full	overt	Canonical	Cinque (2005)

triggers Predicate Inversion of the Burmese (etc...) RC.
In this connection, it is also worth noticing that Greenberg (1990b, p. 181, 185, 191) as well as Greenberg (1990a, p. 238) claims that the complex:

(313) Numeral + Classifier

that in the so-called Numeral Classifier languages is a modifier / Predicate typically combined with a Subject NP in order to express number (see Ch. 1) is to be interpreted as an apposition, a construction that is actually the reduced counterpart of a full non-restrictive RC.
The contrast between (314a) and (314b) makes this correspondence adamantly clear:

(314) English (Heringa, 2007, pp. 76-77)

a. His girl friend, who is a modest person, laughs about that

b. His girl friend, a modest person, laughs about that

In this light, Numeral Classifier constructions are a subtype of Inverted Predication, whose Inverted Predicate is a (reduced and non-restrictive) RC - a phenomenon that den Dikken & Singhapreecha (2004) and den Dikken (2006) do not take into consideration [11].
Evidence for Greenberg (1990a)'s claim is mainly typological and relies upon data from Japanese and Gilbertese (an Oceanian language, also called Kiribati). Interestingly, the numeral classifier constructions he mentions display the order:

(315) Modifier / Predicate + Subject NP

Furthermore, a particle can intervene between the modifier / Predicate and the Subject NP (in the descriptive and typological literature it is called LIGATURE, ASSOCIATIVE PARTICLE etc...):

(316) [Numeral + Classifier]$^{Modifier/\ Predicate}$ + (Particle) + Subject NP

To make the discussion more concrete:

(317) Japanese (Greenberg, 1990a, p. 237)

sam-biki no inu
three-CLF *no* dog

'three dogs'

(318) Gilbertese (Trussell, 1979, p. 111)

[11] As Tanase Dogaru (2008, p. 302) points out, den Dikken & Singhapreecha (2004)'s and den Dikken (2006)'s account of Numeral Classifier constructions is not satisfactory also from another point of view: they do not single out any null Head to explain the Predicate Inversion taking place in such constructions. Accordingly, there is developed here an alternative account, which however is consistent with den Dikken (2006)'s framework. In fact, it is meant to provide evidence for the existence of a null Head triggering Predicate Inversion, whose lack is precisely the object of Tanase Dogaru (2008)'s criticism.

tengaun ma iti-man aomata
ten and seven-CLF people

'seventeen people'

Greenberg (1990a) provides empirical support for identifying the complex Numeral plus Classifier with a (non-restrictive) RC relying upon the following contruction, found in Gilbertese:

(319) Gilbertese (Trussell, 1979, p. 111)

nam'akaina aika teni-ua
month RELATIVE C three-CLF

'three months'

(319) is in fact a construction commutable with (318) and the particle *aika* in Gilbertese generally functions as a relative complementizer, given that it introduces RCs and it is accompanied by an anaphoric pronoun (Trussell, 1979, p. 182). In (319) this latter is covert (see immediately below for more details).
On these grounds, (317), (318) might be plausibly regarded as Predicate Inversion structures and their inverted predicate as a (reduced and non-restrictive) RC. As for the particle that in Japanese (317) intervenes between the modifier / predicate and the Subject NP, it can be taken to be a Linker after den Dikken & Singhapreecha (2004, p. 42), and therefore, by this line of reasoning, its Gilbertese counterpart (318) features a covert Linker in the same position. This state of affairs is illustrated as follows:

(320) Japanese (Greenberg, 1990a, p. 237)

[0 sam-biki]^{RC} 0^{RELATOR} no^{LINKER} inu t^{RELATOR} t^{RC}
pro three-CLF RELATOR LINKER dog

(321) Gilbertese (Trussell, 1979, p. 111)

[0 tengaun ma iti-man]^{RC} 0^{RELATOR} 0^{LINKER} aomata t^{RELATOR} t^{RC}
pro ten and seven-CLF RELATOR LINKER people

On the other hand, Gilbertese (319) is an instance of Canonical Predication, where in the Head C no pro occurs (its overt counterpart *aika* being found instead) and hence there is no need to trigger Predicate Inversion:

(322) Gilbertese (Trussell, 1979, p. 111)

nam'akaina 0^{RELATOR} [aika teni-ua]^{RC}
month RELATOR RELATIVE C three-CLF

English, however, appears to raise counterexamples to the analysis carried out so far (summarized in (312) and in Fig. 5.1).
To begin with, English restrictive RC can be *reduced* and *postnominal* (e.g. *the man I saw yesterday*), contrary to the generalization made here (cp. Fig. 5.1). On one side, in fact, it instantiates a kind of Canonical Predication where the Head C of the modifier / predicate is occupied by pro, which yields a reduced relative clause. On the other side, Predicate Inversion does not occur, contrary to the expectation.

The counterexample is only apparent, if one considers the following piece of data: pro is commutable with *that* (*the book that I bought*). This (pretheoretical) fact can be interpreted as a way of content-licensing pro of the English reduced relative clause not on the syntagmatic axis (Spec, CP) but on the paradigmatic axis - that is by virtue of Saussure (1916)'s *rapport associatif*. Rizzi (1986) admits this strategy (which he dubs THEORY OF ARB, Arb for Arbitrary Interpretation) as an alternative way to content-licensing pro, in order to account for the following Italian sentences:

(323) Italian (Rizzi, 1986, p. 503)

Il bel tempo invoglia a restare
The nice weather induces pro to stay

'The nice weather induces people to stay'

In (323) the verb *invogliare / to induce* obligatorily takes an object, but this latter is found only in English (*people*). Then, in Italian a covert counterpart of *people* must occur for semantic reasons, which forces to posit an Object pro in the syntactic structure. But this instance of pro, upon closer scrutiny, is not content-licensed, because there is no coreferential overt element close to it. In this case, the content of the Object pro, according to Rizzi (1986, p. 523), has been already licensed in the Lexicon, prior to its merging in Syntax: pro is associated with a particular instance of the pronoun *loro* 'they'.
This is the pronoun *loro* having a generic reading (cp. English *people*), which in this usage is only M:

(324) Italian

loro sono tutti ladri
they are all.M thieves.M

'The people are all thieves'

In fact, the F counterpart of (324) is fine, but loses its generic meaning:

(325) Italian

loro sono tutte ladre
they are all.F thieves.F

'They (F) are all thieves'

Rizzi (1986) claims that the Object pro in (323) is content-licensed via (paradigmatic) association with the generic *loro* exemplified in (324) because, among others, a modifier referring to an Object pro can be agreed with it only for the M PL:

(326) Italian (Rizzi, 1986, p. 505)

Un dottore serio visita nudi
A doctor serious visits pro nude.M

'A serious doctor visits people nude'

On these grounds, pro of the English reduced RC is content licensed via paradigmatic association with its overt counterpart *that* very much as Italian Object pro is content licensed via paradigmatic association with its overt counterpart *loro*.
It is worth remarking in this regard that the covert C (pro) of the Gilbertese Numeral Classifier construction realized as *reduced* RC (318) cannot invoke the same strategy because it is not able to get paradigmatically associated with its overt counterpart *aika*. This is due to the fact that the Gilbertese Numeral Classifier construction realized as a *full* RC (319) drops the Subject anaphoric pronoun and *aika* takes over also the function of this latter (Trussell, 1979, pp. 182-183).
Within the present framework, this means that *aika* can be ambiguously interpreted either as a Subject pronoun sitting in Spec, CP or as a Head C (see Section 5.2.1). It is exactly this ambiguity that prevents the covert C (pro) to identify *aika* as its overt counterpart and then to get content-licensed by this latter. Accordingly, Predicate Inversion takes place to content-license pro, which yields (318) in compliance with the generalization summarized in (312) and in Fig. 5.1.
English non-restrictive RC poses the same problem of its restrictive counterpart: it can be *reduced* and *postnominal* (see 314b), contrary to the generalization made in Fig. 5.1. In this case too, the apparent anomaly is straightforwardly explained in terms of pro-licensing via a Saussurean *rapport associatif*. Crosslinguistic data evidence that semantically and diachronically, RCs are closely related with adverbial clauses (Heine & Kuteva, 2007, 251ff.) and this especially holds true for non-restrictive RCs, which can also function as adverbial clauses (see Frascarelli & Puglielli 2006 and the Somali data reported in (329) and in fn. 12 below).
In this light, it is expected that the covert complementizer (pro) of the English reduced non-restrictive RC be licensed by an overt complementizer introducing an adverbial clause.
In the vein of Abney (1987) and Kaiser (1999), it is possible to identify this kind of complementizer with the suffix *-ing* found in the so-called ABSOLUTE CONSTRUCTION. This is exemplified by a partial paraphrase of the full non-restrictive RC (314a), namely *His girl friend, being a modest person, laughs about that*. The gist of Abney's and Kaiser's reasoning is that the absolute construction works as a sentence and therefore it must be introduced by a CP headed by a functional item, which they identify with *-ing*. It ensues that the absolute construction is assigned a syntactic representation along the following lines:

(327) His girl friend $[_{CP}$ Op $[_{C'}$ be+ing ...

A direct implication of (327) for the present study is that in English the overt C *-ing* of the absolute construction licenses pro of the reduced non-restrictive RC. More generally, English data are a motivated exception to the working hypothesis entertained here (the generalization illustrated in (312) and Fig. 5.1) and only an 'Anglo-centric' approach to the RC prevents the analysis from capturing this generalization.
Having established that the Burmese, Gilbertese and Japanese noun enters into a predication relationship with a reduced RC in compliace with the above mentioned generalization, the question arises of what is the exact nature of the Relator mediating such a relationship. The answer plausibly lies in a common feature shared by all the prenominal RCs: they have non-restrictive meaning.
Such a semantic characterization of the prenominal RCs is weak for the Burmese RC, since it

can also convey restrictive meaning, (cp. Den Dikken's translation of (310)) but is strong for Gilbertese and Japanese. Recall in fact that Greenberg (1990b,a) brings evidence in favor of the fact that the complex Numeral plus Classifier diachronically develops out of a non-restrictive ('appositional') RC.

Thus, whatever the syntactic representation assigned to the semantic feature 'non-restrictive', it seems safe to maintain that it holds for the Burmese RC only under certain interpretive circumstances, while it is always true for the complex Numeral plus Classifier. With this caveat, it seems that the semantic feature 'non-restrictive' is represented in Syntax as a conjunction. This hypothesis goes back at least to the linguistic and philosophical speculation of the school of Port-Royal, where it is theoretically evidenced in terms of propositional logic (see Chomsky 1966 and references therein).

Port-Royal's position has been revived in modern linguistics, where both distributional and typological evidence has been culled to support it empirically. First, in Indo-european languages (e.g. English) non-restrictive RCs are commutable with conjoined clauses:

(328) English (Ross, 1967, p. 435)

a. Enrico, and he is the smartest of us all, got the answer in seven seconds

b. Enrico, who is the smartest of us all, got the answer in seven seconds

Second, in non-Indoeuropean languages the meaning of non-restrictive RC is conveyed by a conjoined clause *tout court*. For example, in Somali, a Hamitic language, the conjunction *oo* 'and' introduces a non-restrictive RC:

(329) Somali (Frascarelli & Puglielli, 2006, p. 310)

Cali	oo	Maryam	la	hadlayá	waa	walaalkay
Cali	and	Maryam	with	speaking.is	FOCUS MARKER	brother-my

'Cali, who is talking to Maryam, is my brother'

These data lead Ross (1967) and Frascarelli & Puglielli (2006) to assign the non-restrictive RC the following representation (usually referred to as MATCHING ANALYSIS: see Leu 2008 for more details and references):

(330) $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ and, oo ... $[_{\text{YP}}$ RC]]]

(328,329) seem therefore to be evidence enough that the Relator intervening between the Subject NP and the non-restrictive reduced RC in Burmese, Japanese and Gilbertese (310,317,318) is to be identified with a (covert) Conjunction, heading a Conj(unction)P(hrase) and whose Compl is a RC (ie a CP):

(331)

$[_{\text{RP}}$ Spec	$[_{\text{R}'}$ Head	$[_{\text{YP}}$ Compl]]]	=	$[_{\text{ConjP}}$ Spec	$[_{\text{Conj}'}$ 0	$[_{\text{CP}}$ Compl]]]
Subject	Relator	Predicate	=	Subject	Conjunction	Predicate

(331) is an interpretation of the non-restrictive RC originally put forward by Ross (1967) for English and by Frascarelli & Puglielli (2006) for Somali and extended here to the just discussed instances of RP[12].

[12]Under Frascarelli & Puglielli (2006)'s approach the Complement RC is introduced by a silent CP, where the anaphoric pronoun sitting in Spec CP is Op and the Head C is pro. This approach converges with the

The picture that emerges from the discussion so far is that a CP headed by pro is a good candidate for movement and Predicate Inversion. The CP at issue synchronically is the Complement either of a DegP (272) or of a ConjP (331) - in more traditional terms, it is the predicate either of a comparative or a non-restrictive relationship of predication.
In greater detail, the non-restrictive RC, if reduced, is typically prenominal, ie an Inverted Predicate. The counterexamples found in English are only apparent and can be explained resorting to the Saussurean notion of *rapport associatif*. Alternatively, counterexamples to this pattern involve Abstract Incorporation (reanalysis), as evidenced in the previous Section through Dutch and English examples.
On the other hand, looking at ConjP (331) and DegP (272) under a diachronic perspective reveals that the latter derives from the former. The bulk of evidence for such an assumption, discussed in Section 5.2.1 is that the Inverted Predication (256) is commutable with the Canonical Predication featuring *like* (255), which in some relic constructions still has a clear non-restrictive meaning (*the like of which*: cp. (263), (267a)).
There is no principled reason, however, to assume that DegP (comparison) must be *always* etymologically tied with ConjP (non-restrictiveness) and that it has not a semantic status of its own.
Strong empirical support for allowing instances of DegP (comparison) independent on ConjP (non-restrictiveness) is the fact that in some cases the DegP having undergone Predicate Inversion is *not* commutable with a Canonical Predication counterpart featuring *like*. This happens when the inverted predicate / modifier of DegP is associated either with the adverb *too* or the demonstrative *that* (den Dikken, 2006, p. 177ff.)[13]:

(332) English (based on den Dikken 2006, pp. 162-164)
A jewel of a village / a village like a jewel
That idiot of a doctor / *a doctor like that idiot

The inability of *that idiot of a doctor* in (332) to be commutable with a Canonical Predication structure featuring *like* has a non-trivial implication: given that *like* is responsible for the relative clause reading of the Inverted Predicate *that idiot* (see Fig. 5.1), its absence in the

reduced RC analysis proposed here for the Burmese, Japanese, Gilbertese RC. It would be then expected that, all the other things equal, the covert pro of the Somali non-restrictive RC triggers Predicate Inversion but clearly this is not the case in Somali, which displays Canonical Predication (see (329)). This apparent exception is straightforwardly explained along the lines of the Gilbertese non-restrictive RC (318) by invoking the paradigmatic axis. The Somali conjunction *oo*, in fact, can also convey a temporal meaning:

(1) Somali (Frascarelli & Puglielli, 2006, p. 321)

Isagak	oo	isbitaal-ka	ku jira	buu		dhintay
he	and	hospital-ART	in stayed	FOCUS MARKER.he		died

'He died while he was in the hospital'

In this case, *oo* performs the function of the temporal complementizer *while*, which is usually assumed to sit in Spec, CP (Haegeman, 1991, Ch. 3) and in Somali is realized as *inti*: consequently, *oo* can be ambiguously interpreted either as a Head (Conj of (330)) or as a temporal complementizer sitting in Spec, CP (= *while*). It is exactly this ambiguity that enables the covert C (pro) to identify *oo* as its overt counterpart and then to get content-licensed by this latter: *oo*, when located in Spec, CP not only shares with the Head pro the syntactic feature 'complementizer' but also enters with the Head pro into a Spec - Head relationship (see Section 5.2.1).

[13] Both the constructions under discussion are DegP because the inverted predicate / modifier has the semantic property of SCALARITY, as it will be explained shortly.

Canonical Predication structure of *that idiot of a doctor* (**a doctor like that idiot*) makes highly questionable interpreting its Inverted Predicate as a RC.
The only alternative left is taking the Inverted Predicate *that idiot* to be a NP and consequently the pro responsible for Predicate Inversion, if any, is to be sought within this domain.
One may even push this line of reasoning to its extreme limits and claim that DegP in its most typical form is adorned with *too, that* and features a modifier / Predicate which is *not* an RC. By the same token, one may also speculate that ConjP has a complementary behavior: it features a modifier which is an RC but not associated with *too, that*.
In this light, the manifestations of DegP whose modifier is an RC not associated with *too, that* (255, 256, 272) discussed in Section 5.2.1 are hybrid forms: hence, a ConjP pattern would have been attracted into a DegP pattern, thus yielding the hybrid forms at issue. This is precisely what Abstract Incorporation (reanalysis) does, as discussed at length in the previous Sections. Hence, the overall typology depicted in Fig 5.2 obtains. This hypothesis is further evidenced

Table 5.2: NP: typology of RP

Semantics	Relator	Realization	Predicate	Example
Non-restrictiveness	Conjunction	and, oo	RC	Enrico, and he is the smartest...
Comparison	Degree	-	NP	that idiot of a doctor
Both	Degree	like	RC	a bloke like a bear

in the next Section.

5.2.4 Predication and comparison

To obtain a clearer understanding of why and how in the construction exemplified by *that idiot of a doctor* the pro triggering Predicate Inversion manifests itself within the NP, rather than within an RC, it seems convenient to look more closely at the syntactic profile of the NP in general.
In MP the NP *qua* lexical projection, can be expanded with some functional projections, merged on its left under an antisymmetric approach (see Section 4.5.3). The topmost projection of NP is DP, which adds to NP an additional piece of (functional) information, namely definiteness. Because of its being the topmost projection of NP, DP since Abney (1987) has been brought into comparison with CP.
Among the possible phonological realizations of DP is the article: this latter, in fact, is crosslinguistically found in a NP-initial position (Lyons, 1999, p. 64). Giusti (2002) adopts a cartographic / antisymmetric approach and provides support in favor of the claim that the NP-initial article is the Head D of a functional projection DP higher than NP, to the effect that the morphemic material preceding the article (e.g. demonstratives) is taken to be hosted in Spec, DP[14]:

(333) Modern Greek (Alexiadou et al., 2007, p. 122)

a. afto to vivlio
this the book
'this book'

[14]The demonstrative sits in Spec, DP subsequent to movement: see Campbell (1996) and Giusti (2002).

b. [$_{\text{DP}}$ afto [$_{\text{D}'}$ to [vivlio

Another property that DP has in common with CP is that of being a complete chunk of semantico-syntactic information, a property referred to in the minimalist literature as PHASE (see e.g. Chomsky 2008 and references therein). In fact, distributionally an embedded clause (CP) has the ability to replace a DP: contrast *John's proof of the theorem...* with *that John proved the theorem...*

This similarity leads one to think that Spec, DP and the Head D are involved in the same mechanism of coindexing holding between Spec, CP and the Head C (see Section 5.2.1), which runs into the same problems if the head D is pro. Giusti (2002)'s account demonstrates that in the standard case coindexing applies, because Spec, DP (say *afto*) and the Head D (*to* of (333)) are identical in their agreement features:

(334) [$_{\text{DP}}$ afto$_{SG}$ [$_{\text{D}'}$ to$_{SG}$ [vivlio

According to Giusti (2002), this holds true also when either Spec, DP or the Head D are not phonologically realized, like in English:

(335) [$_{\text{DP}}$ this$_i$ [$_{\text{D}'}$ pro$_i$ [book

(336) [$_{\text{DP}}$ Op$_i$ [$_{\text{D}'}$ the$_i$ [book

Notice that in (335) the covert Head D is identified with pro because it has as its overt counterpart *the*, which is clearly of pronominal origin. In (336) the covert Spec, DP is identified with the same covert Op found in Spec, CP of the RC because the former plays the same semantico-syntactic role of the latter: *afto, this* and *Op* in terms of propositional logic are *all* Subjects of the lower NP (see Section 5.1).

Finally, Giusti (2002) takes NPs not adorned with articles or demonstratives to occupy Spec, DP. A mass-noun like *water* in and by itself does not convey definiteness nor indefiniteness and in fact can occur in both the contexts. Giusti (2002) therefore assumes that *water* raises to Spec, DP where it gets definiteness via coindexing with the overt pro:

(337) [$_{\text{DP}}$ water$_{DEF}$ [$_{\text{D}'}$ pro$_{DEF}$ [t

In this light, the (first) NP of the DegP exemplified in (332), namely *that idiot*, is better understood as a DP, so that these instances of Inverted Predication have typically the form *that / too DP of DP*.

Focusing on the first DP / Inverted Predicate, it is invariably SCALAR (Matushansky, 2002, p. 274), that is denotes an increase or decrease of some amount.

If the increased / decreased amount is physical, scalarity is conveyed by including in the first DP a particle called INTENSIFIER Matushansky (2002, p. 232 and references therein): in particular, in the Predicate Inversion *DP of DP* it is realized as *too*:

(338) English (Troseth, 2009, p. 38)
(Attila is) too good of an athlete

If the increased / decreased amount is metaphorical, scalarity is intrinsic in the DP itself, since this latter must have a meliorative / pejorative meaning and in fact in the latter case is often used also as insult. English *bastard* is a case in point (Matushansky, 2002, p. 233 and references therein). Similar DPs also occur as the first DP of the Predicate Inversion *that DP of DP*

(339) English (den Dikken, 2006, p. 164)
That idiot of a doctor just wrecked my car

den Dikken (2006, Ch. 5) and Troseth (2009) after him interpret *of* of this construction as a Head F / Linker that attracts the first DP *too good, that idiot*, base-generated as the Complement of RP, along the same lines of (304, 306):

(340) [Spec of$_i$ [an athlete / a doctor R [too good$_i$ / that idiot$_i$]]

(341) [too good$_i$ / that idiot$_i$ R+of$_i$ [an athlete / a doctor]] ...

Remarkably, if one espouses Giusti (2002)'s view, both the instances of scalar Inverted Predication *DP of DP* exemplified in (338) and (339) feature a pro in the Head D[15]:

(342) [$_{\text{DP}}$ that [$_{\text{D}'}$ pro [idiot

(343) [$_{\text{DP}}$ too good [$_{\text{D}'}$ pro [t

(342), (343), in fact, are no more than (335), (337), respectively, merged as first DP of the scalar Inverted Predication *DP of DP*.
At this point there arises a problem involving definiteness. On one side, the Spec-Head relation imposes that in *that (pro) idiot, too good (pro t)* Spec, DP and the Head D be coindexed for their features, among them (in)definiteness (see (337)). On the other side, den Dikken (2006) points out that *that (pro) idiot, too good (pro t)* are neither definite nor indefinite: they are specific.
As for Inverted Predication taking place in primary predication, specificity of the inverted predicate is a property that can be diagnosed by inserting a colon between the inverted predicate and its subject (Higgins 1979 and den Dikken 2006, pp. 91, 153 after him):

(344) English (den Dikken, 2006, p. 91)
An excellent doctor is Brian →
An excellent doctor is, for instance, Brian

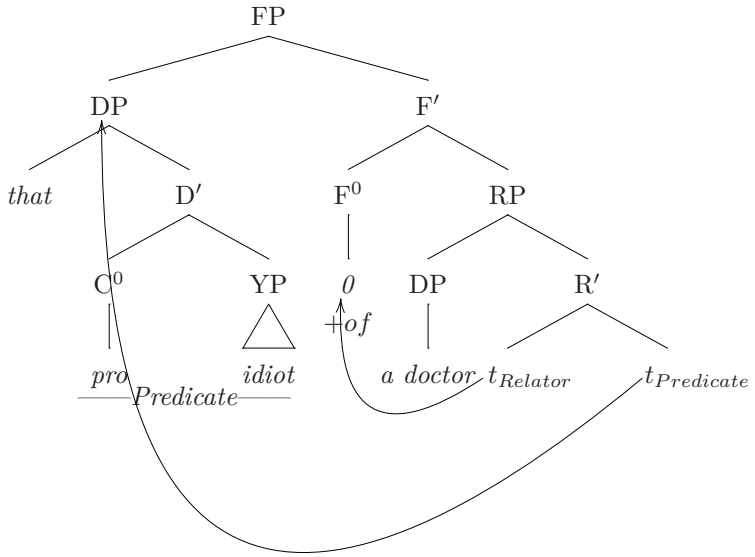
(345) English (den Dikken, 2006, p. 91)
Examples of this are the Vietnam War and the Gulf War →
Examples of this are, for instance, the Vietnam War and the Gulf War

Capitalizing on Enç (1991) (see Ch. 3), one may add that in the Inverted Predication occurring in secondary predication, specificity of the Inverted Predicate is a function of the presence of extra material following it: the complex *of DP / Subject*. In particular, the (pretheoretical) fact that in this construction the first DP is *always* followed by the complex *of DP / Subject* implies that the first DP is *always* specific.
But, recalling from Ch. 3 that specificity is a kind of 'weak' definiteness, it follows that the demonstrative *that* and the NP *too good* sitting in Spec, DP and characterized as specific are

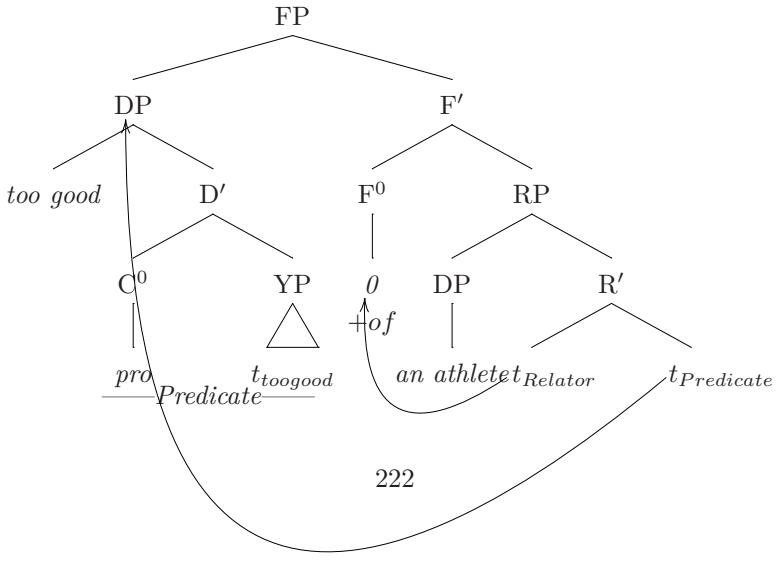
[15] Again, the account proposed here while preserving the spirit of den Dikken (2006) departs from it in its technical implementation. On den Dikken (2006, pp. 179-180, 184)'s view, in fact, the null Head whose licensing triggers Predicate Inversion is an abstract numeral. Leaving aside the intrinsic validity of this proposal, identifying the null Head with pro seems more appropriate, because it makes it possible to realign this instance of Predicate Inversion to the general pro-pattern discussed in the previous Sections.

neither totally definite, nor totally indefinite. On the other hand, the Head D / pro carries either a definite or an indefinite feature.
Accordingly, the demonstrative *that* and the NP *too good* might be plausibly have raised to Spec, DP to get coindexed with pro for other features (e.g. agreement: cp. (334)), but *not* for that of (in)definiteness.
As a corollary, the (in)definiteness feature of the covert Head D (pro), being not coindexed with the overt element hosted in Spec, DP, is not content-licensed in this local configuration and needs to find an overt element in another local configuration for the same purpose: therefore, as expected, it raises to Spec, FP. To put it differently, Predicate Inversion is triggered by the need of licensing pro that heads DP (licensing being in its turn due to the specificity encoded in Spec, DP, which prevents this latter from coindexing with pro).
An overall look at the syntactic representation of the scalar Inverted Predication *DP of DP* shows that a strong parallelism can be set between this latter and the Inverted Predication featuring a reduced RC: in both cases the Inverted Predicate is introduced by a phase headed by pro.

(346) Predicate Inversion (= 'scrambled' Canonical Predication)



(347) Predicate Inversion (= 'scrambled' Canonical Predication)



To recapitulate this Section, two kinds of Inverted Predication plausibly occur in secondary predication. They obtain by applying Predicate Inversion to the following Canonical Predication structures:

- ConjP, the syntactic manifestation of non-restrictiveness. Its modifier is a reduced RC (= CP headed by pro). A subtype of ConjP is the Numeral Classifier Construction
- DegP, the syntactic manifestation of comparison. Its modifier is a scalar DP headed by pro

A third type of Canonical Predication structure, namely

- DegP, the syntactic manifestation of comparison. Its modifier is a reduced RC

does not undergo Predicate Inversion. This is a mixed type (a DegP featuring a Complement typical of ConjP), and its particular behavior is due to a process of reanalysis.
A refinement of den Dikken (2006, Ch. 5)'s original account has been put forward, in order to show that in all these predicates Predicate Inversion is triggered by the need of licensing a pro sitting in the Head of CP / DP.
This interpretation, albeit departing from den Dikken (2006, Ch. 5)'s account in its implementation, is totally in line with its theoretical core and, in addition, has the advantage of providing an unified representation of the syntactic structure of the Inverted Predicate, lacking in den Dikken (2006, Ch. 5)'s account.
Capturing the parallelism between the just mentioned Inverted Predicates has implied a revision of the notion of NP, which is better understood as DP.
Nevertheless, for expository purposes the more familiar term NP will be used in the remainder of this work, unless otherwise stated.

5.2.5 The Reinstated Word Order

The derivation undergone by the Canonical Predication does not end up with Inverted Predication. This latter, in fact, can be further manipulated to yield a new syntactic structure, which will be designated here as REINSTATED WORD ORDER from reasons that will become clear shortly.
The manipulation of the structure in its essence consists of the fronting of the Subject NP (along with the Linker) over its inverted Predicate / modifier:

(348) Canonical Predication → Predicate Inversion → Inverted Predication
Inverted Predication → Fronting of the Subject NP → Reinstated word order

The Reinstated Word Order can be exemplified by the following Thai construction, where the noun *rôm* 'umbrella' is followed by the particle *thîi*, an adjective, a complex numeral plus classifier and a demonstrative:

(349) Thai (den Dikken, 2006, p. 232)

rôm thîi jàj sāam khan nán
umbrella LINKER big three CLF this

'those three big umbrellas'

This phenomenon seems to correlate with a particular informational profile:

(350) Thai (den Dikken, 2006, p. 232)

[rôm]^{NEW} thîi [jàj sāam khan nán]^{OLD}
umbrella LINKER big three CLF this

'those three big umbrellas'

Perhaps a step back is needed to clarify this point. In Canonical Predication (say, (351)) the *initial* Subject NP qualifies the entity it refers to as old information, while its Predicate / modifier is a piece of new information for the listener. After undergoing Predicate Inversion (cp. (352)), the Subject NP, now in *final* position, qualifies the entity it refers to as new information, while its inverted Predicate / modifier represents a piece of old information:

(351) English (den Dikken, 2006, p. 231)
Brian^{OLD} is my best friend^{NEW}

(352) English (den Dikken, 2006, p. 231)
My best friend^{OLD} is Brian^{NEW}

den Dikken & Singhapreecha (2004, p. 6) and den Dikken (2006, pp. 230-231) contend that in secondary predication the Subject NP that has undergone Predicate Inversion can be fronted to give prominence to the newness of information encoded in it (cp. (352)). In other words, (349) derives from an Inverted Predication structure along the following lines:

(353) Thai (den Dikken, 2006, p. 232)

[jàj sāam khan nán]^{OLD} thîi [rôm]^{NEW}
big three CLF this LINKER umbrella

Empirical support for positing (353) is twofold. The first piece of evidence is syntactic and lies in the fact that the particle *thîi* intervening between the Subject NP and the modifier / Predicate in (349) is exactly the same particle *thîi* that in Burmese acts as a Linker (cp. (310)), which by definition signals Inverted Predication.

It is worth observing in this connection that Thai *thîi* qualifies as a Linker also because of its inability to mark a sentential object of verbs of thinking and saying, a behavior ruling out its complementizer analysis (see Kern (2010)'s syntactic diagnostics mentioned at the beginning of Section 5.2.3). Iwasaki & Ingkaphirom (2005, Ch. 21) in fact report that such verbs in Thai introduce their sentential object by means of the particle *wā*[16].

The second piece of evidence is semantic: the informational profile of the Subject NP of (349) (new information) is typical of the Inverted Predication (cp. (352)).

But fronting of the Subject NP has a non-trivial side-effect for Syntax: the same word-order observed in Canonical Predication obtains. For this reason, (349) is designated here as Reinstated Word Order.

As just seen, however, this construction differs from Canonical Predication by:

- a Linker intervening between the (reinstated) Subject NP and its modifier / Predicate

[16]Even verbs of evaluation and emotion, which are said to introduce their sentential object by means of *thîi*, can combine this latter with *wā*, yielding *thîi-wā* (Iwasaki & Ingkaphirom, 2005, pp. 255-256). This state of affairs further supports the complementizer analysis of *wā* and the Linker analysis of *thîi*.

- a different information packaging

In a minimalist perspective, fronting of the Subject NP is generally accounted for in terms of movement. The two usual questions then arise: why and how?
To explain why movement in this case applies, den Dikken (2006) dismisses the theory of pro he invokes in the case of Predicate Inversion and opts for the standard account, which combines the Feature Checking Hypothesis with the existence of uninterpretable features (see the end of Section 4.5.2 and (218)). This choice appears to threaten the internal coherency of the theory, an issue that will be discussed in detail in Section 5.3.5.
Returning to den Dikken (2006)'s original account, the Subject NP is taken to carry an uninterpretable feature (call it uF) which cannot be 'seen' by C-I and must be deleted in order to make the derivation to converge.
Den Dikken does not further specify which is the exact nature of uF, but assumes that it has an interpretable counterpart (call it iF) located in a functional projection (call it WP) immediately above the functional projection FP headed by the Linker. More precisely, iF is located in W, the Head of WP (den Dikken & Singhapreecha, 2004; den Dikken, 2006).
Then, at a certain point of derivation, the Inverted Predication (353) from which the Reinstated Word Order arises is expanded by adding to it the functional projection WP. (354) rewrites (353) under a simplified form, while (355) illustrates merger of WP into the syntactic structure:

(354) jàj... thîi rôm$_{uF}$...

(355) [Spec W$_{iF}$ [jàj... thîi rôm$_{uF}$]]]...

But this is not enough, because deletion, under this account, is an operation applying only in a local configuration, namely the Spec - Head relation. Therefore movement is invoked to bring the Subject NP to the position of Spec, WP, where it enters into Spec - Head Relation with W and this results in (356):

(356) [$\underline{\text{rôm}_{uF}}$ thîi + W$_{iF}$ [jàj... t $\underline{\text{t}}$]]...

Actually, (356) is a (simplified) syntactic representation of (349): in other words, movement has yield the Reinstated Word Order.
A closer look to (356) also reveals *how* movement takes place: the technology accompanying the transformation from Predicate Inversion to Reinstated Word Order is exactly the same found in the transformation from Canonical Predication to Predicate Inversion (see Section 5.2.2).
The higher Head W, phonologically realized as zero, is the Attractor, while the lower Subject NP *rôm* sitting in Spec, RP is the Attractee (see Section 5.2.2). Accordingly *rôm* raises to Spec, WP to get its uF checked and deleted against iF of the Head W (Compl-to-Spec movement).
There is however a disturbing factor between the required Attractor and the Attractee: this is the Minimal Link Condition, which forces the Head W to attract to Spec,WP the constituent closest to it. But this latter is the Inverted Predicate / modifier *jàj*, *not* the Subject NP *rôm*.
The Subject NP *rôm*, in fact, is the potential Attractee of the Head F, ie the Linker *thîi*:

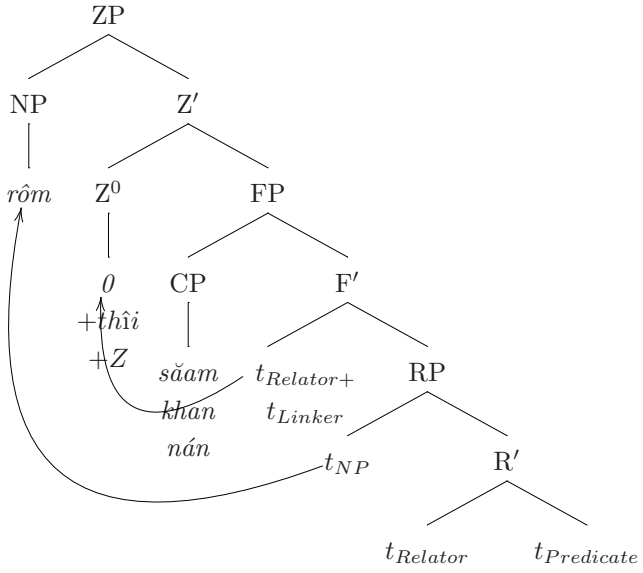
(357) [Spec W [jàj... thîi [rôm]]]
ATTRACTOR #1 ATTRACTEE #1 ATTRACTOR #2 ATTRACTEE #2

This problem is circumvented through Domain Extending Head Movement, by virtue of which the Head F (*thîi*) attaches to the left of the Head W, but in so doing it is still able to attract the Subject NP *rôm*:

(358) [Spec thîi+W [jàj... t [rôm]]]
ATTR.OR #2 + ATTR.OR #1 ATTR.EE #1 ATTR.EE #2

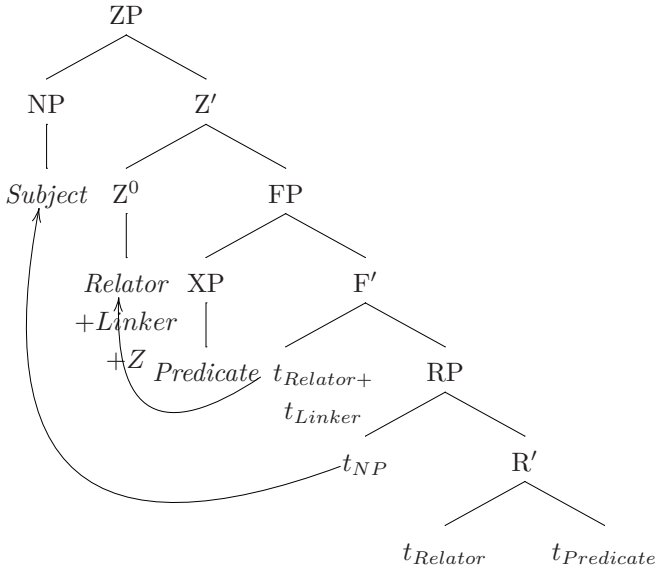
This results in the Reinstated Word Order (349, 356) above. A tree representation of this construction is given below (recall from the previous Section that the modifier / Predicate of *rôm* is a reduced RC because it features a complex Numeral plus classifier)[17]:

(359) Predicate Inversion *cum* Reinstated Word Order



Or, more abstractly:

(360) Predicate Inversion *cum* Reinstated Word Order



[17]See fn. 7 on the departure point of this derivation (Canonical Predication).

5.3 Towards an unified syntactic theory of predication

Den Dikken's *syntactic* theory of predication is couched in the same antisymmetric framework of Moro (2000)'s. The common syntactic scenario may lead one to wonder whether these two theories can be brought into comparison and, if the case, to which extent.
At first look, they seem to sensibly differ relative to the two main phenomena they are meant to capture, namely Canonical Predication and Inverted Predication.
Canonical Predication, in fact, is regarded by Moro (2000) as a non-primitive entity, actually a POS that has undergone nonhead movement (see end of Section 4.5.3 and (243)), while den Dikken (2006, p. 62 *et passim*) explicitly dismisses this view. Focusing on secondary predication, den Dikken (2007) goes even further, claiming that Canonical Predication is a primitive entity *tout court*, in the sense that it cannot be derived via movement by *any* legitimate syntactic structure displaying X-bar format.
The only common point between the two theories is that Canonical Predication is thought of in purely antisymmetric terms, and hence assigned an X-bar format (RP). This state of affairs can be summarized as follows:

(361) Canonical Predication

a. Moro: POS → nonhead mov. → RP (cp. (243))
b. Den Dikken: RP (cp. (260))

As for Inverted Predication (FP), both the theories conceive it as a non-primitive entity that obtains through *nonhead* movement. Nevertheless, they differ in which syntactic structure is manipulated to yield it. Moro (2000) takes the syntactic structure at issue to be the same POS yielding Canonical Predication (see (244)), den Dikken (2006) to be the Canonical Predication (260) itself. In so doing, den Dikken (2006) introduces a further step in the derivation, namely *head* movement:

(362) Inverted Predication

a. Moro: POS → nonhead mov. → FP (cp. (244))
b. Den Dikken: RP → nonhead mov. + head mov. → FP (cp. (261))

There are however some reasons to believe that the just illustrated contrasts are more apparent than real, as it will be seen in the next Sections.

5.3.1 Relators and Base-generation

Den Dikken (2007) proposes to reconsider Ouhalla (2004)'s interpretation of the Semitic NP, which is based, among others, on both Amharic and Arabic data:

(363) Amharic (Ouhalla, 2004, p. 294)

lïj-u yä-gäddälä-w ibaab
boy-the GM-killed-the snake

'the snake the boy killed'

(364) Amharic (Ouhalla, 2004, p. 295)

yä-lïj-u däbtär
GM-boy-the notebook

'the boy's notebook'

(365) Levantine Arabic (Ouhalla, 2004, p. 291)

baTT-it illi ʔakal-naa-ha
duck-*at* RELM ate-we-her

'the duck we ate'

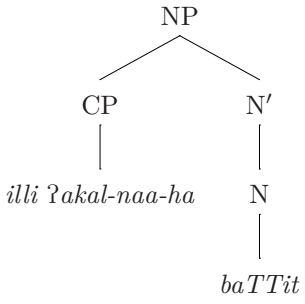
(366) Levantine Arabic (Ouhalla, 2004, p. 292)

suur-it l-bint
picture-*at* the-girl

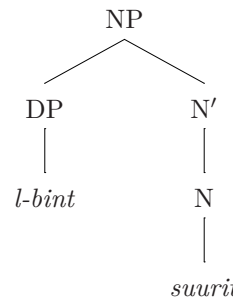
'the girl's picture'

Ouhalla (2004) develops Feghali (1928)'s remark that (365) and (366) share some structural properties in spite of the different modifiers they feature (RC and genitival phrase): e.g. in both (365) and (366) the prenominal article is lacking and the adnominal marker 'tā' marbūṭah ' surfaces as *-t* (see Ch. 3).
Ouhalla (2004) extends this analysis to the Amharic equivalents of Arabic (363, 364) and capitalizing on the parallelism between the Semitic RC (365,363) and genitival phrase (366,364), he assigns these two (apparently) different constructions the same position in the syntactic structure. Ouhalla (2004) casts his claim in an antisymmetric framework, to the effect that this position is identified with Spec, NP, as exemplified by the Arabic data[18]:

(367)



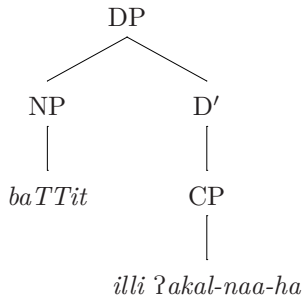
(368)



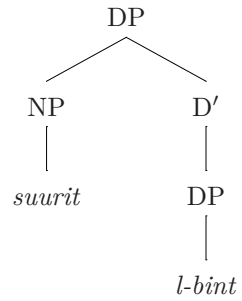
[18]For a non-antisymmetric approach to the Arabic NP, see Bardeas (2009). One of the guiding ideas of this work is that Head-to-Spec movement is allowed in Syntax, à la Matushansky (2006). See also end of Section 4.5.3 for more details.

He also puts forward that the surface word order is the result of nonhead movement, which (without going into technicalities) raises the Head N as a a part of a larger phrase (NP). In so doing, Ouhalla (2004) extends to the Semitic RC the antisymmetric analysis worked out by Ritter (1991) and further developed by Cinque (2003) and Shlonsky (2004) for the Semitic Construct State. Slightly simplifying, the syntactic structure that all these works assign to the surface word order is the following:

(369)



(370)



den Dikken (2007, p. 305) refines Ouhalla (2004)'s account observing that the Amharic NP (363, 364) behaves as the Arabic NP in that the RC and the genitival phrase fulfill the same syntactic position, but differs from it for the word order of the modification-structure:

(371) Modifier + Noun

and, furthermore, for the presence of a particle *yä*. The co-occurrence of these two phenomena prompts den Dikken (2007) to interpret the Amharic NP as an instance of Inverted Predication and hence to regard *yä* as a Linker. In this connection, den Dikken (2007) also points out that contrary to the expectation, the Linker is not found between the modifier / Predicate and the Subject NP, but precedes them: in his view, this is due to further movement of *yä* to some syntactic position higher than the modifier / Predicate (an issue that will be not further discussed here).

As a corollary, he takes the null Head intervening between the Subject NP *ibaab, däbtär, baT-Tit, suurit* and its modifier / Predicate (RC, genitive phrase) to be a Relator. This leads him to 'base-generate them [= the RC and the gentive phrase] as predicates, to the right of their subject' (den Dikken, 2007, p.304).

Now, interpreting some functional heads of the surface word order of the Semitic NP as Relators and Linkers in principle is not inplausible (although this idea calls for further empirical validation). What seems to be problematic is the very notion of base-generation of the modifier / Predicate of the Semitic DP to the right of its Subject NP: this strongly violates Antisymmetry (see Section 4.5.3).

Accordingly, a return to a *status quo ante* is preferable: (363,364,365,366) derive from syntactic structures along the lines of (367,368).
The syntactic structures (367,368) themselves, however, look suspicious upon closer scrutiny. The argument is twofold. First, they already in the Kaynean antisymmetric framework are made up of just *two* terminals, being a complex Spec - Head. Recall in fact that at the end of Section 4.5.3 the conclusion was arrived at that

(372) The substantive item (N) is assigned the following syntactic structure: complex Spec - Head

(372) reproduces (240) and is based on Fig. 4.5 and (219).
Second, a Spec and a Head can be likened to a Head - Head construction under certain circumstances, namely if in secondary predication the Spec behaves as a clitic. In similar cases the nonterminals at issue are *identical* and therefore create a particular kind of POS. As dealt with in detail in Section 4.5.4, the Dutch sentence (245), repeated here as (373):

(373) Dutch (Moro, 2000, p. 50)

wat heeft hij voor romans geschreven?
what has he for novels written

'what novels has he written?'

features this kind of POS. It involves the nonterminals *heeft - wat* in the derivational step (248), as illustrated in Fig. 4.11 (reproduced below as Fig. 5.1.a).
This is tantamount to saying that (367,368) are potentially the POS in Fig. 5.1.a, if it can be

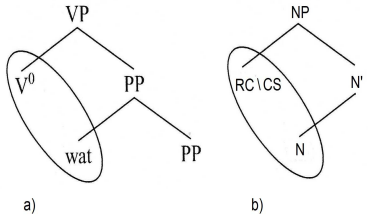


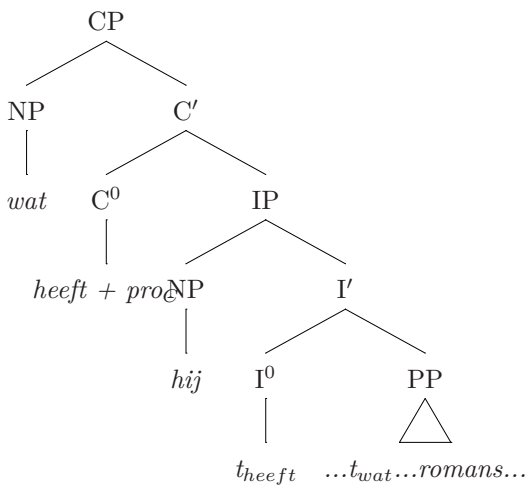
Figure 5.1: The POS Relative Clause / Construct State - Noun

demonstrated that the Arabic RC and the Arabic C(onstruct) S(tate) sitting in Spec, NP have clitic-like properties. This hypothesis is illustrated in Fig. 5.1.b.
This hypothesis, if proved true, forces Den Dikken's Relator to derive from a POS: hence, (361.b) can be in principle realigned with (361.a).

5.3.2 Heads and Points of Symmetry

The just mentioned POS *heeft - wat*, typical of secondary predication, has an interesting property: it invokes head movement to break symmetry. This is illustrated in the derivational step (249), here repeated as (374):

(374)



It seems convenient to return on this derivational step, taking into account the notion of Domain Extending Head Movement discussed in Section 5.2.2.

The Head *heeft* and the Head *wat* create a POS which is usually neutralized by moving *wat* to the Spec of the phrase immediately above it. This latter, in principle, is the VP headed by *heeft*, but movement in this case cannot apply because Spec, VP already hosts the subject *hij*. Accordingly, *wat* searches the first Spec available immediately above VP. Given that the functional projection immediately above VP is CP (see (238)), *wat* is expected to raise to Spec, CP. This is not possible, however, because of the Minimal Link Condition: the nonterminal closest to the (null) Head C is the subject *hij* sitting in Spec, VP, not the object *wat* sitting in Spec, PP - this latter, in fact, is the nonterminal closest to the Head V *heeft*.

As discussed at length in Section 5.2.2, Syntax gets around this problem through head movement: the Head V raises to the (null) Head C, attaching to its left. But *wat* is still seen as the nonterminal closest to the Head V *heeft*, which allows its raising to Spec, CP. The POS *heeft* - *wat* is thus neutralized.

At this point, it is crucial to observe that movement of *wat*, under Moro (2000)'s perspective, is an instance of Predicate Inversion and the whole construction *wat..romans* is an instance of Inverted Predication, because he analyzes *wat* as a modifier / Predicate. This instance of Inverted Predication, however, clearly involves not only nonhead movement, but also head movement, to the effect that it is structurally identical with that posited by den Dikken (2006) (cp. Section 5.2.2).

A synoptic view of the syntactic structures of *een beer van een vent* (256) and of *wat heeft hij voor romans geschreven?* (373) well illustrates this point:

(375) [wat [heeft + pro [hij [t_{heeft} [t_{wat} ...
[een beer [$0^{Relator}$ + van^{Linker} [een vent [$t_{Relator}$ [$t_{eenbeer}$ ...

Moro (2000)'s theory (362.a) is thus realigned with den Dikken (2006)'s theory (362.b) within a precise context, namely all the manifestations of Inverted Predication in the NP other than the first.

Recall in fact from (247) that scrambling of *wat* in (373) has as its starting point the POS *romans* - *wat*, which triggers Predicate Inversion without resorting to head movement.

Therefore, there are at least two aspects to the unified treatment of Inverted Predication exemplified in (375).

First, it is still to be explained why the Head-like constituent (e.g. *wat*) breaks the POS, while the proper Head (e.g. *heeft*) does not, although this is in principle possible. Second, under Moro (2000)'s approach, Predicate Inversion involving only nonhead movement remains a viable option when Inverted Predication is the first step of the derivation, but this is not the case within den Dikken (2006)'s framework, where such a kind of Inverted Predication is barred.
Tackling first the issue of Inverted Predication involving only nonhead movement, it seems convenient to call again into the picture the Inverted Predication structure whose predicate is a DP associated with *that, too* (*that / too DP of DP*).
In Section 5.2.2 this construction was characterized as having a comparative reading and exemplified by (339), here repeated as (376):

(376) English (den Dikken, 2006, p. 164)
That idiot of a doctor just wrecked my car (COMPARATIVE)

It was also shown that the comparative meaning encoded in (376) was realized in Syntax as a DegP, functioning as a Canonical Predication structure (RP), subsequently affected by Predicate Inversion. In the previous Section, a strict application of the antisymmetric assumptions underlying den Dikken (2006)'s theory has led to assign this instance of Inverted Predication the following syntactic structure, also compatible with Moro (2000)'s approach:

(377) POS → RP → nonhead mov. + head mov. → FP

Nevertheless, according to den Dikken (2006, p. 164) in some contexts the construction *that DP of DP* does not have a comparative reading - rather, it is by and large synonimous with the familiar English modification-structure in which an adjective precedes the noun. Den Dikken refers to the construction *that DP of DP* having this reading as ATTRIBUTIVE. Without reproducing his arguments in full, it is sufficient to note that the construction *that DP of DP* when used 'attributively' can undergo deletion of the sequence *of a*, but this is not possible when the same construction gets a comparative reading (say, (376), repeated below as (379)):

(378) English (den Dikken, 2006, p. 164)
a. That idiot <u>of a</u> doctor prescribed me the wrong medicine (ATTRIBUTIVE)
b. That idiot doctor prescribed me the wrong medicine (ATTRIBUTIVE)

(379) English (den Dikken, 2006, p. 164)
a. That idiot <u>of a</u> doctor just wrecked my car (COMPARATIVE)
b. *That idiot doctor just wrecked my car (COMPARATIVE)

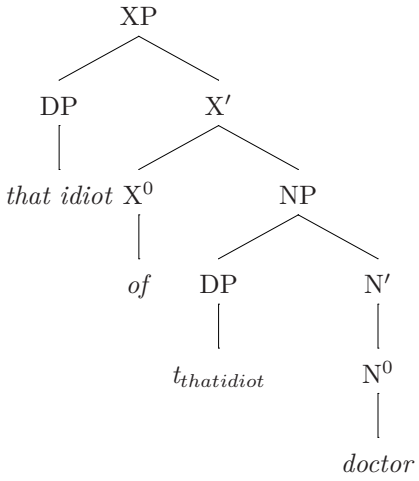
The subtle interpretive difference lies in that in (378) the referent of the entire expression is an idiot '*in his capacity* as a doctor, while he may be quite knowledgeable otherwise' (den Dikken, 2006, p. 163). On the other hand, in (379) the property of being idiot applies to 'an individual that happens to be a doctor by profession (and may, in fact, excel in that capacity)' (den Dikken, 2006, p. 163).
Plausibly, in (378) 'pruning' of *of a* takes place because the meaning of comparison is totally lacking and therefore the syntactic structure can be easily assimilated to the more familiar attributive pattern of English (modifier plus noun *tout court*). Clearly, this is not the case for (379).

On syntactic level, this is tantamount to saying that (378) instantiates no DegP and hence no Canonical Predication (RP), as den Dikken (2006, p. 172) explicitly states.
On this view, no Predicate Inversion affects no Canonical Predication (RP / DegP), so that (378) cannot be taken to be an instance of Inverted Predication: den Dikken (2006, pp. 168ff.), in fact, claims that in the attributive *that DP of DP* (378) the word order modifier / Predicate - Subject NP is not derived via movement, but base-generated as such in X-bar format.
This configuration, which he dubs REVERSE PREDICATION, whatever its precise functional nature, is as basic as the Canonical Predication (RP) (den Dikken, 2006, pp. 13, 43 *et passim*) and then the particle *of* intervening between the modifier / Predicate and the Subject NP in (378) cannot be longer seen as a Linker, but as a Relator:

(380) $[_{XP}$ Spec $[_{X'}$ Head $[_{YP}$ Compl $]]]$ = XP
Predicate Relator Subject = functional projection

Troseth (2009, pp. 38-39) remarks that the construction *too DP of DP* under certain circumstances undergoes deletion of the preposition *of* (e.g. *Attila is too good (of) an athlete*) and that it can be given either an attirbuitve or a comparative reading. In other words, formally and semantically the behavior of the construction *too DP of DP* parallels that of the construction *that DP of DP* and therefore the former is an instance of Reverse Predication as much as the latter is.
However, it has been shown in the previous Section that assuming the X-bar format to be base-generated, as den Dikken (2006) does for Reverse Predication (380), strongly violates antisymmetry, because of Compl merged at the left of the Head (see also Section 4.5.3). It has also been suggested that the X-bar format becomes in principle legitimate if it is regarded as derived from a POS, to be neutralized via movement and hence triggering raising of one of its two poles.
From this vantage point, the Predicate sitting in the Spec of (380) fulfills this position subsequent to movement: it is a pole raised in order to break a POS. But raising of the pole - Predicate is according to Moro (2000) the most typical instance of Predicate Inversion, as schematized in (362.a) and exemplified in (244).
It ensues that at least in principle (and pending further empirical validation) den Dikken (2006)'s Reverse Predication can be realigned with Moro (2000)'s Inverted Predication. On these grounds, the first derivational step of the attributive construction *that / too DP of DP* can be represented as in (381):

(381) Reverse Predication (revised)



Notice that in (381) the noun *doctor* is not adorned with the indefinite article *a*: recall from Section 4.5.3 that the (indefinite) article sits in D and therefore is merged into the syntactic structure in the final stage of derivation[19].

The theoretical move summarized in (381) has the advantage of economizing the theory: den Dikken (2006)'s Canonical Predication and Reverse Predication are no longer two independent linguistic entities, but the two possible outputs of neutralization of one and the same POS, through movement of either the pole Subject (Canonical Predication) or the pole Predicate (Reverse Predication). As a corollary, (381), if proved true, makes it possible to realign den Dikken (2006)'s theory (362.b) with Moro (2000)'s one(362.a) and therefore to give Inverted Predication an unified account.

Nevertheless, (381) again brings into prominence an issue left unsettled above, concerning the 'special' POS made up of a Head and of a phrase functioning as a Head (e.g. *heeft - wat* of Fig. 5.1.a).

It is expected in fact that both the proper Head *heeft* and the Head-like item *wat* can act as the pole to be moved to neutralize the POS, but actually only *wat* undergoes movement, while *heeft* never does so. Moro (2000) provides no explanation for this phenomenon and this depends of the fact that in his work he does not deal with the phenomenon of head-movement (Moro, 2000, p. 93, p. 126, fn. 54).

This issue is tackled in the next Section.

5.3.3 More on Heads and Points of Symmetry

The inability for Syntax to raise the proper Head *heeft* in order to neutralize the 'special' POS *heeft - wat* seems to correlate with two structural properties of *heeft qua* proper Head.

They are:

- the inability for a given Head to break symmetry via movement[20]

[19]It is temptatively suggested here that derivation of the attributive *that / too DP of DP* proceeds along the lines of Kayne (1998), involving massive phrasal movement. Technically, derivation of *that / too DP of DP* is supposed here to be highly reminiscent of the derivation posited by Kayne (1998, p. 138) for negative sentences featuring the negative item *no* (his *John considers no linguist smart*). See the Appendix 5.3.6 for a detailed account of the derivational parallelism at issue.

[20]It will be seen at the end of this Section that the inability for a given Head to break symmetry via movement

- the ability for a given Head combined with another Head to yield a single word (Word Formation)

Considering the former property first, it exists insofar as Dynamic Antisymmetry is taken as a working hypothesis.
Suppose in fact that a given Head X moves to a higher position to neutralize a POS, thus attaching to the left of a Head Y. This results in a configuration in which the two Heads are 'on the same footing' (Head - Head construction), that is they create a POS, as already discussed in Section and depicted in Fig. 4.11. Thus X, itself involved in a POS, if head-moved to Y creates a new POS.
Suppose also that in order to neutralize the new POS either X or Y is subsequently raised: in both cases, they create a new POS, because either X or Y if head-moved to a higher Head (say, Z), attaches to its left and therefore the moved Head and Z are on the same footing, which creates a new POS etc...
Movement of a Head that has already moved to another Head is known in the (pre-)minimalist literature as EXCORPORATION and is barred under Baker (1988)'s Incorporation Theory and Kayne (1994, p. 150, fn. 14)'s Antisymmetry. The just sketched account plausibly demonstrates that the prohibition on excorporation can be also derived in terms of Dynamic Antisymmetry as the *inability for an excorporated Head to neutralize a POS*.
Notice in passing that also raising of the entire complex X + Y to a higher position, although possible under Baker (1988)'s / Kayne (1994)'s approach (see Section 4.5.3), does not neutralize the POS: the mutual symmetrical relationship between X and Y holds irrespective of the portion of syntactic structure they are located.
Therefore, Head-movement does not seem in principle to be an appropriate strategy to neutralize a POS: true, it is potentially able to neutralize a POS, but in so doing it creates a new one, thus giving rise to a *progressus ad infinitum*.
This plausibly explains the inability for Syntax to raise the proper Head (e.g. *heeft*) in order to neutralize the 'special' POS (e.g. *heeft* - *wat*), that is why in (373) *wat* raises while *heeft* does not: *wat* sits in Spec, PP and hence is immune to the problems affecting the items sitting in a given Head, differently than *heeft*, which sits in Head, I.
It should be remarked in this connection that the instance of head-movement designated as Domain Extending Head Movement in Section 5.2.2 does not make exception to this claim. From the standpoint of Dynamic Antisymmetry this phenomenon is no more than head-movement taking place to help nonhead-movement, in its turn occurring to neutralize a POS (see (374) and (375)).
Yet, the problem still remains that a Head - Head construction manifesting itself to help nonhead-movement - and to neutralize a POS - creates itself a new POS which cannot be neutralized via further movement.

is not a property intrinsic to Heads. Rather, a given Head in principle has the ability to do so, but Syntax does not immediately resort to it to break symmetry via movement for reasons of economy: head-movement is a costly operation (where 'economy' and 'costly' are to be understood in a technical sense, as discussed below). Such an approach leaves open the possibility that under certain circumstances head-movement can take place: this is the case of the so-called Domain Extending Head Movement discussed in the previous Sections, as well as of some instances of clitic-movement discussed in Moro (2000, Ch. 3), an issue that will be taken up shortly. Thanks to Andrea Moro for helpful discussion about the theoretical topics dealt with in this and in the next Sections.

Upon closer scrutiny, in fact, *heeft* of (374, 375) attaches to the left of the Head C to help movement of *wat*, but in this derivational step the POS *heeft -wat* is neutralized and the new POS *heeft - C* is created.
Domain Extending Head Movement makes the problem of the POS *heeft - wat* more complicated:

(382) why cannot *heeft* raise as much as *wat* does in order to neutralize the special POS *heeft - wat*, but does raise after raising of *wat*?

The explanation given above (*heeft* does not undergo head-movement in order to avoid to create a new POS) is therefore not convincing, because head-movement of *heeft*, when Domain Extending Head Movement is required, *does* apply and creates a new POS.
To get things worse, one may raise the simple objection that even setting aside Domain Extending Head Movement, more generally in the (pre-)minimalist literature head-movement *does* take place and is a widespread phenomenon, in spite of the fact that it creates a new POS.
To this, one may add that in Syntax there occurs a phenomenon highly reminiscent of head-movement, namely Abstract Incorporation, as discussed at length at the end of Section 5.2.1. Recall in fact that a given Head, if *adjacent* with a *coindexed* Head embedded into the Spec immediately below it (Abstract Incorporation) behaves as a Head attaching to the left of another Head (head-movement).
Recall also that Abstract Incorporation, according to Baker (1988) and Julien (2002) 'behaves as' head-movement in a technical sense: both of these *syntactic* phenomena have one and the same *phonological* effect, namely *word-formation*.
It ensues that Abstract Incorporation, if structurally equivalent to head-movement, is a Head - Head construction and hence constitutes a POS like head-movement.
The issue of Abstract Incorporation therefore adds another problem to the one above (382):

(383) a Head-Head construction, either resulting from head-movement or from Abstract Incorporation, creates a POS that cannot be neutralized via movement. So, how could it be a widespread and legitimate phenomenon in Syntax?

Nonetheless, the issue of Abstract Incorporation has the merit of bringing into examination the latter structural property of Heads, namely *the ability for a Head combined with another Head to yield a single word (Word Formation)*.
This is a phenomenon first observed in Baker (1988) and studied within an antisymmetric / cartographic framework in Julien (2002). It has been seen throughout this Chapter that a Head can combine with another Head (at least) in two ways:

- via head-movement (*derivationally*: as a result of movement)
- via Abstract Incorporation (*representationally*: as a result of *local* coindexing)

These two processes are schematized below in a simplified form in (384, 385), where { } indicates a single word:

(384) $[_{\mathrm{YP}}\ \mathrm{Y}\ [_{\mathrm{XP}}\ \mathrm{X}\ ]] \rightarrow [_{\mathrm{YP}}\ \{\mathrm{X{+}Y}\}\ [_{\mathrm{XP}}\ \mathrm{t}_X\ ]]$

(385) $[_{\mathrm{YP}}\ \mathrm{Y}_i\ [_{\mathrm{XP}}\ \mathrm{X}_i\ ]] \rightarrow [_{\mathrm{YP}}\ \{\mathrm{Y}_i\ [_{\mathrm{XP}}\ \mathrm{X}_i\}\ ]]$

They are exemplified below:

(386) Chichewa[21] (Baker, 1988, p. 148)

Catherine ana-kolol-ets-a mwana wake chimanga
Catherine AGR-harvest-made-ASPECT child her corn

'Catherine made her child harvest corn'

(387) Moroccan Berber (Baker, 1988, p. 196)

Y-ss-ttc u-rjaz a-rba
3S.SBJ-made-eat SBJ-man OBJ-boy

'The man made the boy eat, fed the boy'

A representation of the relevant portion of (386) involving head-movement (384) is the following:

(388) $[_{\text{YP}}$ ets $[_{\text{XP}}$ kolol $]] \rightarrow [_{\text{YP}}$ {kolol+ets} $[_{\text{XP}}$ t_{kolol} $]]$

A representation of the relevant portion of (387) involving Abstract Incorporation (385) is the following:

(389) $[_{\text{YP}}$ ss_i $[_{\text{XP}}$ ttc_i $]] \rightarrow [_{\text{YP}}$ $\{\text{ss}_i$ $[_{\text{XP}}$ $\text{ttc}_i\}$ $]]$

Of particular relevance for Dynamic Antisymmetry is that the Head - Head construction arising from head-movement or Abstract Incorporation more precisely are two poles bearing two kinds of information, syntactic / nonterminal and phonological / terminal (see beginning of Section 4.5.4). Each pole is therefore better seen as a pair nonterminal - terminal: X, x and Y, y.

More concretely, in (388) the former pole is the pair Head V, *kolol*, the latter the pair Head V, *ets*. In (389) the former pole is the pair Head V, *ss*, the latter the pair Head V, *ttc*.

What is needed now is eliminating the POS by eliminating one of the two poles it is made of. The basic idea behind Moro (2000)'s proposal is the following: the problem of a POS is that it is an output of Syntax, where the poles A and B are not ordered. Hence A-P is not able to 'decide' whether the pole A precedes the pole B or vice versa. If one of the two poles is deleted, this problem no longer exists[22].

More precisely, what is needed is eliminating what can be seen by A-P and disturbs it, namely one terminal: either *kolol* or *ets* in (388), either *ss* or *ttc* in (389). There is no reason to eliminate either the former nonterminal V or the latter one in (388, 389): A-P cannot see them.

If, however, eliminating a terminal from a pole of the POS is fine with the module A-P, this raises a non-trivial problem for the module Syntax: its property of Encapsulation / Structure Preservation Principle, in fact, imposes that no information be eliminated in the course of computation (see beginning of Section 4.5).

This gives rise to the following scenario: on one side, the module A-P needs to see only one terminal of the output of Syntax, made up of two terminals x,y, on the other side the module Syntax cannot delete the phonological information carried by the two terminals x,y, because of Encapsulation / Structure Preservation Principle.

[21] A Bantu language spoken in Central and South Africa.

[22] Recall from Section 4.5.3 that an alternative way to solve the same problem would be, at least in principle, making the poles A and B ordered in Syntax by adding more structure. But in the present case this solution cannot apply: head-movement and Abstract Incorporation affect the X-bar format, a full-fledged syntactic structure that does not tolerate any addition of nonterminal (see Section 4.5.3 and in particular Fig. 4.8).

It ensues that the ideal compromise would be preserving all the phonological information carried by the two terminals x, y (as required by Syntax) while converting them into just one terminal (say, z), as required by A-P.
This is precisely what Word Formation does: while leaving the phonological information carried by the terminals x,y unaltered, (as required by Syntax) eliminates one terminal (as required by A-P) by fusing it with the other terminal, that is making one terminal z out of the two terminals x,y. In so doing, Word Formation neutralizes the POS.
Adopting a more familiar notion, it is as if x and y are two summands (say, 2 and 5). Before addition is performed, summands are possessed of the property of COMMUTATIVITY - in simpler words, summands are *unordered*: one can arrange them both in a sum left-to-right, and right-to-left, with no change in the final result (2 + 5 = 7, 5 + 2 = 7).
On the other hand, the terminal z can be likened to the sum of summands. After addition is performed, the sum of summands by definition is not possessed of commutativity (it is just one number: say, 7) but preserves the numerical value of its summands (2 + 5) and in fact their equality can be asserted by an equation (e.g. 2 + 5 = 7 etc...). By the same token, x + y = z insofar as the phonological content is concerned.
The picture that emerges from the discussion so far is that the POS in this case is not neutralized by moving *either* of the poles (movement), in order to leave just *one pole* of them in situ, but fusing *both* of them into *one pole* (Word Formation). Regardless of the strategy, what A-P sees in both the cases is one and the same result: one and only one terminal is associated with the two nonterminals, to the effect that there is no set of two unordered terminals (POS) inhibiting A-P from assigning them an order on the time axis (Saussurean linearity). This is illustrated in Fig. 5.2, which is a refined version of Fig. 4.9.

Therefore Word Formation, if combined with Dynamic Antisymmetry, explains how head-

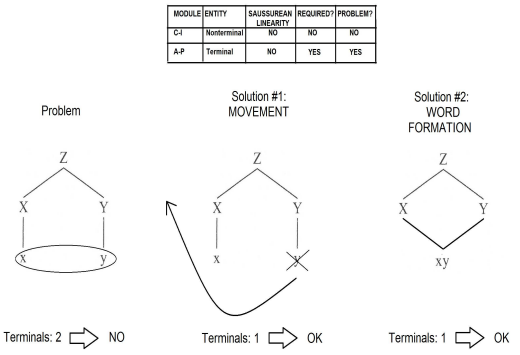


Figure 5.2: Neutralization of a POS

movement and Abstract Incorporation become legitimate pieces of syntactic structure. In consequence of this, the problem raised in (383), here repeated as (390):

(390) a Head-Head construction, either resulting from head-movement or from Abstract Incorporation, creates a POS that cannot be neutralized via movement. So, how could it be a widespread and legitimate phenomenon in Syntax?

is answered as follows:

(391) a Head-Head construction, either resulting from head-movement or from Abstract Incorporation, creates a POS that is not neutralized derivationally (via movement) but *representationally (via word-formation)*. Representational neutralization of the POS makes the Head - Head construction legitimate in Syntax.

This extension of Moro (2000)'s Dynamic Antisymmetry has three important corollaries:

(392) What is traditionally called 'word' is (recursive) 'packaging' of two items that cannot be located on the time axis and hence are unpronounceable. Their 'packaging' enables them to be located on the time axis and to be pronounced. Technically, (recursive) 'packaging' is representational neutralization of a POS

(393) An Head - Head construction / POS resulting from Abstract Incorporation blocks movement of the poles it is made of because it neutralizes its POS representationally via Word Formation. Hence movement, which performs the same task derivationally, is not needed.

(394) An Head - Head construction / POS resulting from Head-movement blocks further movement of the poles it is made of because it neutralizes its POS representationally via Word Formation. Hence further head-movement, which performs the same task derivationally, is not needed.

The first corollary has many theoretical and empirical implications that cannot be explored in this work. Only two aspects of it will be pointed out here. The former concerns the notion of monomorphemic word, the latter revolves around how the morphological types are conceived within the current linguistic debate.
As for the former issue, (392) entails that a monomorphemic word is actually to be reinterpreted as a bimorphemic word made of a covert and an overt item. In the best case, the covert item is in diachrony an overt item that has been dropped for some phonological reason. This hypothesis is evidenced by (German and) English apophonic plurals.
Simplifying, they are plurals resulting from the change of a stem-vowel of the SG form, as exemplified by the English pair *foot - feet*.
Diachronic linguists (see e.g. De Dominicis 2003, and references therein) have long observed that apophony is no more than generalization of a simple process of vowel-change caused by a surrounding vowel: for example, *feet* is a former stem *foot-* to which the PL morpheme *-i* is suffixed. It is precisely addition of *-i* to *foot-* that turns *foot-* into *feet-*: */foot-i/ > [feet-i]*. Later, the PL morpheme *-i* of *feet-i* is dropped, which results in a zero PL morpheme.
From the vantage point of Dynamic Antisymmetry, both *feet-* and *-i* are Heads (the former substantive, the latter functional) and therefore they create the POS *feet - i*, which is representationally neutralized via Word Formation (394), yielding the bimorphemic word *feet+i*.
The bimorphemic word at issue is made of two overt items and clearly belongs to the agglutinative type[23], because each of the two morphs it is made of encodes a single feature (*feet-* $\leftrightarrow$ *FOOT*, *-i* $\leftrightarrow$ PL). This amounts to saying that:

(395) the POS Head - Head construction featuring two overt poles and neutralized via Word Formation is what is traditionally called (process of) agglutination

[23]See fn. 56 of Ch. 3 for an overview of morphological types.

Notice that when the functional item (PL) is not phonologically realized, subsequent to dropping of *-i*, the POS Head - Head construction no longer exists, because one of its poles has been eliminated for independent phonological reasons. At this point one may temptatively assume that (representational) neutralization of a POS applies 'blindly', although conditions for its applications are no longer met.
This results in the POS *feet* - *zero*, which undergoes Word Formation (*feet+zero*). This is a bimorphemic word made of an overt and a covert item. One may also speculate that the bimorphemic word featuring a covert item is reanalyzed as a monomorphemic word *feet*. Technically, two nonterminals (*FOOT*, PL) are associated with one terminal (*feet*). Plainly, the monomorphemic word at issue belongs to the fusive type, because one single morph (*feet*) encodes two features (FOOT, PL). In other words,

(396) the POS Head - Head construction featuring a covert and an overt pole and neutralized via Word Formation is what is traditionally called (process of) fusion

Pushing this line of reasoning to its limits, it could be temptatively assumed that the isolating type is no more than the POS Head - Head construction featuring two overt poles, typical of agglutination (395), which for some unknown reason does not undergo neutralization via Word Formation.
This scenario is summarized in Fig. 5.3)[24].
These speculations about the monomorphemic word then bring the discussion to the issue of

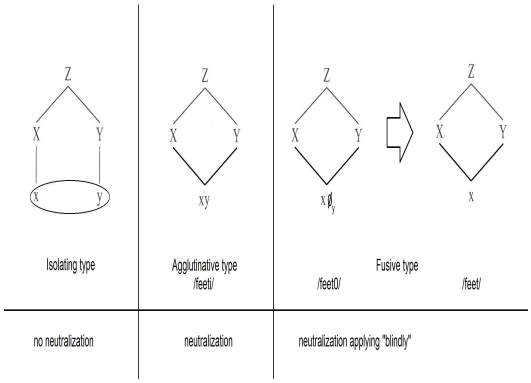


Figure 5.3: Morphological types and the POS Head - Head construction

how agglutination and fusion can be thought of in the light of Dynamic Antisymmetry, and, more generally, of the current linguistic debate.
Under the perspective of Dynamic Antisymmetry the just described agglutinative and fusive types, far from being 'inherently unnatural' and 'a disease, a pathology of language' (Aronoff 1998, p. 413 and Thornton 2005 after him) are a natural and logical consequence of the interaction between Syntax and A-P.

[24]It could be added that in generative phonology (see Chomsky & Halle 1968 and Section 4.3) association of two nonterminals with just one terminal (e.g. *TAKE*, PAST ↔ *took*) is designated as Phonetic Representation. Hence:

(1) A phonetic representation diachronically arises from reanalysis of the POS Head - Head construction featuring a covert and an overt pole

Agglutination and fusion, in fact, are to a less or more degree Word Formation (395,396), which amounts to saying that they are a representational strategy to make A-P understand the output of Syntax (technically, to neutralize a POS).
If one espouses Chomsky et al. (2002)'s evolutionary view on human language, agglutination and fusion are a natural and logical consequence of the shift that turned Syntax and A-P from separate to interacting modules (see previous Chapter): once Syntax and A-P are no longer incommunicado (as it still happens in other primates etc...), Word Formation is needed to ensure exchange of information between them, very much as X-bar format and movement are.
The second and third corollaries are more relevant for the present work.
(393) forces a reconsideration of den Dikken (2006)'s assumption that in Canonical Predication Abstract Incorporation (reanalysis) blocks movement of the modifier / Predicate because it licenses pro. A case in point is (255), reproduced below as (397):

(397) Dutch (den Dikken, 2006, p. 174)

$[_{\text{DegP}}$ een vent $[_{\text{Deg}'}$ als$_i$ $[_{\text{CP}}$ Op$_i$ $[_{\text{C}'}$ pro$_i$ [een beer]]]]]

'a bloke like a bear'

This account, in fact, takes pro (C) to get licensed against an overt counterpart coindexed with it (Deg), made available to it by Abstract Incorporation. Consequently, pro no longer needs to move higher in the tree to find such a counterpart. But den Dikken (2006)'s account has nothing to say about the phenomenon of Word Formation typical of Abstract Incorporation (see end of Section 5.2.1).
On the other hand, Dynamic Antisymmetry explains not only why Abstract Incorporation blocks movement (manifestation of an alternative strategy to neutralize a POS), but also why Word Formation occurs (deletion of a terminal).
Thus, pro licensing is no longer thought of as a factor blocking movement (and yielding Word Formation). Rather, both movement blocking and pro licensing are consequences of Word Formation: den Dikken (2006) while succeeding in isolating the phenomenon of movement blocking, mistakes an effect for the cause.
Consider again in this light Abstract Incorporation of English *like*, illustrated in (270), (274) here repeated as (398), (399):

(398) 0_{ge} + $\underline{\text{like}}$ Op →
0_{ge} $\underline{\text{like}}$ + Op

(399) $[_{\text{Deg}'}$ 0_{ge} $[_{\text{CP}}$ $\underline{\text{like}}_i$ + Op$_i$ $[_{\text{C}'}$ pro$_i$ →
$[_{\text{Deg}'}$ 0_{ge} + $\underline{\text{like}}_i$ $[_{\text{CP}}$ Op$_i$ $[_{\text{C}'}$ pro$_i$

as well as Abstract Incorporation of Dutch *als* (275), (278) here repeated as (400), (401):

(400) al + $\underline{\text{s}}$ →
al $\underline{\text{s}}$ + Op

(401) $[_{\text{Deg}'}$ al $[_{\text{CP}}$ $\underline{\text{s}}_i$ $[_{\text{C}'}$ pro$_i$ →
$[_{\text{Deg}'}$ al + $\underline{\text{s}}_i$ $[_{\text{CP}}$ Op$_i$ $[_{\text{C}'}$ pro$_i$

0_{ge} / al- and *like / -s* are an Head (Deg) and a phrase (DP / AdvP) sitting in Spec, CP. Moreover, the phrase *like / -s* syntactically exhibits an ambiguous behavior, highly reminiscent

of clitics and *wat* (see Section 4.5.4).
On one side, *like / -s* behaves as a phrase (in technical sense, ie X-bar configuration) because it does not tolerate addition of Dependents (see the discussion surrounding Fig. 4.7 and Fig. 4.8 in Section 4.5.4).
On the other side, the phrase *like / -s* behaves as a Head. This latter, in fact, in the vast majority if not all of the examples reported by Cinque (1999, Ch. 3) is monosyllabic (cp. Fula Fulfulde *-id-* in (235), the Italian clitic *lo* in (287) and *wat*) and so is *like / -s*: English writing, in fact, obscures the fact that the bisyllabic graphic word *li-ke* is actually monosyllabic ([laik]). From the vantage point of Dynamic Antisymmetry, the ambiguous behavior of *like / -s* makes the complex 0_{ge} */ al-* and *like / -s* function as an Head - Head construction (see Section 4.5.4), which creates the POS 0_{ge} */ al - like / s*.
Now, *like / -s* also undergoes Abstract Incorporation (reanalysis), which results in Word Formation. But Word Formation has the ability to (representationally) neutralize a POS, and hence it neutralizes the POS 0_{ge} */ al - like / s*.
As a consequence, movement of Spec, CP *like / -s*, which performs the same task, is no longer needed (393). Another side-effect of Word Formation is that the item *like / -s*, previously sitting in Spec, CP and now incorporated into Deg, preserves its coindexing with pro located in C, as depicted in (399) and (401) (recall that coindexing between Spec, CP and C, in fact, is the expected result of the Spec - Head relation). Pro is thus licensed (see end of Section 5.2.1 for more details).
If Spec, CP does not undergo Abstract Incorporation, movement must apply to neutralize the POS Deg - Spec, CP: hence Spec, CP raises to a higher position. Plausibly, Spec, CP moves along with the entire CP in order to allow pro-licensing and this results in Inverted Predication. Although this analysis needs refinement[25] it is clear that the trigger of movement for Inverted Predication is neutralization of a POS, *not* pro-licensing. (256) here repeated as (402) illustrates this scenario:

(402) Dutch (den Dikken, 2006, p. 174)
$[_{\mathrm{FP}} \boxed{[_{\mathrm{CP}}\ \mathrm{Op}\ [_{\mathrm{C'}}\ \mathrm{pro}_i\ [\mathrm{een\ beer}]]]}\ [_{\mathrm{F'}}\ 0{+}\mathrm{van}_i\ [_{\mathrm{DegP}}\ \mathrm{een\ vent}\ \mathrm{t}_0\ \mathrm{t}_{CP}\ ]]]$
'a bear of a bloke'

If pro-licensing is not a trigger of movement, the question arises of what is responsible for this phenomenon in the patterns belonging to the generalization summarized in (312) and in Fig. 5.1 and reproduced below as (403) and Fig. 5.3, which in Section 5.2.3 was motivated in terms of pro-licensing[26].

(403) covert C (pro) ('Reduced RC') → RC + Noun (Kern, 2010)
overt C ('Full RC') → Noun + RC (Cinque, 2005)

[25] As it will be seen in Ch. 8, this analysis can be improved taking into account the more accurate representation of Canonical Predication and Inverted Predication hinted at in fn. 7. In particular, a refinement in the syntactic analysis of Canonical Predication makes it possible to dispense with pro licensing in the explanation of movement of CP.

[26] For the sake of completeness, it should be added that in Section 5.2.3 some apparent counterexamples to this generalization were explained invoking the Saussurean idea of *rapport associatif*, and especially its syntactic implementation by Rizzi (1986) (Theory of Arb).

Table 5.3: Syntactic position of RCs and realization of their complementizer

Position	Type	Complementizer	Predicate	Source
Prenominal	Reduced	covert (pro)	Inverted	Kern (2010)
Postnominal	Full	overt	Canonical	Cinque (2005)

The postnominal position of the full RC can be accounted for as a phenomenon of Abstract Incorporation (reanalysis) between a (null) Relator and Spec, CP along the lines of (396). Middle English RC (264) repeated here as (404)

(404) Middle English (Radford, 2004, p. 229)

a. ... in every peril which that is to drede ...

b. ... in every peril $[_{\text{CP}}$ which $[_{\text{C}'}$ that [is to drede ...

is illustrative of this proposal:

(405) 0 + which + that →
0 which + Op + that

Or, more accurately:

(406) $[_{\text{R}'}$ 0 $[_{\text{CP}}$ which $[_{\text{C}'}$ that →
$[_{\text{R}'}$ 0 + which $[_{\text{CP}}$ Op $[_{\text{C}'}$ that

Indirect evidence for Abstract Incorporation (405), (406) comes from phonological deletion of the *wh-* relative pronoun diachronically documented in English (a possible equivalent of (404) in modern English is *in every danger that is...*).
In this language, in fact the functional *Head* intervening between a noun and its modifier tends to be covert, differently from e.g. Modern Greek, as the contrast between (333) and (335) illustrated below shows[27]:

(407) a. afto to vivlio
this book / *this the book

b. $[_{\text{DP}}$ afto $[_{\text{D}'}$ to $[_{\text{YP}}$ vivlio
$[_{\text{DP}}$ this $[_{\text{D}'}$ 0 $[_{\text{YP}}$ book

(408) a. yellow book *yellow the book

b. $[_{\text{XP}}$ yellow $[_{\text{X}'}$ 0 $[_{\text{YP}}$ book

Hence, turning of the *wh*-element from an overt item (Middle English) to a covert one (Modern English) is puzzling if it sits in a Spec, but perfectly natural if it sits in a Head subsequent to Abstract Incorporation (reanalysis) and Word Formation.
Dynamic Antisymmetry seems to be superior to den Dikken (2006)'s theory in that the former is able to derive deletion of *wh*-element left unexplained in the latter. At the same time, Dynamic Antisymmetry is compatible with Rizzi (1986)'s Theory of Arb, which derives English counterexamples to (403) and Fig. 5.3 (see fn. 26).
The counterexamples at issue are (restrictive and non-restrictive) RCs which are postnominal

[27]Notice that in (408) a covert Head must be posited also if Dynamic Antisymmetry is adopted. According to Moro (2000, p. 54), in fact, an adjective juxtaposed to a NP cannot be interpreted as a POS.

but reduced - say, *his girl friend, a modest person, laughs about that* and *the book I bought*. In dynamic antisymmetric terms, they are the result of the same process of Abstract Incorporation taking place in the full postnominal RC, with no additional machinery. The only difference lies in the fact that C in this case is covert (pro) rather than overt (*that*) but the overt or overt realization of C is taken care of by Theory of Arb (see Section 5.2.3) and is not relevant for Abstract Incorporation: the Head C is found outside the local domain Abstract Incorporation operates over. This is illustrated in the following:

(409) 0 + which + pro →
0 which + Op + pro

Or, more accurately:

(410) $[_{R'}$ 0 $[_{CP}$ which $[_{C'}$ pro →
$[_{R'}$ 0 + which $[_{CP}$ Op $[_{C'}$ pro

Under the perspective of Dynamic Antisymmetry, the null Head and the clitic-like item *who* originally sitting in Spec, CP create a POS (see (396)), which is representationally neutralized via Word Formation typical of Abstract Incorporation, to the effect that no movement of Spec, CP is required.

Turning to the prenominal position of the reduced RC, Dynamic Antisymmetry accounts for it along the following lines. A reduced RC typically exhibits not only a covert relative complementizer, but also a covert Subject relative pronoun (Op). The latter, rather than the former is taken here to be the real trigger of movement.

If Spec, CP hosts a covert Op, in fact, Abstract Incorporation (reanalysis) cannot apply because it by definition involves Word Formation, a *phonological* operation requiring that at least one of the two items affected by it be phonologically realized (see (396) and Section 5.2.1).

Since representational conditions for Word Formation via Abstract Incorporation are not met, the only alternative left is neutralizing the POS derivationally, via movement. Accordingly, Op sitting in Spec, CP moves as a part of the larger chunk CP to a higher portion of the syntactic structure, along the lines of (402).

On the whole, Dynamic Antisymmetry derives the generalization that a full / postnominal RC features an overt C and that a reduced / prenominal RC features a covert C (pro) calling into the picture two facts:

- in the full / postnominal RC, an *overt* C is paired with an *overt* relative pronoun
- in the reduced / prenominal RC, a *covert* C (pro) is paired with a *covert* relative pronoun (Op)

It is precisely the overt realization of the relative pronoun paired with the overt C that is responsible for Word Formation and hence movement blocking. The pairing overt relative pronoun - overt C in the full / postnominal RC is crosslinguistically documented either in synchrony (Buli and Akan, two Niger-Congo languages spoken in Ghana: Cinque 2008, p. 115 and Saah 2010, pp. 92-94 as well as 'many dialects of Germanic languages': de Vries 2001, p. 230, fn. 18) or in diachrony (Italian and French: *il quale, lequel* < *ille qualis* as per Giacalone Ramat 2005, English: see above). As for the pairing covert relative pronoun (Op) - covert C (pro), crosslinguistic evidence is provided by the language sample reported in Kern (2010), partly reproduced in Section 5.2.2.

Thus, Dynamic Antisymmetry derives the syntax (lack vs. presence of movement) of the RC types depicted in Fig. 5.3 focusing on the overt vs. covert realization of the Subject relative pronoun, and in so doing unveils a richer correlation in the RC: the two term correlation postnominal RC - overt C, vs. prenominal RC - covert C observed in it is actually a three term correlation postnominal RC - overt Spec, CP - overt C, vs. prenominal RC - covert, Spec, CP - covert C. Consequently, the generalization summarized in (312) and in Fig. 5.1 is rewritten as (411) and Fig. 5.4.

(411) covert Spec, CP (Op) ↔ covert C (pro) ↔ RC + Noun ('Reduced RC')
overt Spec, CP ↔ overt C ↔ Noun + RC ('Full RC')

Table 5.4: Syntactic position of RCs and realization of their complementizer

Position	Type	Relative Pronoun	Complementizer	Predicate
Prenominal	Reduced	covert (Op)	covert (pro)	Inverted
Postnominal	Full	overt	overt	Canonical

The issue of pro-licensing as (alleged) trigger of movement will be further elaborated on in Section 5.3.5.

Finally, (394) solves one problem (383) but makes the other (382) worse. Consider again the phenomenon of Domain Extending Head Movement in the Dutch sentence (245) *wat heeft hij voor romans geschreven* in the light of the finding that head-movement is legitimate in Syntax. Concretely, the phenomenon under scrutiny is that head-movement of *heeft* applies in order to help nonhead-movement of *wat*, in its turn triggered by the need to break the POS *heeft* - *wat*. The finding is that, true, head-movement creates a new POS, but this latter is then neutralized via Word Formation. This scenario makes it possible for Syntax to resort to Domain Extending Head Movement: *heeft*, in fact, when head-moving to the left of the (covert) C, creates a new POS that is neutralized through Word Formation.

Given that Word Formation allows *heeft* to raise to the Head C without any problems, the question (382), here repeated as (412) is particularly cogent:

(412) why cannot *heeft* raise as *wat* does in order to neutralize the special POS *heeft* - *wat*, but does raise after raising of *wat*?

The answer plausibly is the following: the strategy of neutralizing a POS by means of Word Formation in Syntax is a viable option, but is not economical, and is thus avoided by Syntax whenever possible. An explanation along these lines implies that Syntax prefers to raise *wat* rather than *heeft* because the movement of the former is in some sense more economical than the latter. Moreover, it implies that *heeft* in a later step of derivation is raised because movement is required to neutralize a POS and no more economical option is available.

To give this idea a more precise formulation, the notion of (syntactic) 'economy' is to be understood here in the technical sense of Chomsky (1995) (see e.g. pp. 78-92 and 181-186): *the lower the number of syntactic operations that build a syntactic structure, the more economical the syntactic structure.*

Chomsky (1995, p. 91) refers to the syntactic operations as 'derivation' (see end of Section 4.3) and therefore summarizes the (axiomatic and syntactic) Economy Principle under the form of

the maxim: 'Make derivations as short as possible'.
The Economy Principle thus formulated appeals to a notion of *non-local* economy, since it takes into account an entire derivation - ie *all* the syntactic operations involved into a *complete* syntactic structure (a DP or a sentence / CP). This is clearly at odds with a key-property of Syntax, namely Locality. As already discussed in Section 4.5.1, this is no more than the empirical hypothesis that memory costs force to interpret any given syntactic structure as the (recursive) sum of many 'mini'-syntactic structures, each of them in its turn being a combination *two* items A and B.
Then an Economy Principle, if any, is at work not in the 'macro'-syntactic structure (DP, CP) resulting from the 'mini'-syntactic structures thus characterized, but in these latter, as explicitly stated in more recent versions of MP (Boeckx, 2006, p. 101, fn. 31 and references therein). Accordingly such a principle can be stated as follows (Andrea Moro, pc)[28]:

Minimalist definition 5.4 (Economy Principle)**.**

Each step of the derivation must be as short as possible.

Recall from the previous Chapter that Dynamic Antisymmetry inherits from MP the idea that two are the operations carried out in Syntax, and provides them an unified account. The former operation a is purely syntactic one and is the recursive combination of two items A and B (Merge), which yields a symmetrical structure (POS). The latter looks at A-P and consists of displacement of either A or B (Move): it applies in order to neutralize their symmetry - that is, simply put, it applies to make A and B able to be located on the time axis, when spelled-out.
In this Section it has also been discussed that another syntactic operation looking at A-P is Word Formation and that it can be interpreted in terms of Dynamic Antisymmetry as a strategy alternative to displacement. Word Formation, in fact, enables Syntax to neutralize the symmetry between A and B: displacement is derivational, Word Formation is representational.
Since this interpretation of the very notion of syntactic operation crucially rests on Dynamic Antisymmetry, what is needed is a reformulation of the Economy Principle taking Dynamic Antisymmetry into account. Following Delfitto & Melloni (2009) Economy Principle can be therefore stated as follows: the lower the number of syntactic operations that build a syntactic structure *in order to neutralize a POS*, the more economical the syntactic structure.
As Boeckx (2006, p. 100) remarks, the core of the Chomskian version of Economy Principle is that it brings into comparison *two* sets of syntactic operations, in order to count them and then to choose between them the set involving the lower number of syntactic operations.
The question then arises of whether the two sets of syntactic operations can be defined. To answer this question, one may consider that the syntactic structure they build is no more than the result of the recursive combination of *two* items A and B (Locality: see immediately above and Section 4.5.1). Accordingly, Economy Principle and the two sets of syntactic operations have to do with such a binary combination: what Economy Principle brings into comparison are

[28]Thanks to Andrea Moro for helpful discussion on this issue. He also remarks that the non-local formulation of the Economy Principle ('Make derivations as short as possible') encounters not only theoretical problems but also empirical ones. It seems to imply, in fact, that a *wh*-element only moves to an intermediate position in the syntactic structure, avoiding a higher one, because the former is shorter than the latter, but as a matter of fact in some natural languages, e.g. Italian, the *wh*-element does move not only to an intermediate position (*sai cosa$_i$ Gianni ha detto t$_i$?*) but also to a higher one (*cosa$_i$ sai t$_i$ che Gianni ha detto$_i$?*). The Italian examples are Moro's.

the items A and B or, more precisely, *the set of syntactic operations each of A and B triggers* (Boeckx, 2006, p. 101, fn. 31 and references therein).
Notice that this is even more so if Dynamic Antisymmetry is adopted, because each of the items A and B *must* trigger a set of syntactic operations in order to break the POS intrinsic to them.
A non-trivial implication of characterizing Economy in purely syntactic terms is that this principle, in the biological perspective of MP, turns out to be a constraint *internal* to the module Syntax *qua* imposed by organic limitations (e.g. working memory: Boeckx 2006, p. 101). That is, it is plausibly a Condition of Optimal Design as much as the Minimal Link Condition (cp. Section 4.5.1).
But characterizing Economy Principle as a Condition of Optimal Design *internal* to the module Syntax in its turn entails that it must interact with constraints *external* to Syntax (Bare Output Conditions), which ensure that the syntactic output (ie a given syntactic structure) as a whole can be assigned an appropriate spell-out (at A-P) and an appropriate meaning (at C-I).
It follows that Economy Principle is not absolute and rather is always influenced by the Bare Output Conditions: accordingly, Economy Principle imposes that a given set of syntactic operations be small *in Syntax*, while Bare Output Conditions impose that a given set of syntactic operations be compatible *with A-P* - in terms of Dynamic Antisymmetry, Bare Output Conditions impose that such a set must be able to break a POS to be seen by A-P.
Suppose that for some reason the combination of A and B (possibly a POS) does no longer display the smaller one of the two sets of syntactic operations typically associated with A and B. Consequently, only the set of syntactic operations involving more effort (e.g. of working memory) is left.
On the one hand, this is not desirable for Syntax, but on the other hand, such a set of syntactic operations has the ability to neutralize a POS (or, according to the Feature Checking hypothesis, to delete uninterpretable features), so that it is the only way to satisfy the Bare Output Conditions at A-P (or at C-I, following the Feature Checking hypothesis). If Syntax did not choose it, the overall derivation would crash.
It follows that a bigger set of syntactic operations (a less economical derivation) is legitimate in Syntax if and only iff a smaller one (a more economical derivation) is not available. This scenario is designated in MP as Last Resort (Condition) and is stated by Boeckx (2006, p. 102, fn. 33) as follows:

Minimalist definition 5.5 (Last Resort Condition)**.**

A syntactic operation may apply only if the derivation would otherwise result in an illegitimate representation at the interfaces.

Looking at the lack of raising of *heeft* through the lens of Economy Principle implies that this latter compares *heeft* and *wat* and detects that *wat* involves a number of syntactic operations lower than the number of syntactic operations associated with *heeft*.
It is put forward here that this is actually the case: *wat*, in fact, can neutralize the POS just raising as a phrase (one syntactic operation) to Spec, CP, while *heeft* performs the same task raising as a head to C, which results in a new POS to be further neutralized via Word Formation (*two* syntactic operations).
Thus, it is Economy Principle (in a technical sense) that bars raising of *heeft* in favor of *waat*. The single syntactic operation instantiated by the latter in fact is more economical than the

two syntactic operations instantiated by the former:

(413) *wat*: POS → nonhead movement$_1$ → neutr.
heeft: POS → head movement$_1$ → POS → Word Formation$_2$ → neutr.

But this is not the whole story: a more accurate look at the POS exemplified by *waat - heeft* also reveals that the Last Resort Condition too is at work in it.
It has been just seen that the Head *heeft* does not raise to C to neutralize the POS *heeft - wat* because *in this step of derivation* another alternative is viable, namely raising of *wat* to Spec, CP. But the Minimal Link Condition prohibits raising of *wat* because Spec, CP has not *wat* as its closest constituent, this latter being *hij* (see Section 5.2.2).
The only way to get around this problem is raising the Head *heeft* to the Head C (Domain Extending Head Movement): given that alternative solutions are not viable *in this derivational step*, the Head *heeft* raises to C - giving thus rise to a Head - Head construction / POS, to be neutralized via Word Formation. Plainly, in this scenario the Last Resort Condition applies.
On these grounds, the solution to the problem stated in (412) could be more accurately formulated as follows:

(414) X cannot raise as much as a XP does in order to neutralize the special POS X - XP, because of Economy Principle. After that XP has been selected for movement, X is too because of Last Resort. XP cannot raise to its landing site because of an intervening YP (Minimal Link Condition) and only movement of X make XP and YP equidistant to the landing site, thus allowing XP to raise to it (Domain Extending Head Movement).

(414) has an important consequence: it significantly reduces application of head-movement, which turns out to be a by-product of nonhead-movement (Domain Extending Head Movement). Notice that taking head-movement to be connected with nonhead-movement requires that the former very much as the latter has to do with A-P rather than with C-I: this is precisely what currently MP argues on independent grounds, pointing out, among others, that head-movement has no semantic import (Boeckx & Stjepanovic 2001 and references therein).
Interpreting head movement as Domain Extending Head Movement in principle allows head movement to undergo subsequent head-movement: Domain Extending Head Movement in fact is invoked whenever nonhead-movement needs its help to neutralize a POS. *This is actually the only reason for allowing head-movement to apply recursively*: recall from (394) that in the normal case a single instance of Domain Extending Head Movement does not undergo subsequent head movement because Word Formation intervenes.
It has been also pointed out throughout this Section that such a characterization of head movement and Abstract Incorporation is based on Moro (2000)'s Dynamic Antisymmetry framework, but, interestingly, it seems to be also compatible with an account along the lines of den Dikken (2006).
First, in den Dikken (2006)'s account of secondary predication, head movement applies only as part and parcel of Domain Extending Head Movement (see Sections 5.2.2 and 5.2.5), which is totally in line with (414). Second, he explicitly states that Abstract Incorporation is a strategy typical of a Head - Head construction (see Section 5.2.2).
In other words (an extended version of) Moro (2000)'s Dynamic Antisymmetry comports with den Dikken (2006) as to *how* a Head behaves. Yet, they still differ as to *what* the Head exactly is (the distinction between Relators and Linkers, crucial in den Dikken 2006 has no theoretical

status in Moro 2000) and as to *why* a Head behaves so (e.g. neutralizing a POS vs. pro-licensing).
These two issues are taken up in the next Sections.

5.3.4 Heads and copulae

It has been seen in the previous Sections that strict application of Antisymmetry has the advantage of providing an unified account of den Dikken (2006)'s and Moro (2000)'s interpretation of the syntactic structures underlying the different kinds of predication. It is put forward in this Section that strict application of Antisymmetry to den Dikken (2006)'s framework has also non-trivial consequences for a better understanding of the syntactic behavior of his Relators and Linkers.
An overall view at the Relators and Linkers exemplified in the previous Sections plausibly shows that they can be kept distinct on the basis of both phonological and structural properties.
According to den Dikken (2006), in fact, a Relator and a Linker generally have different morpho-phonological realizations (cp. Dutch *als* - *van*, English *like* - *of*, Thai zero - *thii*). Structurally, the distinctive property of a Relator is heading a base-generated configuration, that of a Linker is its co-occurrence with a Relator (contrast Inverted Predication with Canonical Predication and Reverse Predication).
There are, however, at least two aspects that falsify this claim, as already hinted at in the previous Sections.
First, according to den Dikken (2006) himself the particle *of, van* etc... is a Relator when a *that / too DP of DP* construction is attributive (den Dikken 2006's 'Reverse Predication'), but a Linker when the same construction has a comparative reading (Inverted Predication). This is exemplified by the contrast between the English attributive Reverse Predication (378) and the English comparative Inverted Predication (379), here repeated as (415) and (416):

(415) English (den Dikken, 2006, p. 164)
  a. That idiot <u>of a</u> doctor prescribed me the wrong medicine (ATTRIBUTIVE)
  b. That idiot doctor prescribed me the wrong medicine (ATTRIBUTIVE)

(416) English (den Dikken, 2006, p. 164)
  a. That idiot <u>of a</u> doctor just wrecked my car (COMPARATIVE)
  b. *That idiot doctor just wrecked my car (COMPARATIVE)

Clearly in the attributive Reverse Predication (415) morpho-phonological realization of the Relator (*of*) points to an equivalence rather than to a differentation between Relators and Linkers.
Second, considerations of Antisymmetry force to reinterpret Reverse Predication (415) as an instance of Inverted Predication: the pre-nominal modifier / Predicate cannot be base-generated, and hence it must be the result of Predicate Inversion à la Moro (2000). This gives rise to an impasse.
If Reverse Predication (415), in fact, is reinterpreted as an Inverted Predication à la Moro (2000) the particle *of* cannot be regarded as a Relator, because this implies that the configuration at issue is base-generated, which is not the case. Nor can *of* be regarded as a Linker, because *of* is not combined with any Relator, and a Linker, under den Dikken's analysis is by definition

combined with a Relator.
The only relevant property of *of* in this construction is that traditionally ascribed to the copula: mediating the relationship between a Subject NP and modifier / Predicate, regardless of word order of the constituents involved and of the more or less complex semantics (comparative vs. attributive reading of the *that / too DP of DP* constructions).
In short:

- one and the same functional item may function either as a Relator and as a Linker
- a Linker does not necessarily imply the presence of a Relator, occurring alone as much as the Relator does
- a Relator is *not* base-generated, which is precisely the behavior of a Linker

It ensues that it makes little sense to draw a dividing line between a Relator and a Linker and that the more familiar notion of copula is to be preferred instead.
However, this is only in part a return to a *status quo ante*. Consider again den Dikken (2006)'s Relators and Linkers in the light of the revised scenario sketched out in this Section. If den Dikken (2006)'s Relator is *not* base-generated, this means that it makes available an empty Spec for the nonterminal undergoing movement, regardless of whether the nonterminal is a modifier / Predicate ('Reverse Predication') or a Subject NP (Canonical Predication)[29]. In a likewise fashion, den Dikken (2006)'s Linker makes available an empty Spec for the modifier / Predicate undergoing movement, as discussed at length in Section 5.2.2 (see e.g. (304) and (305)).
The (alleged) Relators and Linkers thus characterized now share a new syntactic property: they make available an empty Spec for the nonterminal undergoing movement.
Remarkably this is precisely how Moro (2000, pp. 106-107) syntactically characterizes the copula in his framework - a desirable outcome, since den Dikken (2006)'s (former) Relators and Linkers are realigned with Moro (2000)'s copula.
The just outlined unified account of *what* the (syntactic) copula is crucially relies upon movement. This is an aspect that potentially threatens integration of den Dikken (2006)'s and Moro (2000)'s theories: as dealt with in the previous Sections, the former has it that movement has to do with C-I (Feature Checking Hypothesis), the latter with A-P (Dynamic Antisymmetry). It is not totally clear, then, how den Dikken (2006)'s account harmonizes with Moro (2000)'s in this repect. This topic is treated in the next Sections.

5.3.5 Triggers of movement

If, on one side, the way *how* movement applies seems to be the same in both the theories, den Dikken (2006) does not completely dovetail with Moro (2000) as for the reason *why* movement applies. True, one may concede that in both the theories movement is not triggered by Syntax in and by itself: in both the cases, in fact, movement manifests itself only in order to allow the syntactic structure to be 'seen' by some module other than Syntax, be it A-P or C-I (Bare Output Conditions).

[29] Recall from Sections 5.3.1 and 4.5.4 that considerations of Antisymmetry strongly suggest that movement of the nonterminal at issue is triggered by the need to neutralize a POS, along the lines of Moro (2000)'s Dynamic Antisymmetry.

One may also concede that the considerations of Antisymmetry made in Sections 5.3.1 and 4.5.4 force to explain den Dikken (2006)'s Canonical Predication and attributive Inverted Predication (Reverse Predication) in terms of Moro (2000)'s Dynamic Antisymmetry, to the effect that these two constructions involve just one trigger of movement: the need to neutralize a POS, or more generally, to make the syntactic structure compatible with A-P.
Nevertheless, if two other contructions discussed within the theory of Relators and Linkers are taken into consideration, namely comparative Inverted Predication and Reinstated word order, the rationale for movement proposed by den Dikken (2006) is clearly at odds with Moro (2000)'s one: a given constituent raises to the Spec of a given phrase because its feature needs to enter into relationship with some feature found on the Head of that phrase.
More precisely, in the case of (comparative) Inverted Predication, the Predicate / modifier undergoing Predicate Inversion contains a covert pro that raises to Spec, FP in order to be paired with its overt counterpart sitting in F / Linker (formal licensing of pro: see Section 5.2.2).
The Dutch construction (256) discussed in the previous Sections and here repeated as (417) for the sake of commodity exemplifies comparative Inverted Predication (recall that Relators and Linkers in the present work have no longer theoretical status):

(417) Dutch (den Dikken, 2006, p. 174)

een beer^{PREDICATE} van^{LINKER} een vent^{SUBJECT}
a bear of a bloke

'a bear of a bloke'

The relevant portion of derivation (304, 305, 306) is reproduced below as (418, 419, 420):

(418) [Spec F [een vent R [...pro...]]]
ATTRACTOR #1 ATTRACTEE #1 ATTRACTOR #2 ATTRACTEE #2

(419) [Spec R+F [een vent t [...pro...]]]
ATTR.OR #2 + ATTR.OR #1 ATTR.EE #1 ATTR.EE #2

(420) [pro$_i$ R+F$_i$ [een vent t t]]
ATTR.EE #2 ATTR.OR #2 + ATTR.OR #1 ATTR.EE #1

As for Reinstated Word Order, it is the Subject NP that raises to Spec, WP because it contains an uninterpretable feature needing to be paired with its interpretable counterpart sitting in W. In so doing, it undergoes deletion (Feature Checking: see Section 5.2.5).
The Thai construction (349) discussed in Section 5.2.5 and here repeated as (421) for the sake of commodity is illustrative of Reinstated Word Order:

(421) Thai den Dikken (2006, p. 232)

rôm thîi jàj sāam khan nán
umbrella LINKER big three CLF this

'those three big umbrellas'

The relevant portion of derivation (354, 355, 356) is reproduced below as (422, 423, 424):

(422) [jàj... thîi rôm$_{uF}$]]...

(423) [Spec W$_{iF}$ [jàj... thîi rôm$_{uF}$]]...

(424) [rôm$_{uF}$ thîi + W$_{iF}$ [jàj... t t]]...

Therefore, in both the cases movement appears to be triggered by the need to make the syntactic structure compatible with C-I rather than with A-P. This account, however, presents some difficulties.

On the whole, in both Spec-Head relations under scrutiny, the core of this (alleged?) semantic mechanism is unclear, because the exact nature of the feature encoded in F that enters into relationship with pro (420) as well as the (un)interpretable features carried by the Subject NP and W (424) are unknown.

Even setting aside this problem, the behavior of the Spec-Head relations (420) and (424) casts doubt on the validity of den Dikken (2006)'s Feature Checking account for other reasons.

In the Spec-Head relation typical of the Inverted Predicate (420), Spec, FP (pro) is *always* coindexed with the copula F (Den Dikken's Linker). Such a characterization of the copula F implies the incorrect assumption that this latter has *always* a pronominal nature, but there is no compelling and principled reason to assume that this is always the case: the copula R (Den Dikken's Relator), for example, is coindexed with pro only under certain circumstances (Reanalysis / Abstract Incorporation: see Section 5.2.1).

More generally, nothing in both the traditional and minimalist definition of the copula (Moro, 2000; den Dikken, 2006) hinges on the intrinsic presence of a pronominal element, feature etc... in it.

As for the Spec-Head relation instantiated by the Reinstated Word Order (424), it exhibits a puzzling behavior, given that it is optional. Contrast e.g. Thai (421) with Burmese (310), here repeated as (425), where the Inverted Predication structure featuring *thii* is not further manipulated to yield the Reinstated Word Order:

(425) Burmese (den Dikken & Singhapreecha, 2004, fn. 23, p. 36)

ca-naw weh thii sa-ouq
I buy *thii* book

'the book I bought'

No convincing explanation is given by den Dikken & Singhapreecha (2004) and den Dikken (2006) for optionality of the Spec - Head relation taking place in WP, as observed in Burmese and Thai. This phenomenon, however, cannot be dismissed so easily: for example, one may wonder how the Spec - Head relation deletes the uninterpretable feature encoded on the Subject NP if movement does not apply. A possible solution would be assuming that no uninterpretable feature is carried by the Subject NP when movement to Spec, WP is not observed, and therefore to characterize the Burmese NP of (425) in this way. But this assumption in its turn relies upon a principled explanation of why sometimes the Subject NP carries the uninterpretable feature while sometimes does not - an explanation not offered in den Dikken & Singhapreecha (2004)'s and den Dikken (2006)'s work.

On these grounds, relating movement of the comparative Inverted Predication (417) and of the Reinstated Word Order (421) to C-I seems to be highly questionable, and the only alternative left to account for the same phenomenon is relating it to A-P, that is resorting to Dynamic Antisymmetry.

In particular, if these speculations are on the right track, manifestation of pro-licensing in comparative Inverted Predication is no longer thought of as a trigger of movement, but as one of its

consequences. This claim is consistent with the conclusion arrived at on independent grounds in Section 5.3.2, where it has been demonstrated that pro-licensing does not plays any role neither in blocking movement of Canonical Predication (Abstract Incorporation) nor in allowing it (comparative Inverted Predication), representational vs. derivational neutralization of a POS being the explanation of these phenomena.
The overall picture that emerges from this Chapter is that three out of four manifestations of predication in the NP, namely Canonical Predication, attributive Inverted Predication and comparative Inverted Predication are very likely to be accounted for in terms of Dynamic Antisymmetry. This leads to think that Dynamic Antisymmetry could be invoked to explain Reinstated Word Order too. Such a state of affairs is summarized in the Table 5.5.
However, the culled evidence for positing a POS deriving Canonical Predication is mainly the-

Table 5.5: Predication structures in the NP: trigger of movement

Predication Structure	Nonterminal	Mov.	Problem	Solution	Trigger
Canonical Predication	Spec, RP	NO	Violation of antisymmetry	Dynamic Ant.	POS
Attr. Inv. Predication	Spec, RP	NO	Violation of antisymmetry	Dynamic Ant.	POS
Comp. Inv.Predication	Spec, FP	YES	Anomalous copula F	Dynamic Ant.	POS
Reinstated Word Order	Spec, FP	YES / NO	Optionality	?? (⇑)	?? (⇑)

oretical and further empirical arguments are needed to corroborate this claim. More generally, the validity on this hypothesis mainly rests on three points:

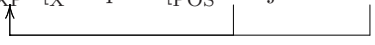
- a POS made up of a Head and of a clitic-like projection is to be found in deriving Canonical Predication
- a pure (= non-pronominal) manifestation of the copula F is to be found
- lack of movement to Spec, WP must be accounted for

Notice also that the discussion so far leads to refine Moro (2000)'s hypothesis that the copula is a Head making available an empty Spec for the syntactic constituent undergoing movement (either a modifier / Predicate or a Subject NP) (see Section 5.3.4).
Suppose that the scenario depicted in Table 5.5 is basically right: movement in *all* the instances of predication observed in the NP is triggered by the need to make the syntactic structure compatible with A-P. Then, it is the copula that allows movement and therefore that allows the syntactic structure to be given a pronunciation at A-P. Technically:

(426) The copula is a Head merged into the syntactic structure to provide a landing site with the raised pole of a POS and hence to make the derivation converge at A-P

This is illustrated in (427):

(427) $[_{XP}\ [_{X'}$ copula $[_{POS}$ Subject Predicate



5.3.6 Appendix

The derivational parallelism between the attributive *that / too DP of DP* and the negative sentence featuring the negative item *no* (Kayne 1998, p. 138 *John considers no linguist smart*)

alluded to in fn. 19 is illustrated in what follows in greater detail.
Consider the derivational step taking place after that depicted in (381), a higher phrase (call it YP) is recursively merged into the syntactic structure:

(428) ... $[_{YP}\ [_{Y'}\ Y\ [_{XP}$ that idiot $[_{X'}$ of $[_{NP}$ ...doctor...

Compl, YP, that is the whole XP of (428) *that idiot doctor* raises to Spec, YP:

(429) ... $[_{YP}\ \boxed{[_{XP}\text{ that idiot }[_{X'}\text{ of }[_{NP}\text{ ...doctor...}]]]}\ [_{Y'}\ Y\ t...$

This results in the syntactic structure (429) parallel to Kayne (1998, p. 139)'s derivational step *considers no linguist smart*:

(430) [covert Y + that idiot of + doctor]
[considers + no linguist + smart]

Then, a higher phrase (call it ZP) is recursively merged into the syntactic structure:

(431) ... $[_{ZP}\ [_{Z'}\ Z\ [_{YP}\ \boxed{[_{XP}\text{ that idiot }[_{X'}\text{ of }[_{NP}\text{ ...doctor...}]]]}\ [_{Y'}\ Y$ ...

At this point, only a portion of the XP sitting in Spec, YP raises to Spec, ZP. Raising of a portion of syntactic structure is a phenomenon known in the (pre-)minimalist literature as SUBEXTRACTION and it is affected by the so-called LEFT BRANCH CONDITION, according to which in some languages the leftmost constituent of a nominal expression cannot be extracted out of the expression containing it (Ross 1967 and much subsequent work). This results in the following scenario.
The Kaynean Antisymmetry in principle allows either the Spec *that idiot* or the Compl *doctor* of the XP *that idiot of doctor* to raise to Spec, ZP and in the normal case, Minimal Link Condition forces Syntax to raise the Spec *that idiot* to Spec, ZP and to leave the Compl *doctor* in situ, because the former is closer to ZP than the latter. But the Left Branch Condition blocks movement of the Spec *that idiot*, so that the Compl *doctor* raises to Spec, ZP instead:

(432) ... $[_{ZP}\ \boxed{[_{NP}\text{ ...doctor...}]}\ [_{Z'}\ Z\ [_{YP}\ \boxed{[_{XP}\text{ that idiot }[_{X'}\text{ of }t_{NP}]]}\ [_{Y'}\ Y$ ...

The resulting syntactic structure is (432), parallel to Kayne (1998, p. 139)'s derivational step *smart considers no linguist*:

(433) doctor + covert Z + [covert Y + that idiot of + t]
smart + covert Head + [considers + no linguist + t]

Finally, the functional projection DP is recursively merged into the syntactic structure:

(434) ... $[_{DP}\ [_{D'}$ a $[_{ZP}\ \boxed{[_{NP}\text{ ...doctor...}]}\ [_{Z'}\ Z\ [_{YP}\ \boxed{[_{XP}\text{ that idiot }[_{X'}\text{ of }t_{NP}]]}\ [_{Y'}\ Y$ ...

The remainder of XP, namely *that idiot of* raises from Spec, YP to Spec, DP (REMNANT MOVEMENT), plausibly resorting to head movement of (the covert) Z to D, as standardly assumed (Domain Extending Head Movement: Section 5.2.2):

(435) ... $[_{DP}\ \boxed{[_{XP}\text{ that idiot }[_{X'}\text{ of }t_{NP}]]}\ [_{D'}$ a $[_{ZP}\ \boxed{[_{NP}\text{ ...doctor...}]}\ [_{Z'}\ Z\ [_{YP}\ t_{XP}\ [_{Y'}\ Y$ ...

There obtains the syntactic structure in (435), parallel to Kayne (1998, p. 139)'s derivational step *no linguist smart considers*. Abstracting away from head movement, this parallelism is illustrated in the following:

(436) that idiot of + a + doctor + covert Z + [covert Y + t + t]
no linguist + covert Head + smart + covert Head + [considers + t + t]

As expected, merging the topmost functional projection DP in (434) closes the derivation of *that idiot of a doctor*, while the syntactic structure *no linguist smart considers* undergoes further movement (of the entire phrase *[considers t t]* across *no linguist smart*), which results in the sentence *(John) considers no linguist smart* (Kayne, 1998, p. 139).

5.4 Summary

The present Chapter has proposed a substantial revision of den Dikken (2006)'s minimalist theory of predication. Considering phrase structure first, it has been argued that crosslinguistically three kinds of Canonical Predication occur, all of them featuring one and the same Predicate / modifier: a phase (CP / DP) headed by pro.
In greater detail, the three kinds of Canonical Predication at issue are:

(437) ConjP, the syntactic manifestation of non-restrictiveness. Its modifier is a reduced RC, that is a CP headed by pro. A subtype of ConjP is the Numeral Classifier Construction.

(438) DegP, the syntactic manifestation of comparison. Its modifier is a scalar DP headed by pro

(439) DegP, the syntactic manifestation of comparison. Its modifier is a reduced RC (CP headed by pro)

The two former undergo Predicate Inversion, the latter does not. This is due to the fact that it is a mixed type (a DegP featuring a Complement typical of ConjP) arising from a process of reanalysis.
It has been also shown that formally reanalysis can be characterized along the lines of both Baker (1988)'s Abstract Incorporation, according to which the two items involved in it must be

- Heads
- coindexed for some feature

and Julien (2002), which refines Baker's account as follows:

(440) A given Head and the Head-within-Spec immediately below it undergo reanalysis

Accordingly, reanalysis can be represented as follows:

(441) X + $\underline{\text{Y}}$ $\rightarrow$
X $\underline{\text{Y}}$ + Op

Or, in a more accurate formulation:

(442) $[_{\text{R}'}$ X $[_{\text{CP}}$ $\underline{\text{Y}}$ $[_{\text{C}'}$ pro $\rightarrow$
$[_{\text{R}'}$ X + $\underline{\text{Y}}$ $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro

Reanalysis thus characterized has three interesting properties:

- X incorporates into it Y embedded in Spec, CP, in its turn coindexed with C / pro. Hence, X licenses C / pro via a chain (443, 444)

- Incorporation of Y into X blocks movement of CP (the Predicate / modifier)
- Incorporation of Y into X gives rise to Word Formation

(443) $[_{R'}\ X\ [_{CP}\ Y_i\ [_{C'}\ pro_i ...$

(444) $[_{R'}\ \underset{\underset{Reanalysis}{\uparrow___

Head (448.b).
The empty Spec at issue is Spec, CP, the Head is the C hosting pro (447). Nevertheless, the Minimal Link Condition prohibits raising of *wat* because Spec, CP has not *wat* as its closest constituent, this latter being *hij* sitting in Spec, IP.
The only way to get around this problem is raising the Head *heeft* to C (Domain Extending Head Movement): given that alternative solutions are not viable *in this derivational step*, the costly (but legitimate) twofold operation (448.b) is resorted to, *faute de mieux*.
Accordingly the Head *heeft* raises to C, giving thus rise to a Head - Head construction / POS, to be neutralized via Word Formation. This phenomenon is designated in MP as Last Resort Condition.
Hence, Dynamic Antisymmetry characterizes interaction of Economy Principle, Minimal Link Condition (actually, two Conditions of Optimal Design *internal* to Syntax) and Last Resort (a Bare Output Condition, *external* to Syntax) as follows:

(449) X cannot raise as much as a XP does in order to neutralize the special POS X - XP, because of Economy Principle. After that XP has been selected for movement, X is too because of Last Resort. XP cannot raise to its landing site because of an intervening YP (Minimal Link Condition) and only movement of X make XP and YP equidistant to the landing site, thus allowing XP to raise to it (Domain Extending Head Movement).

An interesting corollary of (446) and (449) is the following:

(450) Given that in the normal case a single instance of head movement does not undergo subsequent head movement because Word Formation intervenes (446), recursive head movement applies iff and only head movement is Domain Extending head Movement (449): that is, iff and only if nonhead-movement needs help from head movement to neutralize a POS.

Insofar as nonhead-movement is concerned, the overall picture that emerges from this Chapter is that three out of four manifestations of predication in the NP, namely Canonical Predication, attributive Inverted Predication and comparative Inverted Predication instantiate it as a way to neutralize a POS. This leads to think that Dynamic Antisymmetry could be invoked to explain Reinstated Word Order too. Such a state of affairs is summarized Table 5.6.
However, the culled evidence for positing a POS deriving Canonical Predication is mainly the-

Table 5.6: Predication structures in the NP: trigger of movement

Predication Structure	Nonterminal	Mov.	Problem	Solution	Trigger
Canonical Predication	Spec, RP	NO	Violation of antisymmetry	Dynamic Ant.	POS
Attr. Inv. Predication	Spec, RP	NO	Violation of antisymmetry	Dynamic Ant.	POS
Comp. Inv.Predication	Spec, RP	YES	Anomalous copula F	Dynamic Ant.	POS
Reinstated Word Order	Spec, FP	YES / NO	Optionality	?? (⇑)	?? (⇑)

oretical and further empirical arguments are needed to corroborate this claim. More generally, the validity on this hypothesis mainly rests on three points:

- a POS made up of a Head and of a clitic-like projection is to be found in deriving Canonical Predication

- a pure (= non-pronominal) manifestation of the copula F is to be found
- lack of movement to Spec, WP must be accounted for

Revising den Dikken (2006)'s theory in the light of Moro (2000) has led to a reconsideration of den Dikken (2006)'s Relators and Linkers. It has been argued here that they have no theoretical status, and that Moro (2000)'s characterization of copula is to be preferred instead. This move casts the traditional conception of the copula as an element mediating the relationship between a Subject and a Predicate in dynamic antisymmetric terms.
In this light, the copula is the Head making available an empty Spec for either the Subject NP or the modifier / Predicate undergoing movement in order to neutralize the POS they created subsequent to their combination:

(451) $[_{XP}\ [_{X'}$ copula $[_{POS}$ Subject Predicate

Technically:

(452) The copula is a Head merged into the syntactic structure to provide a landing site with the raised pole of a POS and hence to make the derivation converge at A-P

These findings impact on the traditional notion of word and of morphological type, which are likely to be different manifestations of the POS. They also unveil a three term correlation postnominal RC - overt Spec, CP - overt C, vs. prenominal RC - covert, Spec, CP - covert C, summarized in in Fig. 5.7. Apparent counterexamples to this generalization have been accounted for in terms of Rizzi (1986)'s Theory of Arb.

Dynamic Antisymmetry derives this correlation along the following lines. The constituent

Table 5.7: Syntactic position of RCs and realization of their complementizer

Position	Type	Relative Pronoun	Complementizer	Predicate
Prenominal	Reduced	covert (Op)	covert (pro)	Inverted
Postnominal	Full	overt	overt	Canonical

sitting in Spec, CP and the copula hosted in the Head immediately above it create a POS. If the pole sitting in Spec, CP is overt, it has the ability to undergo Word Formation (via Abstract Incorporation) with the pole - copula (the Head immediately above it). Word Formation, in fact, by definition requires the pole sitting in Spec, CP to be overt. Word Formation *representationally* neutralizes the POS. As a consequence, the RC is left in situ (postnominal) which results in Canonical Predication. In the exceptional case, covert pro located in C is licensed by the overt pole sitting in Spec, CP (see above).
If, on the other hand, the pole sitting in Spec, CP is covert, it does not have the ability to undergo Word Formation (via Abstract Incorporation) with the pole - copula. Word Formation, in fact, by definition requires the pole sitting in Spec, CP to be overt. Word Formation cannot neutralize the POS *representationally*. The onyl alternative left is for Syntax to neutralize it *derivationally*, via movement. As a consequence, the pole sitting in Spec, CP raises to a higher position (Spec, FP) as a part of the larger chunk CP. This results in Inverted Predication. Covert pro located in C is licensed by the overt Head F.
The account based on Dynamic Antisymmetry, in particular, concentrates on the overt vs.

covert realization of the Subject relative pronoun sitting in Spec, CP and on the role that Word Formation (through Abstract Incorporation) plays in blocking Predicate Inversion. Under this approach, pro-licensing is a (beneficial) side-effect of neutralization of a POS, be it representational or derivational.

It ensues that trigger of movement is neutralization of a POS, not pro-licensing, contrary to what den Dikken (2006) claims. Dynamic Antisymmetry in this case provides a principled explanation of movement that seems to be superior to den Dikken (2006)'s one because the former is able to derive not only Word Formation but also deletion of English *wh*-element in English RCs (*his girl friend, a modest person, laughs about that, the book I bought*) left unexplained in the latter. The more or less traditional notions of Canonical Predication , (attributive / comparative) Inverted Predication, Reinstated Word Order, copula, Word Formation and reanalysis in the next Chapters are to be understood in the technical sense assigned to them in this Chapter.

Part III

Copulae and Classifiers in the Arabic NP

Chapter 6

Phrase structure

6.1 Aims and introduction

Chapter 3 has offered diachronic and typological evidence in favor of the hypothesis that the following modification structures found in the Arabic NP:

- Noun + Relative Clause (RC)
- Noun + (Extended) 'Adjective'[1] (Adj)
- Noun + Complement of Specification (or annexed noun: Axd)

are to be assigned one and the same syntactic syntactic structure, namely:

(453) Arabic modification = Noun + Copula + Non-restrictive RC

This hypothesis is meant to provide a more accurate formulation of Pennacchietti (1968)'s idea, summarized in the following as (454):

(454) Arabic modification = Noun + Copula + RC, Adj, Axd
Copula = *nota relationis, designationis, genitivi*

It has been also observed at the end of Chapter 3 that syntactic considerations threaten viability of the unified hypothesis in (453). The syntactic structure of the Arabic modification structure, as well as the categorial status of some constituents involved in it, in fact, are not totally clear and therefore (453) is in need of more syntactic evidence, especially of diagnostic tools assessing correctness of syntactic analysis.
In greater detail, the interpretive problems related to the syntax of (453) have been summarized at the end of Chapter 3 as follows:

(I) Can some diagnostics be developed to assign the 'adjectival article' *l-* and the relative marker *llaḏī, illi* a more precise categorial status?

(II) Which is the exact syntactic position of the subject covert pronoun posited by Arab Grammarians in the '(extended) adjective'- structure?

[1] Recall from Chapter 3 that the Arabic adjective is actually an apposition.

(III) Which is the exact categorial status of the copula usually designated as Case (Nyberg, 1920), if it cannot be taken to be of pronominal origin?

(IV) How can the allomorphy NOM / ACC, OBL be explained?

(V) Where are the expected quantifier in the three term system (160), (161) (see Section 1.1.3)?

(VI) Which is the singulative counterpart of the SOUND PL-ending *-na*?

(VII) How can the Depluralization Hypothesis (145) be integrated within both the Agglutination Hypothesis and the Classifier Hypothesis?

(VIII) Why does the noun associated with the copula *-n* stand alone under certain circumstances (Demarcative Hypothesis), if the copula by definition requires a modifier - predicate (Copula Hypothesis)?

(IX) Why and how is *l-* subsequently used outside the modification-structure and found at the left of the noun?

This Chapter as well as Ch. 7 present diagnostics for identifying the categorial status of the Arabic relative marker, 'Case' and *tanwīn*.
They also aim at describing the phrase structure of the Arabic modification structure in antisymmetric / cartographic terms. In so doing, they will answer the questions (I) - (IX).
Finally, Ch. 8 interprets the phrase structure of the Arabic modification structure motivated in the previous Chapters as the result of transformations (movement), invoking Dynamic Antisymmetry as their driving force.
The minimalist Antisymmetry entails looking at (morpho)syntax of the Arabic NP through the prism of agglutination, especially if a cartographic and / or dynamic version of Antisymmetry is adopted (see beginning of Ch. 4, end of Section 4.5.3, Section 5.3.3 and Fig. 5.3). This amounts to saying that Arabic (morpho)syntax is more fine-grained than traditionally assumed, a perspective strictly intermingled with Giuliano Lancioni's (pc) and Karel Petracek's contention that Arabic in the main falls into an agglutinating type:

> "La structure de l'arabe apparaît maintenant comme agglutinative plus qu'auparavant et a certains égards plus proche des structures des langues africaines."
> (Petracek, 1981, p. 172)

In particular, insofar as the Arabic adnominal markers (Case, tanwīn etc...) are concerned, Antisymmetry is totally in line with the Agglutination Hypothesis pursued here (see e.g. Section 3.3.1).

6.2 Reducing the typology of the Arabic RC

The manifestations of the modification structures featuring a RC or an Adj in Arabic are a great many ones and give rise to a scenario that is not heartening (see Tables 3.24, 3.13, 3.14). This is much more so, if one considers that the data concerning the RC and the Adj set forth in Chapter 3 are but a fragment of the rich linguistic variation characterizing Arabic. However, the sample is significantly reduced if two diachronic data are taken into account:

- adnominal markers belong *either* to the Amorite *or* to the Pre-semitic parastrate (Language Contact Hypothesis: Ch. 3)

- opposition definite vs. indefinite is a late development (Specificity Hypothesis: Ch. 3)

As a matter of fact, these two criteria filter out the vast majority of the constructions, and only two of them are left for the analysis: such constructions are *either* Amorite *or* Pre-semitic and are specific in meaning. One way or the other, all the other RC types are a mixture of Amorite *and* Pre-semitic elements (and undergo the definite / indefinite polarization).
The Amorite elements found in the structural makeup of the RC are the bare N (Ch. 2, p. 41), the particle *la* (see (169)), the demonstrative particle *d̲-* (Ch. 2, p. 38).
The pre-Semitic elements found in the same domain are the 'case-endings' suffixed to N (Ch. 2, p. 41) and the prenominal and adjectival 'article' *l-* (Ch. 2, p. 41).
The Amorite and Pre-semitic RC structures under discussion are summarized in Table 6.1, which also reports for expository purposes the CA RC type, namely:

(455) Classical Arabic (= 115)

l-malik-u llad̲ī ya'dilu
the-king-NOM RELM (,) 3M.SG-act.with.justice

'The king (,) which acts with justice'

Under the approach pursued here, this is actually a mixed and late type, but it is illustrated in order to show how misleading would be taking CA as a departure point for the analysis.
It is worth observing that the twofold typology of Arabic RC put forward in this work on the basis of Garbini & Durand (1994)'s Language Contact Hypothesis converges with the twofold typology of *Semitic* RC set by the received view (*Stammbaumtheorie*)[2].
Adopting that approach, what are taken here to be the Amorite and the pre-Semitic RC types are usually referred to as pronominal and construct RCs respectively. Even more remarkable is that the received view conceives these RCs as two separate entities as much as the Language Contact Hypothesis does (albeit for different reasons, since in this latter the *two* distinct types of RC at issue are the reflex of the *two* parastrates / varieties whose mixture forms Arabic): 'in terms of Proto-Semitic relatives, there are differences between pronominal relatives (with *llad̲ī*) and construct relatives. These differences seem to indicate that these relatives are nor related to each other (i.e., did not develop one of the other)' (Na'ama Pat-el, pc). The types depicted

Table 6.1: Arabic Relative Clause and Adjective: revisited typology

Type	Structure	Parastrate	Control set	Specificity	Variety	Example
l Noun Relm Agr RC	l-N-u llad̲ī RC	Amorite	la-, d̲-		OA, CA	(455)
		Pre-sem.	-u			
l Noun RC	l-N-u RC	Pre-sem.	-u, l-	X	OA	(500)
Noun Relm RC	N illi RC	Amorite	bare N, la- (d̲-)	X	Lev. Ar.	(456)

in this Table are discussed in greater detail in the next Sections.

[2]Thanks to Na'ama Pat-el for helpful discussion about how the Semitic RC is interpreted within a more traditional diachronic scenario.

6.3 The Amorite Relative Clause Type

This type of RC is exemplified by (122) here repeated as (456) (see Ch. 3, p. 117 for the specific reading of this construction):

(456) Levantine Arabic (Ouhalla, 2004, p. 291)

baTT-it illi ʔakal-naa-ha
duck-*at* RELM ate-we-her

'a / the duck we ate'

The core of this RC type is the relative marker *illi*, which Barth (1913, p. 159) derives from a form *illaḏī* via dropping of *ḏ*. Although *illaḏī* is very likely to have been a peripheral form in usage from OA onwards (Vicente, 2006), it is still attested in a few old and modern varieties of Arabic (see immediately below) and is the original form of the Amorite RC marker.
Very low usage of *illaḏī* and deletion of *ḏ* in it are intuitively on a par with the very low frequency of CA roots displaying adjacent *l*, *ḏ*: a careful look at the tripartite list of CA roots compiled by Greenberg (1950) shows that only 9 CA roots featuring adjacent *l*, *ḏ* are documented out of 3375 CA roots.
Poor co-occurrence between *l* and *ḏ* depends on a phonological constraint first observed by Greenberg (1950) - though Greenberg does not apply such a constraint to the phonemes under discussion here. This is the tendency for Arabic to avoid the clustering of two adjacent homorganic consonant irrespective of their sequencing and of intervening vowels, because of the difficulty in pronunciation (Greenberg, 1950, pp. 162-163).
Ḏ and *l* are homorganic, if one considers that not only *ḏ* is interdental, as standardly assumed, but also *l*: according to Al-Qahtani (2000) both the so-called MURAQQAQ ('light') *l* of OA described by early Arab Grammarians and the present-day *l* of modern Colloquial Varieties are interdental sounds.
The common articulation point of *l* and *ḏ* results in the following constraint:

(457) *l(V)ḏ, *ḏ(V)l

It follows that the phonological change from *illaḏī* to *illī* via *ḏ*-deletion hypothesized by Barth (1913) is empirically grounded along the lines of (457), *pace* Rabin (1951, p. 165), whose claim that 'no such process is proved for Arabic' is contrary to the empirical data.
Accordingly, the following phonological rule of *ḏ*-deletion must have taken place in OA:

Phonological Rule 6.1 (*ḏ*-deletion)**.**
ḏ ⟶ *0 /l(V)_/*

The form *illaḏī* occurs in the Amorite RC type in Jewish Baghdadi Arabic, a Sedentary variety spoken by the Jewish community of Baghdad (Owens, 2006, p. 144ff. and references therein):

(458) Jewish Baghdadi Arabic (Pat-el, 2009, p. 34)

bạḥur ellaḏi qa-yezzawwaj
young.man RELM PROG-gets.married

'a / the young man who is getting married'

A cognate form *llad̲ī* also occurred in the Hijazi branch of OA and CA (Rabin, 1951, pp. 154 - 155), although the original RC type in this variety has been expanded through insertion of Pre-semitic material (see Table 6.1). This Hijazi offshoot of the Amorite RC marker displays two interesting properties.

Semantically, according to Gensler (2004), *llad̲ī* in Koran (OA) is associated with non-restrictive meaning: 140 out of 225 occurrences[3].

Formally, *llad̲ī* can be analyzed as a demonstrative base *d̲ī* augmented with the morphemic material *lla-*: *d̲ī* and *llad̲ī*, in fact, are commutable as RC marker not only in Old Hijazi, but also in one of its modern Varieties, namely the Yemenite one (Vicente, 2006, and references therein).

The demonstrative base *d̲ī* is invariable in the modern varieties it is found in (Vicente, 2006) and so is in early stages of OA (Rabin, 1951, p. 155). Remarkably, even in CA, where it is agreed for gender and number with its antecedent, descriptively *llad̲ī* resists case-marking, although 'Case' is normal currency in this variety of Arabic: only DU exhibits case-endings (Garbini & Durand, 1994, pp. 99-100).

As for the morphemic material *lla-*, it has been regarded from Reckendorf (1895, p. 601), Brockelmann (1910, p. 123) and Barth (1913) as made of the article *l-* and the asseverative particle *la*. However, comparative evidence argues against this interpretation: in the Hebrew equivalent of *llad̲ī*, namely *hallāzeh*, *lla-* is a primitive entity (Rabin, 1951, p. 155).

For example, it could be observed that in Hebrew *lla-* cannot be decomposed into *l-* and the asseverative particle *la* because in this language there occurs no morpheme *l-*, even less working as an article. Therefore *lla-* is better understood as the asseverative particle *la* having undergone gemination (for reasons that will become clear in Section 6.3.3). To summarize, the internal structure of the Amorite RC marker is as follows:

(459) (l)la + d̲ī$_{CASELESS}$
ASSEVERATIVE *la* + DEM

It is argued in the next Section that morphosyntax of the Amorite RC marker (459) is a function of its semantics (non-restrictiveness).

6.3.1 Diagnostics

Semantics of *llad̲ī* in the Koran points to a copula analysis of the Amorite RC marker and hence to a non-restrictive nature of the RC signalled by it. Recall in fact from Section 5.2.3 and from (328, 329) that the RC marker intervening between a noun and a non-restrictive RC is a copula (*Enrico, and he is the smartest of us all...* etc...).

[3]According to Gensler (2004), in OA restrictive RCs have a totally different structural make up, having at its core the so-called MĀ...MIN CONSTRUCTION (*mā...min* standing for 'what' and 'from', respectively). The OA restrictive RC is exemplified in the following:

(1) Old Arabic (*Koran*, II: 174)

'mā 'anzala -l̤lah-u mina l-kitāb-i
what has.sent.down God-NOM from the-book-GEN

'what of the Book God has sent down on them'

.

It is worth noting in this connection that Arab Grammarians worked out two-fold syntactic diagnostics to detect a copula in primary predication. It ensues that the copula analysis of the Amorite RC marker *llaḏī* can be evidenced syntactically by extending to this latter (secondary predication) the two diagnostic tools Arab Grammarians used in primary predication.
Perhaps a step back is needed to clarify this point. In CA primary predication, the pronoun *huwa* 'he', *hiya* 'she' etc... can also function as a copula. This happens when both the Subject and the Predicate are definite (e.g. a proper name and a noun adorned with the article *l-*), since mere juxtaposition of two definite items could be also interpreted in Arabic as a modification-structure Noun + Adjective (see Section 3.2.2.1). Accordingly, the construction:

(460) Classical Arabic (based on Wright 1896, p. 291)

'aḷḷāh-u l-ḥayy
God-NOM *l*-living

is ambiguous between the reading *God is the living* and the reading *the living God*. This scenario arises as a consequence of the fact that Arabic generally does not signal primary predication by means of an overt copula[4].
To solve this problem, the pronoun *huwa* intervenes between the Subject and the Predicate:

(461) Classical Arabic (based on Wright 1896, p. 291)

[4]In OA, CA, MCV the item *kāna* along with its variants *kan* etc...shows up in the copular sentence when the event takes place in the past and is usually translated as a perfective copula ('was'). This is exemplified in (1):

(1) Moroccan Arabic (Benmamoun, 2000, p. 46)

Omar ma-kan-š mriḍ
Omar *ne*-was-*pas* sick

'Omar was not sick'

Both in the Arab linguistic tradition and in the modern generative paradigm it has been nevertheless observed that in OA, CA and MCV the past copular sentences (displaying *kāna, kan...*) show diagnostic properties different from their present counterparts (see Goldenberg 1988 for OA, CA and Benmamoun 2000 for MCV). For example, in Moroccan Arabic lack of *kan* in the (present) copular sentence allows the predicate to host the sentential negation *ma...š*:

(2) Moroccan Arabic (Benmamoun, 2000, p. 45)

Omar ma-mriḍ-š
Omar *ne*-sick-*pas*

'Omar is not sick'

while presence of *kan* in the (past) copular sentence inhibits the predicate from hosting it:

(3) Moroccan Arabic (Benmamoun, 2000, p. 46)

*Omar ma-mriḍ-š kan
Omar *ne*-sick-*pas* was

'Omar was not sick'

Contrasts like the one just illustrated point to a *non-copular* nature of *kāna, kan...*, which is therefore better understood as a tense marker (Goldenberg 1988 and Benmamoun 2000, Ch. 3). In more familiar terms, a comparison between Arabic and English, if any, would state that *kāna, kan...* of the former is like the *auxiliary was* of the latter.

’aḷḷāh-u l-ḥayy
God-NOM *l*-living

‘the living God’

(462) Classical Arabic (Wright, 1896, p. 291)

’aḷḷāh-u huwa l-ḥayy
God-NOM he *l*-living

‘God is the living’

Now, the complex Subject + Predicate of (462) differs from the complex Noun + Adjective of (461) in that it is more marked, since it features one more item: the pronoun *huwa*. It is the presence of this latter, therefore, that inhibits the ambiguous reading of the complex Subject + Predicate. In this case *huwa* no longer functions as a pronoun (what refers to a previously established or implied noun) but as a copula, because it mediates the relationship between a Subject and a Predicate.
Arab Grammarians were aware of this state of affairs and distinguished between *huwa* etc... acting as a coreferential pronoun (KINĀYAH) and *huwa* etc... acting as copula (FAṢL) (Peled, 2006).
In order to test if *huwa* etc... stands for a copula, Arab Grammarians made use of two diagnostics tools. The former is stated and exemplified in the following (from Peled 2006, 558):

Syntactic Diagnostics 6.1 (Demonstrative / Pronoun as a Copula #1)**.**
The pronoun acting as a copula (ḍamīr al-faṣl) is devoid of Case (‘i’rāb).

(463) Classical Arabic (*Šarḥ al-Mufaṣṣal*, vol. II, *Faṣl ḍamīr al-faṣl ’a l-‘imād*, p. 331)

ẓanan-tu Zayd-an huwa l-qā’im-a
thought-I Zayd-ACC he the-standing-ACC

‘I thought Zayd was the one standing’

In (463) the verb of thinking *ẓanantu* is followed by a sentential object, namely *Zaydan huwa l-qā’ima*. The noun-like constituents of this latter (nouns, adjectives, pronouns) must be marked for ACC. This is actually the case for the Subject *Zayd* and the Predicate *l-qā’im*, but not for *huwa*, whose ACC form (‘him’) is *(’iyyā)-hu*. It follows that *huwa* is not a pronoun, but something else: a copula.
The latter diagnostic tool can be stated as follows (from Peled 2006, 559):

Syntactic Diagnostics 6.2 (Demonstrative / Pronoun as a Copula #2)**.**
The asseverative particle la (lām al-ta’kīd) may be attached to the pronoun acting as a copula (ḍamīr al-faṣl).

(464) is illustrative of *la*-insertion:

(464) Classical Arabic (Peled, 2006, p. 559)

’in kāna Zayd-un la-huwa l-‘āqil-a
indeed was Zayd-NOM *la*-he the-clever-ACC

‘Indeed, Zayd was the intelligent one’

Turning to secondary predication, extending this twofold diagnostics to the Amorite RC marker *ḏī* yields the expected result, since this latter

- is combined with *la* (see (459)), like the copula *huwa* in (464)
- resists case-marking (see (459)), like the copula *huwa* in (463)

It ensues that:

(465) the Amorite RC marker *ḏī* is a former demonstrative base turned into a copula: hence, the non-restrictive reading associated with it.

6.3.2 The minimalist analysis

Development of the Amorite demonstrative base *ḏī* into a copula takes place within a precise syntactic context: the non-restrictive RC. Hence, it is more accurately characterized as a shift from a relative pronoun of demonstrative nature *ḏī* into a copula.
The relative pronoun status of *ḏī* is still documented in Modern South Arabian, where it occurs in the variant *de* (see Rabin 1951, p. 39 for the genetic relationship between OA *ḏī* and Modern South Arabian *de*). In this language, the relative pronoun status of *ḏī / de* can be inferred from the fact that it works as a proximal demonstrative outside the RC. Contrast (466) with (467):

(466) Modern South Arabian (Soqotri) (Shibatani & Bin Makhashen, 2009, p. 12)

rəj de jəʿāzen birhe jədəḥ
man *de* loves children came

'a man who loves children came'

(467) Modern South Arabian (Soqotri) (Shibatani & Bin Makhashen, 2009, p. 13)

de mokšəm
this boy

'this boy'

On the other hand in OA, where *lla-ḏī* is a copula (see previous Section), *ḏī* is not used productively as a proximal demonstrative, this function being generally performed by the bases *ḏā, t-, l-* (Fleisch, 1961, vol. II, Ch. 2) - and CA simply inherits this state of affairs[5]. To recap:

(468) Modern South Arabian: *ḏī* = Dem → Relative Pronoun
OA *ḏī* ≠ Dem → Copula

At this point it is clear *which* is development of the Amorite RC marker *ḏī*, but it is still an unsettled issue *why* and *how* this development applies at all.
The explanation explored in the present work is that:

[5]In OA / CA a demonstrative base *ḏi* highly reminiscent of *ḏī* also occurs. Whether *ḏi* is etymologically connected with *ḏī* or not, is not relevant here because in both the cases it is an isolated and opaque form in the demonstrative paradigm. It is found only in the F SG form *hāḏīhi* and even in this cell of the paradigm the presence of the additional morphemic material *hā...hi* makes it difficult to detect it. Such material develops out of the demonstrative base *h-*, also found in the pronoun *huwa* (Fleisch, 1961, vol. II, pp. 27, 33ff., 57-60).

(469) The Amorite RC marker *d̲ī* turns from a relative pronoun into a copula because it is a clitic-like element found in a non-restrictive RC. Its development takes place via Abstract Incorporation (reanalysis).

This working hypothesis invokes the dynamic antisymmetric framework, especially the phenomenon of Abstract Incorporation dealt with in the previous Chapter:

(470) X + Y →
X Y + Op

Recall that a desirable consequence of this approach, if proved plausible, is that movement of the RC is blocked via Word Formation, which derives its postnominal position.
Abstract Incorporation involves a Head acting as a copula and the relative pronoun sitting in the Spec immediately below it. They create a POS, representationally neutralized via Word Formation:

(471) $[_{R'}$ X $[_{CP}$ Y_i $[_{C'}$ pro_i...

(472) $[_{R'}$ X+Y_i $[_{CP}$ Op_i $[_{C'}$ pro_i...
Reanalysis

In greater detail, it is put forward here that the Amorite demonstrative base *d̲ī* is originally a relative pronoun hosted in the Spec, CP immediately below a null copular Head. This is the scenario defined in Section 5.3.3 as *blind* POS, that is a POS that applies even if one of its poles is covert (see also (396) and Fig. 5.3).
In the normal POS, two overt poles are not fine with AP, because they are unordered in the 'atemporal' Syntax, so that AP is not able to decide their ordering onto the time axis at Spell-Out.
Thus, Syntax puts them together (Word Formation) in order to yield just one pole - this is the representational neutralization of a POS. AP now is fed with only one pole, so that no problem of ordering of the poles on the time axis arises (cp. the property of commutativity in mathematics, which by definition holds for the summands, not for their sum: 4+1, 1+4 vs. 5).
It ensues that a complex of two poles made up of a *covert* and an *overt* pole is not a POS for AP, because this latter only sees overt elements, ie just one pole, something it is able to handle.
But the fusive morphological type in natural languages strongly suggests that a covert and an overt pole *do* create a POS: e.g. in English the PL *feet* diachronically develops out of the POS *feet-*0_{PL} (in its turn from *feet-*i_{PL}, a proper POS: see Section 5.3.3). *Feet-*0_{PL}, in fact, is a word and the word, as just seen, in dynamic antisymmetric terms is no more than a way to neutralize a POS. *Feet-*0_{PL} is therefore the result of the *blind* POS made up of *feet* and 0_{PL}, probably due to a phenomenon of syntactic overgeneration.
Returning to Arabic, *d̲ī* incorporates into the null copular Head in order to neutralize the blind POS they create.
On this view, the relative pronoun *d̲ī* incorporates into the null copula that precedes it along the same lines of the (Middle) English relative pronoun *who* (cp. (396, 405, 406)). This is due to the fact that *d̲ī*, like *who*, is monosyllabic and Dependent-less.
The parallelism is illustrated in (473, 474):

(473) a. 0 + who →
b. 0 who + Op

(474) a. 0 + d̲ī →
b. 0 d̲ī + Op

A more accurate representation of (473, 474) couched within a minimalist framework is (475, 476). For lack of a better designation, the term Conj that has been used so far to refer to the conjunction acting as a copula introducing a non-restrictive RC is extended here to indicate any item playing the same role.

(475) a. [$_{R'}$ 0 [$_{CP}$ who →
b. [$_{R'}$ 0 + who [$_{CP}$ Op

(476) a. [$_{Conj'}$ 0 [$_{CP}$ d̲ī →
b. [$_{Conj'}$ 0 + d̲ī [$_{CP}$ Op

The account in (476) rests on the assumption that an Amorite item at least in origin functions as a relative pronoun following a (null) copula. Such a state of affairs is still documented in two modern Colloquial Varieties.

In Egyptian Arabic there occurs a pronominal base *huwa*, of Amorite origin (see Fig. 2.6 in Ch. 2) and functionally equivalent to *d̲ī*: cp. the caseless *la-huwa* in (464) with the caseless *(l)la-d̲ī* in (459).

Egyptian Arabic *huwa* functions as a relative pronoun in a precise context: as Medjell (2003, p. 543) explicitly states, *if the RC has a non-restrictive reading*. This latter straightforwardly qualifies the RC marker *illi* at least in this case as a Conj copula (see previous Section), to the effect that the relative pronoun *huwa* immediately follows the copula (*illi*). (477) exemplifies this scenario:

(477) Egyptian Arabic (Medjell, 2003, p. 543)

il-king walla malik Harold illi huwa kān malik briṭānya fi-l-wa'ti dā
the-king or king Harold *illi* he was king Britain in-the-time that

'the king or King Harold, who was king of Britain at that time'

The minimalist account developed and defended in Ch. 5 predicts that the postnominal position of the non-restrictive RC (477) is due to Abstract Incorporation of *huwa* into *illi*. On this view *illi*, being hosted in the Head Conj and *huwa*, *qua* relative pronoun sitting in the Spec, CP immediately below it create the POS *illi - huwa*.

Abstract Incorporation, however, applies iff and only if the lowest Head-within-Spec, in this case *huwa*, has clitic-like properties. They are (Section 5.3.3):

- no addition of Dependents
- monosyllabicity

Huwa clearly matches the former criterion. Notwithstanding the appearance, it matches the latter one as well, if one follows Chekayri (2005) in taking the glide of a sequence CVGV (G for Glide) to be absent from its phonological representation (*hua*) and to be the result of a process of GLIDE CREATION triggered by the vowel following it: *hua* > *huwa*.

The conditions for application of Abstract Incorporation are therefore met and the POS *illi - huwa* is neutralized because Abstract Incorporation yields a word-like unit, which is no more than representational neutralization of a POS (see (391)). In line with (476), the word-like unit at issue is a new composite Conj copula, including the former pronoun *huwa*

(478) a. $[_{\text{Conj'}}$ illi $[_{\text{CP}}$ <u>huwa</u> →
b. $[_{\text{Conj'}}$ illi + <u>huwa</u> $[_{\text{CP}}$ Op

The word-like nature of *illi huwa* (actually, *illi-huwa*) appears more transparent from the *same* syntactic complex Conj plus *huwa* instantating the Conj *wa-* instead of *illi*: it is in fact a well known fact among Arabic dialectologists that in Egyptian Arabic *wa-* combined with *huwa* yields *wa-hu*.
It is precisely the representational Word Formation (triggered by *huwa*) that makes application of its derivational counterpart, ie movement, superfluous, thus blocking it, so that the entire RC of (477) fulfills a postnominal position.
To summarize: the POS *illi - huwa* and its neutralization has two (related) outcomes:

- it turns the Amorite relative pronoun *huwa* into a Conj copula (478)
- it is responsible for the postnominal position of the non-restrictive RC of (477)

The dynamic antisymmetric account of (477) thus corroborates the claim that the Amorite RC copula *<u>d</u>ī* is originally a relative pronoun, which only later has turned into a copula (via Abstract Incorporation).
Last but not least, recall from Section 5.2.1 that a strict application of Baker (1988)'s Abstract Incorporation requires that a given Head-within-Spec incorporates into the Head immediately above it iff and only if they are coindexed for some feature: in Arabic, this is the syntactic feature Demonstrative.
Consider the OA POS *0 - <u>d</u>ī*: 0 is a Conj copula occurring in the Amorite RC type, but this is not the only Conj found in this parastrate. Fleisch (1961, vol. II, pp. 61-62) reports for OA the Conj (operating at intersentential level) *<u>t</u>umma*, which develops out of an adverb of place having at its core the demonstrative base *<u>t</u>-*. Under the Language Contact Hypothesis, the OA demonstrative system is an innovation of Amorite origin (see Ch. 2 and cp. the interdental articulation of *<u>t</u>-* also found in *<u>d</u>(ī)*).
On these grounds, it seems safe to claim that the Amorite null Conj gets it Demonstrative feature by virtue of a *rapport associatif* with the Amorite Conj *<u>t</u>umma*, to the effect that it is coindexed with *<u>d</u>ī* for this feature.
Turning to the POS *illi - huwa*, *illi* is the Egyptian Arabic offshoot of *lla<u>d</u>ī*, hence of demonstrative origin, and so is the pronoun *huwa*, built on the demonstrative base *h-* (see fn. 5).
Another interesting case of study is the behavior of the Amorite base *h-* in an RC type found in Nedroma Arabic, a Colloquial Sedentary Variety spoken in the city of Nedroma (Algeria) and described in Guella (2010). In this language, the Amorite pronoun *h-* found in the RC does not turn into a Conj copula, but nonetheless exhibits its semantic properties.
The Nedroma Arabic RC is characterized by fronting of the syntactic constituent containing the relative pronoun *h-* (or *āš*). This is perspicuous from the behavior of such constituents when playing a semantic role other than the Subject, as exemplified in (480 - 482)[6]. Thus, for example, in (480) the sentence *hlīb-ha šerbet* 'she drank her milk' has the OV order, but it is widely agreed in literature (Lancioni 1996 is but a sample) that in Arabic the basic word order is VO, e.g. *šerbet hlīb-ha* 'id.' (cp. (488) in the next Section), to the effect that OV obtains from VO via fronting of the Object (say, *hlīb-ha* 'her milk') over the verb (*šerbet*):

[6]This RC type coexists with another RC type where fronting does not apply, accounted for in Section 6.3.3.

(479) Nedroma Arabic (Guella, 2010, p. 103)

l-šibānija elli ražel-ha f-el-ħabs
the-old.woman *elli* man-her in-the-prison

'the old woman whose husband is in prison'

(480) Nedroma Arabic (Guella, 2010, p. 103)

l-mra elli ħlīb-ha šerbet
the-woman *elli* milk-her drank

'the woman who drank her milk'

(481) Nedroma Arabic (Guella, 2010, p. 104)

l-mtirqae elli b-āš semmer el-kursi
the-hammer *elli* with-which nailed the-chair

'the hammer with which he nailed the chair'

Similar to (481) is the following example, from Moroccan Arabic[7]:

(482) Moroccan Arabic (Brustad, 2000, p. 107, fn. 14)

el-'ām elli f-āš mšīt
the year *elli* in-which left

'the year in which I left'

But fronting of a given sytactic constituent of the RC is strongly associated with its non-restrictive meaning, since Rebuschi (2003, p. 468, fn. 24) observes that the Basque RC, where the *entire embedded sentence is fronted over the C*, in the unmarked case tends to be *non-restrictive* - in order to make it restrictive, the antecedent must be combined with a further modifier ('I add an attributive Adj(P) here [= (484) below] because the lighter the "articled" nominal expression is, the more likely it is for the right-adjoined SFR [= RC] to be interpreted as non-restrictive'). On this basis, (483) is preferably interpreted as non-restrictive, (484) as restrictive, although in principle both the interpretations are available (in these examples, *(e)n* is the C equivalent to English *that* and the Subject of the RC in (484) is a covert Op, as explicitly stated by Rebuschi 2003, p. 458):

(483) Basque (Rebuschi, 2003, p. 468)

zu, egi-a daki-zu-n-a
you truth-SG know-you-C-SG

'you, who know the truth'

(484) Basque (Rebuschi, 2003, p. 468)

gizon zaharr-a egi-a daki-en-a
man old-SG truth-SG know-C-SG

'the old man who knows the truth'

[7]See Section 6.3.3 for more discussion on this Moroccan RC type.

The very occurrence of fronting in Nedroma Arabic RC thus provides a twofold clue. First, it points to an (original) non-restrictive nature of the RC, to the effect that the RC marker *elli* is a Conj, because under the approach adopted here Conj is the syntactic correlate of the RC non-restrictive meaning (see previous Section). Second, although in Nedroma Arabic the Amorite item *h-* does not evolve into a Conj copula via Abstract Incorporation, nonetheless the former intuitively behaves at least in part as if it were the latter, because by undergoing fronting it helps the Conj copula in the function of conveying non-restrictiveness.

Staying with the lack of development of the relative pronoun *h-* into a Conj copula a little longer, it should be remarked that this is totally expected under a dynamic antisymmetric perspective.

The syntactic constituent immediately below the Conj *elli*, in fact, is not *h-* but the syntactic constituent *h-* is tacked onto (e.g. *ražel-ha* of (479)) and the syntactic constituent at issue doubtlessly does not have clitic-like properties: it is nor monosyllabic neither Dependent-less. This means that no representational neutralization of a POS (Word Formation) via Abstract Incorporation is needed and accordingly no development of *h-* into a Conj copula can take place. Notice however that the postnominal position of this RC type characterized by fronting can be still derived resorting to Dynamic Antisymmetry.

Given that the syntactic constituent immediately below the Conj *elli* is not a clitic and no POS arises, and given that movement is triggered by the need to (derivationally) neutralize a POS, the conditions for application of movement are not met and thus the RC type in (479 - 482) remains postnominal.

Returning on the (semantic) Conj-like behavior of the relative pronoun *h-* in Nedroma Arabic, it lends empirical support in favor of the hypothesis entertained here that the Amorite RC marker tends to evolve from a pronoun into a Conj copula. Therefore Egyptian Arabic and Nedroma Arabic, on the whole, constitute convincing evidence in favor of the Abstract Incorporation analysis within a dynamic antisymmetric framework, as stated in (469) and schematized in (474), or more precisely, in (476).

Sticking to the minimalist framework, the traditional notion of prefixation of *(l)la-* to the copula *(0)d̲ī* of (476) is still to be translated in more accurate structural and syntactic terms. Following Zanuttini (1997, pp. 22-25)'s antisymmetric / cartographic approach, the asseverative particle is taken to be the Head (notationed as Σ by Zanuttini, and here as Assev) of a functional projection merged immediately above that hosting the Subject and the Predicate. In this light, the Subject subsequently raises to Spec, AssevP in order to yield the expected word order. Hence, (485) is the syntactic structure that this work proposes for the Amorite RC type featuring the copula *(l)la-d̲ī*:

(485) ... $[_{\text{AssevP}}$ NP $[_{\text{Assev}'}$ (l)la- $[_{\text{ConjP}}$ t_{NP} $[_{\text{Conj}'}$ 0 + d̲ī $[_{\text{CP}}$ Op ...

Last but not least an overall look to the just set forth Nedroma Arabic data as well as to (477) clearly demonstrate that:

(486) In the Arabic non-restrictive RC the syntactic constituent immediately following the Conj copula is either a relative pronoun or contains it.

6.3.3 Further diagnostics and developments

There are at least two aspects to (485). First, movement of NP from Spec, ConjP to Spec, AsserP is to be justified in terms of Dynamic Antisymmetry. This issue will be dealt with in Ch. 8. Second, implicit in (485) is the structural property that either the relative pronoun or the constituent containing it must be able to immediately follow (the local offshoot of) *(l)la-d̲ī*. Actually, this is the scenario observed in the Egyptian and Nedroma Arabic data discussed in the previous Section (cp. (486)) and can be used as a diagnostic criterion to identify a demonstrative / pronoun acting as a Conj copula:

Syntactic Diagnostics 6.3 (Conj copula)**.**

Either the overt relative pronoun or the constituent containing it must be able to immediately follow the demonstrative / pronoun acting as a Conj copula.

While the Nedroma RC type discussed in the previous Section is introduced by an RC marker *elli* matching the just stated criterion, this is not the case for another RC type found in the same Colloquial Variety.
Guella (2010, p. 101) in fact documents for Nedroma Arabic a type of RC clause which he qualifies as having a 'diverse yet semantically equivalent structure' and therefore ascribes linguistic variation between this type of RC and that discussed in the previous Section to dialectal dynamism rather than to stylistic / communicative needs.
The key-property of this construction is that (the constituent containing) the overt relative pronoun *cannot* immediately follow *elli*. This is exemplified by the following sentences, which are identical with those discussed in the previous Section, except for the position of (the constituent containing) the overt relative pronoun:

(487) Nedroma Arabic (Guella, 2010, p. 103)

l-šibānija elli f-el-ẖabs ražel-ha
the-old.woman *elli* in-the-prison man-her

'the old woman whose husband is in prison'

(488) Nedroma Arabic (Guella, 2010, p. 103)

l-mra elli šerbet ẖlīb-ha
the-woman *elli* drank milk-her

'the woman who drank her milk'

(489) Nedroma Arabic (Guella, 2010, p. 104)

l-mtirqae elli semmer bī-ha el-kursi
the-hammer *elli* nailed with-it the-chair

'the hammer with which he nailed the chair'

Broadening the perspective of the analysis, this type of RC is the most common one in the modern Colloquial Varieties, while that discussed in the previous Section seems to be confined to the varieties spoken along the Algerian - Moroccan border (Guella 2010, Brustad 2000, p. 91, 111) - native speakers of Moroccan from other areas, in fact, found (482) strange (Brustad,

2000, p. 107, fn. 14).
Plainly, in the RC type exemplified in (487, 488, 489) the categorial status of *elli* is not that of a Conj copula, because it does not meet the conditions stated in the just mentioned diagnostics. Which is then the categorial status of *elli* in this case? Brustad (2000, Ch. 3) and Mughazy (2006) interestingly remark that in the MCV where this RC type occurs the relative marker *elli* has the ability to introduce a new clause, which points to its nature of complementizer (see also Section 5.2.1). A telling example is from Levantine Arabic:

(490) Levantine Arabic (Brustad, 2000, p. 104)

‘aẓīm illi ’dirt itxalliṣ b-sur‘a
great *illi* were.able finish with-speed

‘it is great that you were able to finish quickly’

This leads to implement the following diagnostics:

Syntactic Diagnostics 6.4 (Complementizer status of *elli, l-*)**.**
Either the overt relative pronoun or the constituent containing it are not able to immediately follow elli, l-.

One may wonder at this point why *elli* turns from a copula signalling a non-restrictive RC (Conj) to a complementizer (C). The answer explored here is that such development is due to influence from the Pre-semitic parastrate.
This is evidenced by the fact that in Nedroma Arabic the Pre-semitic item *l-* (see Section 6.2) carries out the function of an RC marker as an alternative to *elli* under certain circumstances, namely if the predicate of the RC is an adjective, a participle or a noun (Guella, 2010, pp. 107-108)[8].
In other words, replacing *elli* with *l-* in this particular context gives rise to the construction designated in literature as the Extended Adjective (see (133, 134) illustrated in Ch. 3).
A comparison between the Nedroma Arabic RC (491) and the Nedroma Arabic Extended Adjective (493) makes this adamantly clear.
It is precisely the ability for the C *elli* (491), doubtlessly a C, to replace *l-* (493) that lends empirical support to the C-analysis of this latter hinted at in Sībawayhi and Wright (1896); Diem (1998); Kremers (2003) (see p.120): cp. the Bloomfieldian substitution test as a diagnostic tool dealt with in Ch. 4.
In a likewise fashion, in OA *l-* sometimes introduces the RC typically introduced by *llaḏī*, as illustrated in (495): once again, the fact that in this OA piece of data the constituent containing

[8]Recall from Section 3.2.2.1 and fn. 45 in Ch. 3 that Arab Grammarians subsume these three categories under the same label of noun on morphological and syntactic grounds. In short, these three parts of discourse in Arabic often have the same morphological pattern in common and resist agreement. Guella (2010, p. 107) contends that *(e)l-* is a truncated form of *elli*, omophonous with the article *(e)l-*. He adds that this truncated form shares with the article the same phonological properties: e.g. in (493) the relative marker *(e)l-* undergoes assimilation (*(e)ṣ-*), in (494) resyllabification *(l(e)-* (Guella, 2010, p. 107). The strong similarities between the RC marker *(e)l-* and the article *(e)l-* strongly suggest that there is no principled reason to consider them as two separate entities and accordingly they are regarded here as the one and same item *(e)l-*. Additional evidence for this claim is syntactic and is presented in Ch. 7. Finally, notice that the RC marker *(e)l-* in (493, 494) is given a phonological transcription which abstracts away from the just mentioned phenomena of assimilation and resyllabification.

the relative pronoun is not adjacent to *llaḏ̲ī* forces the C-analysis of this latter and therefore the same holds for *l-*.

Finally, it is convenient to remark that in Nedroma Arabic *elli* and *l-* show a remarkable asymmetry in the way they are associated with (the constituent containing) the overt relative pronoun. Both *elli* and *(e)l-* allow it to occur in a non-adjacent position (491, 493), but only *elli* allows it to be right-adjacent (492, 494). In other words, both *elli* and *l-* work as C (491, 493), but only *elli* works as Conj (492, 494), as per the diagnostics stated at the beginning of this Section:

(491) Nedroma Arabic (Guella, 2010, p. 107)

el-šažra elli ṣūfer wrāq-ha
the-tree *elli* yellow leaves-its

'the tree whose leaves are yellow'

(492) Nedroma Arabic (Guella, 2010, p. 107)

el-šažra elli wrāq-ha ṣūfer
the-tree *elli* leaves-its yellow

'the tree whose leaves are yellow'

(493) Nedroma Arabic (Guella, 2010, p. 107)

el-šažra el-ṣūfer wrāq-ha
the-tree *elli* yellow leaves-its

'the tree whose leaves are yellow'

(494) Nedroma Arabic (Guella, 2010, p. 107)

*el-šažra el-wrāq-ha ṣūfer
the-tree *elli* leaves-its yellow

'the tree whose leaves are yellow'

(495) Old Arabic (Wright, 1896, vol. I, p. 269)

mā 'anta bi-l-ḥakam-i l-turḍà ḥukmat-u-hu
not you at-the-judge-OBL *l*-is.approved sentence-NOM-his

'thou are not the judge whose sentence is approved'

This asymmetry demonstrates that the RC type featuring the C *(e)l-* is the original source of the RC type featuring the C *elli*. It seems safe to maintain that the Conj *elli* has been attracted into the pattern of the C *(e)l-* for phonological reasons.

The morph encoding the functional feature Conj, namely *ḏ̲ī* (see previous Section) has been deleted from the morphemic material *elli*, belonging to the Amorite parastrate, because of the difficulty in pronouncing the cluster *l(V)ḏ̲* (see (457)). Thus Conj is no longer an available interpretation for *elli*. Accordingly, the remaining Amorite morphemic material *elli* is reinterpreted as C by analogy with the C *l-* belonging to the pre-Semitic parastrate. This analogical process might have plausibly been catalyzed by the phonological similarity between the morph *l-* and the morph(s) *elli*.

In short:

(496) The Amorite *elli* is a Conj reinterpreted as C because of dropping of *d̲ī* and phonological analogy of the remaining string *elli* with the pre-Semitic C *l-*

Further evidence to assume that *elli* develops from Conj to C is phonological and lies in the diachronical process of gemination of *la(-d̲ī)* into *lla(-d̲ī)*. The context in which this kind of gemination is found is a sequence of two words (Noun + *la(-d̲ī)*) in which the word-initial consonant of the second word strenghtens to a geminate (Noun + *l̲la(-d̲ī)*). But this is precisely the phenomenon known in literature as RADDOPPIAMENTO SINTATTICO (Passino, 2009, p. 7):

Phonological Rule 6.2 (Raddoppiamento sintattico)**.**

$C \longrightarrow CC \,/\# _/$

Scheer (2008) and Passino (2009)[9] after him explain insertion of a (geminated) consonant typical of *raddoppiamento sintattico* as a strategy of the module A-P to signal what is a phase in the module Syntax: simply put, gemination marks a phase. Recall from Section 5.2.4 that a phase is a complete chunk of semantico-syntactic information and that both DP and CP are phases: distributionally, an embedded clause (CP) has the ability to replace a DP (e.g. *that John proved the theorem* is equivalent to *John's proof of the theorem*).
Scheer (2008)'s claim is corroborated by the Italian data (286) discussed in Section 5.2.1 and replicated below as (497):

(497) Italian (Julien, 2002, p. 38)

dello studente
of.the student

'of the student'

In (497) the second word involved in *raddoppiamento sintattico*, in fact, is *lo*, which is clearly the head of the phase DP because it is an article (see Section 5.2.4). On these grounds, *la(-d̲ī)* turns into *lla(-d̲ī)* because it is reinterpreted as the head of the phase CP (see (496) immediately above).
In minimalist terms, the pre-Semitic RC marker *l-* and the Amorite RC marker *elli* modelled after it by analogy (491, 493), *qua* complementizers, are represented as follows:

(498) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ Op $[_{\text{C}'}$ l-...

(499) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ Op $[_{\text{C}'}$ elli...

Notice that in both (498 , 499) a *covert* relative pronoun (Op) has been posited in the syntactic position immediately following Conj (Spec, CP), although nothing in the foregoing analysis proves this. In the Amorite RC type *not* influenced by the pre-Semitic parastrate (485), in fact, Op is motivated as the result of a process of Abstract Incorporation affecting the overt relative pronoun *d̲ī* (474, 476), but no such process can be ascribed to the pre-Semitic RC type featuring *l-* (498) and to the Amorite RC type influenced by it (499), because the overt relative pronoun *d̲ī* does not belong to the pre-Semitic parastrate.
Yet, the presence of Op in Spec, CP seems to be intuitively necessary, since it has to do with the inability for (the constituent containing) the *overt* relative pronoun to fulfill the same syntactic position. This insight is captured in minimalist terms in Section 6.4.2.

[9] Thanks to Giovanna Marotta for pointing me out the syntactic implications of the phonological phenomenon of gemination.

6.4 The Pre-semitic Relative Clause Type

This type of RC is well documented in the Koran. It occurs in at least two variants, depending on the 'case-ending' marking the antecedent (Jones, 2005, p. 146). They are what is traditionally referred to as NOM, OBL (on which, see Ch. 3). Apart from this, the two manifestations of the pre-Semitic RC are identical, as shown by (127), here repeated as (500) and by (501):

(500) Old Arabic (*Koran*, 2:248)

'an ya'tiy-a-kumu l-tābūt-u fī-hi sakīn-at-u-n
that comes-SBJV-you the-ark-NOM in-it peace.of.reassurance-F-NOM-GM

'that the Ark will come to you, in it a Shechina from your Lord'

(501) Old Arabic (*Koran*, 62:5)

ka-maṯal-i l-ḥimār-i yaḥmilu 'asfār-a-n
as-likeness-OBL the-donkey-OBL carries books-ACC

'as the likeness of an ass carrying books'

Selected translations of Koran, namely Khan (1996); Ali (1934); Arberry (1955)[10] assign (500) a non-restrictive reading and furthermore Khan (1996) assigns this sentence also a specific reading, since he renders it into English as 'that there shall come to you At-Tabut (a wooden box), wherein is Sakinah (peace and reassurance) from your Lord ' (cp. the article 'a').
As for (501), Khan (1996); Ali (1934); Arberry (1955) interpret its antecedent as specific (cp. the article 'a') but do not give the RC a non-restrictive reading. This latter, however, is given by Arab Grammarians the same semantic and structural description of a non-restrictive RC (see next Section).
The pre-Semitic RC type (500, 501) is strikingly similar to the Somali RC discussed in the previous Chapter and exemplified below as (502):

(502) Somali (Frascarelli & Puglielli, 2006, p. 320)

Xasan gabar-tii oo gurig-a u socota buu arkay
Xasan girl-ANAPHORIC ARTICLE and house-the to going.is FOCUS MARKER saw

'Xasan met that (specific) girl, who is now going to home'

in two main respects:

- the antecedent is adorned with an article
- the RC itself is reduced (ie features a covert pronoun)

To this, one may add that specificity is a key-ingredient not only of the pre-Semitic RC type, but also of the Somali RC, as shown by Frascarelli & Puglielli (2006)'s translation of (502): *...that (specific) girl...*
The two RCs *prima facie* differ in one crucial detail: while in the pre-Semitic RC (500) a

[10] All these translations have been selected for *The Quranic Arabic Corpus* compiled at the School of Computing, University of Leeds, under the guidance of Kais Dukes and are available at the following website: *http* : *//corpus.quran.com/*.

'case-ending' intervenes between the Subject NP and the modifier / Predicate, in the Somali RC (502) it is the conjunction *oo* 'and' that fulfills the same syntactic position.
Because of this very position, Pennacchietti (1968) interprets the conjunction *oo* of (502) as a copula. As for the 'case-ending' of (500), having the same distribution of *oo*, he does not say anything about it: Pennacchietti (1968), in fact, does not take into consideration the pre-Semitic RC type represented in (500). Nevertheless, the strict structural parallelism holding between the pre-Semitic (500) and the Somali (502) non-restrictive RC strongly suggests that the pre-Semitic 'case-ending' is no more than the Somali copula *oo* - actually, a Conj copula: see Ch. 5.
This hypothesis, which partly revives Nyberg (1920)'s idea that Arabic case arises from a copular element, rests upon a crucial *distributional* property of it: in OA, 'Case' only occurs between a Subject NP and its modifier / Predicate (Owens, 2006, Ch. 3), ie in a predicational environment. *Syntactic* evidence for this claim is provided in the next Section.

6.4.1 Diagnostics and developments

In Somali, the conjunction acting as a copula has a well-defined semantic property: it works both as a non-restrictive RC and as an adverbial clause (see fn. 12). This is taken here to be useful diagnostics to detect a conjunction acting as a copula:

Syntactic Diagnostics 6.5 (Conjunction as a Copula)**.**
The conjunction acting as a copula has both a non-restrictive and an adverbial reading.

The Late Grammarian al-Zamaxšarī (d. 538 / 1144) in analyzing (501) in his Koranic commentary *al-Kašsāf* explicitly states that both these readings are available for the clause *yaḥmilu 'asfāran*, a phenomenon also observed by Brustad (2000, p. 98, fn. 8):

> "...If you asked: which is the syntactic function of *yaḥmilu ['asfār-a-n]*? I would say: either it functions as an adverbial clause (*ḥāl*), as if it were marked for ACC or as a relative clause (*waṣf*) as if it were marked for [the same case of its antecedent, namely] OBL...".
> (*al-Kašsāf 'an ḥaqā'iq ġawāmiḍ al-tanzīl wa-'uyūn al-'aqāwīl fī wujūh al-ta'wīl*, vol. II, p. 1482)[11]

Al-Zamaxšarī's remarks empirically ground the hypothesis that the OA / CA 'Case' is in reality a conjunction.
It is worth noting in this connection that Arab Grammarians describe the adverbial clause, or ḤĀL, as having (at least) the three following properties (see e.g. Hashemi 2006a):

- it conveys redundant information (FAḌLAH)
- it must contain a pronoun coreferential with the noun preceding it
- this latter is either adorned with the article (definite) or followed by extra material (specific)

[11] '*...fa-'in qulta yaḥmilu mā maḥall-u-hu qultu-l-naṣb-u 'alà l-ḥāl-i 'aw l-jarr-u 'alà l-waṣf...*'

These are the properties that modern linguistics regards as distinctive of a non-restrictive RC (see fn. 10 in Ch. 5).
In particular, (500) and (501) feature a type of adverbial clause in which the pronoun referring back to its antecedent is covert (see Wright 1896, vol. II, p. 331 for details).
Outside OA and CA, both the non-restrictive and adverbial reading are possible also for a kind of subordinate clause found in Levantine Arabic, exemplified in (503): Feghali (1928, p. 134) gives (503) a non-restrictive reading ('qui est innocent'), but at the same time defines it as an adverbial clause (*ḥāl*).
In this construction, the subordinate clause modifies a noun preceding it, and there occurs between them a functional item *w-* that cannot be taken to be a 'case-ending' because this Colloquial Variety of Arabic totally lacks it.
Following the diagnostics proposed in this Section, the ability for *w-* to get both the non-restrictive and adverbial reading qualifies it as a conjunction acting as a copula, as much as OA / CA 'Case' does.
However, this construction differs from its OA / CA cognate in two major respects.
First, the clausal modifier features an obligatorily overt pronoun referring back to the modifiee, phonologically realized as *huwe, hiye* etc... . Second, the conjunction / copula *w-* and the anaphoric pronoun *huwe* can be omitted as a whole:

(503) Levantine Arabic (Feghali, 1928, p. 134)

hnīyâl ẹl-mathûm (w-huwe) bâre
(and-he) innocent

'heureux l'accusé (qui est) innocent'

Remarkably, native speakers of Levantine Arabic identify *w-* of (503) with the conjunction *w-* 'and':

(504) Levantine Arabic (Cowell, 1964, p. 377)

huwwe 'ādami w-'anā b-ḥebb-o
he nice.person and-I love-him

'he is a nice person and I like him'

Lexical considerations therefore confirm the syntactic diagnostics, at least for Levantine Arabic.

6.4.2 The minimalist analysis

The discussion of the pre-Semitic RC type carried out in the previous Section makes it possible to establish a comparison between the Somali non-restrictive RC (502) and the Arabic one (500, 501 ,503). In consequence of this, Frascarelli & Puglielli (2006)'s structural description of the Somali non-restrictive RC should be extended to its Arabic counterpart (503).
Upon closer scrutiny, however, the equivalence is not total. True, Levantine Arabic copula *w-* (503) can be identified with the conjunction *w-* occurring outside the modification structure, as much as Somali *oo* (502) does, nonetheless the former sensibly differs from the latter because it is obligatorily followed by an *overt* anaphoric pronoun *huwwe* and can undergo deletion when associated with this latter.

On the other hand, the OA / CA copula *u, i* (500, 501) like Somali *oo* (502) introduce an RC in which the relative pronoun is covert (Op), but they cannot be identified with any conjunction *w-* found outside the modification structure.
The former problem evaporates as soon as one considers that *huwwe* in (503) is an item belonging to the Amorite parastrate. This is at odds with the RC type under discussion, of pre-Semitic origin. It ensues that *huwwe* might have plausibly been inserted in this pre-semitic RC type under pressure of the Amorite parastrate. This is evidenced by the (pretheoretical) fact that *huwwe*-insertion correlates with deletion of the complex *w-huwwe* as a whole (cp. (503)).
The argument is built as follows. *W-* is a Conj copula, because it signals a non-restrictive RC, while *huwwe* is a pronoun having the same structural properties of the demonstrative base *d̲ī* (see Section 6.3.2) and then behaves exactly as this latter when merged into (503).
The Conj copula *w-* is hosted in the Head Conj and *huwwe*, *qua* relative pronoun, in the Spec, CP immediately below it. Furthermore *huwwe*, being the Levantine Arabic cognate of Egyptian Arabic *huwa* in (477) has clitic like properties, and this results in the POS *w - huwwe*, which is the Levantine Arabic counterpart of the Egyptian Arabic POS *illi - huwa* (see Section 6.3.2).
Word Formation via Abstract Incorporation (reanalysis) manifests itself to break the POS *w - huwwe*. Word Formation can be easily seen not only in Levantine Arabic, where *w-* becomes part of the syllable following it and is never accented (Cowell, 1964, pp. 18, 391), but also in Egyptian Arabic where the equivalent of *w - huwwe* is *wa-hu* (see Section 6.3.2).
As a consequence of Abstract Incorporation, *huwwe* is reinterpreted as (a part of the complex) copula and therefore incorporates into Conj *w-* giving rise to the new copula *w-huwwe*. But the same process of Abstract Incorporation affects *d̲ī*, which after being reinterpreted as a copula incorporates into Conj *0*, yielding *0d̲ī* (see (476)).
This sets a parallelism between the two complex Conj copulae in the speaker's mind, so that when the complex copula *0d̲ī* undergoes deletion because of its adjacency with *(l)la*, so does the complex copula *whuwwe*, although there is no principled phonological reason to do so, deletion of this latter being due to analogy only.
A simplified representation of this analogical process is given in the following, where (505) reproduces (474):

(505) a. 0 + d̲ī →
b. 0 d̲ī + Op

(506) a. w + huwwe →
b. w huwwe + Op

The minimalist account of the same phenomenon can be illustrated along the following lines, where (507) reproduces (476):

(507) a. $[_{\text{Conj'}}$ 0 $[_{\text{CP}}$ d̲ī →
b. $[_{\text{Conj'}}$ 0 + d̲ī $[_{\text{CP}}$ Op

(508) a. $[_{\text{Conj'}}$ w $[_{\text{CP}}$ huwwe →
b. $[_{\text{Conj'}}$ w + huwwe $[_{\text{CP}}$ Op

Then, the process of *whuwwe*-deletion, which can be only motivated invoking the Amorite phenomenon of *(0)d̲ī*-deletion, corroborates the hypothesis that *huwwe* is a spurious item in the Pre-semitic RC type found in Levantine Arabic (503). Notice that Abstract Incorporation

requires that *huwwe* incorporates into *w-* iff and only if they are coindexed for the Demonstrative feature (see above). This seems to be the case, because the POS *w - huwwe*, albeit intrinsically featuring the *non*-demonstrative item *w-*, is clearly modelled after the Amorite POS *0 - d̲ī* (cp. the just discussed phenomenon of *w-huwwe* deletion). Thus the Amorite POS, characterized by coindexing between the Conj and the relative pronoun attracts the POS *w - huwwe* into the same pattern:

(509) Amorite: 0_i - $\underline{d}\bar{\text{i}}_i$ (> Deletion)
Mixed type: w - huwwe$_i$ > w$_i$ - huwwe$_i$ (> Deletion)

The latter problem is partially solved if one considers native speakers's judgments about the vowels of OA / CA. Arab Grammarians, in fact, do not assign the vowel *u* an independent status, but consider it as a (free?) variant of the (semi-)consonant (ḤARF) *w*, designating *u* as the 'small' semi-consonant *w*, or as 'a portion' of the semi-consonant *w* (see Fleisch 1961, vol. I, p. 204-205 and references therein). By the same token, Cowell (1964, p. 9) observes that in Levantine Arabic 'the non-syllabic version of *u* — transcribed '*w*' — is substantially the same sound as an unaccented syllabic *u*'.

In this light, OA / CA 'NOM' *u* is better understood as a variant of *w*, and given that the Levantine Arabic conjunction *w-* is non-syllabic (Cowell, 1964, p. 391), it can be regarded as a vowel-like sound *u*. As a result OA / CA NOM and Levantine Arabic conjunction *w-* are the same thing.

Indirect evidence for this claim comes from Akkadian: in this language NOM and the standard conjunction found outside the modification structure are identical, both of them being phonologically realized as *u* (Moscati et al., 1980, pp. 95, 121).

This account admittedly leaves out of the picture 'case-endings' other than NOM, namely ACC, OBL, whose copular nature will be motivated in Ch. 7.

The picture that emerges from the discussion so far is that the pre-Semitic RC type, both as instantiated in OA / CA (500, 501) and in Levantine Arabic (503) is realigned with the Somali RC (502). In both of them, in fact, the copula is the conjunction 'and' (*w- / oo*, respectively) and both of them feature a *covert* relative pronoun (Op) in the RC.

This parallelism is not surprising in the light of the fact that Somali is a Hamitic language and Hamitic languages, under the approach à la Garbini & Durand (1994) adopted here, preserve pre-Semitic relics (see p. 39).

In particular, the parallelism between the pre-Semitic RC type and the Somali one provides indirect evidence to shed light on an issue left open at the end of Section 6.3.3, namely the presence of Op in the topmost syntactic position of the pre-Semitic RC type. Given that in the Somali RC (502) there is found an Op occupying Spec, CP (Frascarelli & Puglielli, 2006):

(510) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ oo [$_{\text{CP}}$ Op [$_{\text{C}'}$ pro...

the pre-Semitic RC type behaves in the same way:

(511) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ pro...

The Pre-semitic RC type can also feature an overt C *l-* (for its Pre-semitic origin, see the beginning of this Chapter), as shown in the OA type exemplified in (495). In diachrony, this appears to be a type later than (511), as discussed in Section 6.4.3[12]:

[12]C plausibly has the overt manifestation *l-* because it acts as a disambiguating element, meant to avoid that the modification structure noun - RC (secondary predication) be mistaken for a *be*-sentence (primary

(512) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ l-...

Another instantiation of the same type is found in (493), where ConjP has a null realization. The findings of Section 6.4.1 (*w-huwwe* deletion) suggest that this phenomenon is due to influence of the Amorite Parastrate, where ConjP is originally covert (cp. (507)), although this issue will be not further pursued here. It will suffice to say that rather than a pure Pre-semitic type, this is a mixed type because of its Amorite covert Conj:

(513) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ 0 [$_{\text{CP}}$ Op [$_{\text{C}'}$ l-...

The Pre-semitic type (511) if compared with its Amorite counterpart (485) replicated below as (514):

(514) ... [$_{\text{AssevP}}$ NP [$_{\text{Assev}'}$ (l)la- [$_{\text{ConjP}}$ t$_{NP}$ [$_{\text{Conj}'}$ 0 + d̲ī [$_{\text{CP}}$ Op ...

displays a subtle yet important difference: both of them are (originally) non-restrictive RCs, featuring an overt Conj etc...but their Op are not created equal. While Conj and Op of the former are merged into the structure as such, those of the latter are the outcome of Abstract Incorporation:

(515) Pre-semitic: ...[$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op...
Amorite: ...[$_{\text{Conj}'}$ 0$_i$ [$_{\text{CP}}$ d̲ī$_i$... > ...[$_{\text{Conj}'}$ 0$_i$ + d̲ī$_i$ [$_{\text{CP}}$ Op$_i$...

But Abstract Incorporation in the Amorite parastrate is a consequence of the *blind* POS *0 Conj - d̲ī* (see Section 6.3.2). This raises a conundrum: after all, in the Pre-semitic parastrate, exemplified in the OA example (500), the blind POS *w - Op* is found in the same portion of syntactic structure, so why is it not neutralized?

Adopting Dynamic Antisymmetry as a working hypothesis predicts in fact that the blind POS *w - Op* is either neutralized via Abstract Incorporation (representationally) or via movement (derivationally).

Discussion of the Levantine Arabic POS *w - huwwe* plausibly demonstrates that Abstract Incorporation is a viable option for the OA blind POS *w - Op* detectable in (500), but there is a price to pay. In Levantine Arabic, the Conj *w-* is not intrinsically coindexed with *huwwe* for the Demonstrative feature needed by Abstract Incorporation, and gets it only via analogy with the Amorite POS *0 - d̲ī* (see (509)).

This implies that in the Pre-semitic parastrate w and Op are *not* coindexed with Op for the Demonstrative feature and therefore Abstract Incorporation cannot apply if the Amorite blind POS *0 - d̲ī* does not exert its analogical pressure on it.

The only alternative left, in principle, for the pure Pre-semitic parastrate is solving the blind POS *w - Op* invoking its derivational neutralization: movement of Op. This twofold scenario is summarized in the following (⇔ indicates the analogical pressure, □ the structural difference between the pure Pre-semitic *w-* and that exposed to Amorite pressure):

(516) a. ...[$_{\text{Conj}'}$ $\boxed{0_i}$ [$_{\text{CP}}$ $\boxed{\text{d̲ī}_i}$... ⇔ ...[$_{\text{Conj}'}$ $\boxed{\text{w}_i}$ [$_{\text{CP}}$ $\boxed{\text{Op}_i}$... > Abstract Inc.
b. ...[$_{\text{Conj}'}$ w [$_{\text{CP}}$ Op... > Movement

More concretely, the Pre-semitic blind POS *w - Op* mimicking the Amorite blind POS *0 - d̲ī* (516.b) leaves the RC in situ, in the postnominal position and hence it is precisely this

predication): see Section 6.4.3 for indirect evidence corroborating this claim.

analogical POS that accounts for the pre-Semitic RC type displaying the Conj *w* (the putative NOM) exemplified in (500).
Intuitively, the 'pure' Pre-semitic POS (516.a) is in some way responsible for the Conj other than *u* (the putative NOM), namely *a, i* (the putative ACC, OBL), exemplified in (501). This idea will be implemented in Ch. 8.
Turning to (512), this representation has the advantage of solving the problems intrinsic to Kremers (2003)'s RC analysis of the Arabic adjective, answering the questions (I), (II).
(512) in fact provides a structural description also of the 'adjectival' type (cp. (493)), and the C status of *l-* in it is not speculative, having been detected via the diagnostics worked out in Section 6.3.3. More generally, this Chapter has provided syntactic diagnostics to identify the categorial status of OA *u / w* ('NOM') and *lla-d̲ī*, which answer not only the question (I), but also (III).
Moreover, (493) empirically grounds the occurrence of Op in Spec, CP, thus anwering (II): Op acts as the real Subject if the RC is an Adjective, and as an expletive if the RC is an Extended Adjective (ie manifests an RC-final NP):

(517) Nedroma Arabic (based on (493))

(el-šažra$_i$ el-) ṣūfer	wrāq-ha	= Op$_{EXPLETIVE}$	Adj NP-PRN$_i$
(el-šažra$_i$ el-) ṣafrae		= Op$_i$	Adj

This scenario adopts the account known in literature as Expletive Hypothesis (see e.g. Lancioni 1996)[13].
However, the Pre-semitic type schematized in (511, 512) should be taken as a first approximation, since the OA / CA conjunction *w* (in traditional terms: *u*, NOM) appears not to be the only item heading Conj: in principle, the 'case-endings' *a, i* too (ACC, OBL) can occur. The syntactic structure in (511, 512) is legitimate iff and only if the alternation NOM, ACC, OBL is explained. Ch. 7 expands on this point. For the time being it will be assumed that *a, i* too (ACC, OBL) are copulae because *in OA* they are *commutable* with *u* (NOM) (see (30), (31) at the end of Section 2.2.2.4), pending a more accurate syntactic (minimalist) representation of these items in Ch. 8.

6.4.3 Deriving the prenominal article

Implicit in the account of *l-* sketched out so far is the assumption that *l-*, being the complementizer of a postnominal RC, is base-generated in a medial position (ie between a noun and the RC).
This is desirable, because Testen (1998) and Pat-el (2009) provide diachronic and phonological evidence in favor of this hypothesis, dubbed in Section 3.3.3 Clitic Hypothesis. In particular, they remark that the phonological form of *l-* points to its medial origin into the syntactic structure: it is a clitic (Wackernagel's law: see Section 3.3.3).
In so doing, however, the present proposal runs into the same difficulties of Testen (1998)'s and Pat-el (2009)'s account. It is not able in fact to derive the instance of *l-* traditionally referred

[13]The MCV / OA Extended Adjective, exemplified in Nedroma Arabic (493) differs from its CA counterpart in that the former exhibits strict agreement, the latter deflected agreement. The Expletive Hypothesis can explain this phenomenon invoking a different lexical composition of the expletive Op: in MCV / OA it corresponds to English *there* (cp. *there are...*), in CA to English *it* (cp. *it is...*).

to as 'prenominal article', which precedes the antecedent of Arabic RC and, more generally, the Arabic nouns: say, *el-šažra* of (517), be it unmodified or followed by the RC *el-ṣūfer wrāq-ha*. Whatever the explanation, the structure *l*-N is consistent with Garbini & Durand (1994, pp. 102-103)'s remark that emergence of *l*- is relatively recent in Semitic, being documented from the I millennium BC onwards[14].

Bringing into comparison *el-šažra* of (517) with its Somali equivalent in (502), ie *gabar-tii* as well as with the Akkadian piece of data in (518) makes this point clear. Hamitic and Akkadian, which under the Language Contact Hypothesis adopted here are conservative of a Pre-semitic state of affairs, in the same syntactic environment display a *post*nominal determiner, rather than *pre*nominal. Pat-el (2009) too, although following the *Stammbaumtheorie*, contends that in Semitic the postnominal position of the determiner is older than the prenominal one.

(518) Akkadian (Pat-el, 2009, p. 20)

kasp-a-m anni-a-m
silver-ACC-*m* this-ACC-*m*

'this silver'

Comparative evidence therefore indicates that the prenominal *l*- found in OA is at least recent as the adjectival *l*-.

The hypothesis put forward here is that this happens because the prenominal article *is* the adjectival article and vice versa, which is in principle desirable because it provides an unified syntactic account of *l*-. The argument is built as follows.

Consider the following construction, found in OA:

(519) Demonstrative + Noun

This is exemplified in (520)[15] and (521), drawn from Sībawayhi's *al-Kitāb* and from *Lisān al-'Arab* (see Section 2.3.5), respectively, and reported by Rabin (1951, pp. 75-76):

(520) Old Arabic (Rabin, 1951, p. 75)

ḏā ṣabāḥ-i-n
that morning-OBL-n

'that morning'

(521) Old Arabic (Rabin, 1951, p. 76)

qabbaḥa ḷlah-u ḏā fā
put.to.shame God-NOM that mouth

'may God put that mouth to shame'

Rabin (1951) also remarks that this pattern is still preserved in Dhofari Arabic, a modern Colloquial Variety spoken in Southern Yemen:

(522) Old Arabic (Rabin, 1951, p. 76)

[14]Recall from Ch. 2 that in Arabic this *innovation* makes use of *archaic* morphemic material, borrowed from the pre-Semitic parastrate

[15]In particular, Rabin (1951, p. 75) observes that the context in which (520) occurs forces a demonstrative reading of *ḏā*, while the same expression in CA has lost its deictic force and simply means 'one morning'.

ḏē mkān
this place

'this place'

Interestingly, already in OA this construction can be ambiguously interpreted either as a modification structure or as a copular sentence featuring a demonstrative. The latter, in fact, shares with the former the same pattern, as shown by (523):

(523) Old Arabic (*Koran*, II: 178)

ḏalika taxfīf-u-n
that lightening

'that is a lightening'

The interpretive ambiguity at issue can be represented as follows[16]:

(524) ḏalika taxfīf = that lightening
ḏalika taxfīf = that is a lightening

Such ambiguity is even stronger if one recalls that Arab Grammarians regard the Demonstrative of the construction (519) as a Subject and the Noun following it as its Predicate (see Section 3.2.2.1 and 5.2.4 for a modern implementation of this idea), as much as they do in analyzing the constituents of the copular sentence (523).
At the same time the ambiguous state of affairs in (524) is highly reminiscent of the situation of interpretive ambiguity discussed in Section 6.3.1 relative to the construction Definite Noun - Definite Noun (460): in both cases it is not possible to distinguish between a primary predication interpretation and a secondary predication one.
It should be remarked at this point that the ambiguity found in (460) is solved by inserting the disambiguating element *huwa*, characterized by two structural properties:

- it is pronominal
- it intervenes between the Subject and its Predicate

The behavior of *huwa* in (460) evidences that the medial *l-* found in the demonstrative construction in (525) and exemplified in (526) is a disambiguating element, inserted in (519) to solve the interpretive ambiguity in (524)[17] - an idea already hinted at in Pat-el (2009, p. 43):

(525) Demonstrative + *l-* + Noun

(526) Old Arabic (*Koran*, II: 253)

tilka l-rusul-u
those *l*-messengers-NOM

'those Messengers'

[16]Recall from Section 3.3.1 that in OA 'case-endings' only intervene between a noun and its modifier, so that a more accurate transcription of *ṣabāḥin*, *taxfīfun* in (520, 523) is *ṣabāḥ*, *taxfīf*.

[17]Perhaps the disambiguating element at issue is to be identified with Frazier's MORPHOLOGICAL CUE, and the situation of interpretive ambiguity between a syntactic structure in secondary predication and its (quasi-)equivalent in primary predication as Frazier's GARDEN PATH EFFECT, although this point will be not further developed here. See Townsend & Bever (2001) and Carroll (2008) for details.

It is important to stress that the just mentioned pattern (525) *diachronically precedes the pattern featuring the prenominal article* schematized below as (527):

(527) *l-* + Noun

This is explicitly stated by Garbini & Durand (1994, p. 103) and Putzu & Ramat (2001, p. 115), which observe that in Phoenician, the oldest of the Amorite languages (Phoenician, Hebrew, Aramaic, Arabic: see Section 2.2.2.2), the prenominal article tends to cooccur with the demonstrative. This claim is even stronger if one considers that what is taken to be the oldest example of the Phoenician article (variety of Byblos, XIII BC) is characterized by the very combination of the article with the demonstrative (Putzu & Ramat, 2001, p. 115):

(528) Phoenician (Putzu & Ramat, 2001, p. 115)

kl mplt h-dtm 'l
all ruins the-this temple

'all the ruins of the temple'

Setting aside the different word order of the sequence demonstrative - article, which will be not investigated here, the OA (/ CA) pattern (526) is identical with the Phoenician one (528) in the constituents making up the syntactic structure: a noun combined with an article *and a demonstrative*.
It is proposed here that actually (527) is (525) having undergone dropping of the demonstrative. This comes as no surprise if one takes into account that the OA / CA demonstrative is an offshoot of the Amorite demonstrative base *d-*, as illustrated in (520, 521, 523 etc...), and that this base is adjacent to the item *l-*. It has been shown in Section 6.3 that such a phonological cluster is not tolerated in Arabic (457) and tends to be eliminated via *d-*deletion. But if in (525) *d-* is dropped, what is left is (527), as schematized in what follows:

(529)

Dem + *l-* + N	> *d-*deletion	> *l-* + N
(Phoen.) OA, CA		OA, CA

The hypothesis in (529) is evidenced by some formal and semantic properties of the demonstrative base *d-* observed in the modern Colloquial Varieties, which often preserve an old state of affairs (see Ch. 3).
As for its meaning, the Moroccan cognate of the base *d-*, namely *(ha-)d*, when found in the pattern (525):

(530) Moroccan Arabic (Harrell, 1962, p. 147)

had le-ktab
this *l*-book

'this / that book (already mentioned or implied in the discourse)'

conveys no deictic (proximal / distal) information and has instead an anaphoric function (Harrell, 1962, p. 147), as much as the Indo-european definite article does. Thus, albeit *(ha-)d* does not exhibit all the semantic and pragmatic peculiarities of the definite article (see Brustad 2000, pp. 115-117 and references therein for more details), it is nonetheless an instance of the so-called ANAPHORIC DEMONSTRATIVE, which crosslinguistically competes with the definite article, as pointed out by Croft (*apud* Brustad 2000, p. 117).

Moroccan Arabic thus makes it clear enough that in the pattern (525) it is the demonstrative *ḏ-* that functions as an article, *not l-*. This forces to posit a (covert) demonstrative *ḏ-* in (527): if, in fact, this pattern has an article-like behavior and *l-* does not encode it, another element must be responsible for it.
Another piece of semantic evidence comes from a property of *l-* already dealt with in Section 3.2.2.1. This is its ability to occur in a context of indefiniteness, which is clearly at odds with its (alleged) function of definite article[18]. This is shown in (52), repeated below as (531):

(531) Moroccan Arabic (Edzard, 2006, p. 189)

wāḥd əl-mra
one the(??)-woman

'a woman'

(531) lends empirical support in favor not only of dismissing the equivalence *l-* = definite article, but also of identifying *ḏ-* with such a marker. A simple substitution test, in fact, shows that in Moroccan Arabic *wāḥd*, the (indefinite) article of (531) can be replaced by *ḏ-*, the (definite) article of (530), in the same syntactic position.
Insofar as the form of the demonstrative is concerned, it is worth noting that in both OA and Najdi Arabic, a conservative modern Bedouin Variety (see Ch. 2), there is documented at least one case where the same piece of information usually conveyed in CA by the pattern *l-N* is conveyed by the pattern *ḏā l-N*. This is the Arabic equivalent of the English *demonstrative* adverb 'now' (= *this* time).
Contrast Najdi Arabic (533) and OA (534) with CA (535). See also (532), which clearly shows that *l-* is not involved at all in expressing the adverbial demonstrative meaning of *now*, so that a (covert) demonstrative must be posited in (535)[19]:

(532) Old Arabic (Fleisch, 1961, vol. II, p. 68)

ta-ḥīna
this-time

'now'

(533) Najdi Arabic (Ingham, 1986, p. 277)

ḏa l-ḥīn
this *l*-time

'now'

(534) Old Arabic (Fleisch, 1961, vol. II, p. 68)

ta l-'āna
this *l*-time

[18]Rather, the function of *l-* in (531) is that of a disambiguating element, in line with the discussion so far: while (531) basically means '*one* X', its counterpart, in which no *l-* surfaces, means '*some* X': see Edzard (2006) for details.

[19]In (532, 534), *ta* is a demonstrative base that in OA alternates with *ḏ*: see Fleisch (1961, vol. II, p. 68). The OA data reported here are drawn from Fleisch (1961), but they can be also found in *Lisān al-'Arab*, s.v. *Ḥ Y N*.

'now'

(535) Classical Arabic Fleisch (1961, vol. II, p. 469)

l-'ānа
l-time

'now'

Thus, OA and Najdi Arabic well document that the pattern *d̲ā l-N* (525) is the phonological representation of the phonetic representation *l-N* (527) and therefore provide evidence enough in favor of the claim that *l-N* (actually, 0_{dha} *l-N*) is a subcase of *d̲ā l-N*.
In this light, the prenominal article is actually a medial article, base-generated between the subject (the demonstrative) and its predicate (the noun), in compliance with the Wackernagel's law.
To conclude this Section, it will be pointed out that interpreting the demonstrative base *d̲-* as an article straightforwardly derives the so called OTIOSE D̲- of OA (documented in the variants *d̲ū, d̲ā, d̲ī*). This is an instance of demonstrative *d̲-* which is (see Lane 1863, vol. III, p. 986, Rabin 1951, p. 76):

- semantically devoid of meaning
- combined with proper names

and according to Arab Grammarians was current in the speech of Qays, a speech belonging to the Bedouin branch (Rabin, 1951, Ch. 1) (see e.g. *Lisān al-'Arab*, p. 1479, s.v. *D̲ W*: *wa-hwa kaṯ̲īrun fī kalāmi qaysin wa-man jāwara-hum* and Rabin 1951, p. 76)[20]. The otiose *d̲-* is illustrated in the following:

(536) Old Arabic Lane (1863, vol. III, p. 986)

'atay-na	d̲ā	yaman
came-1PL	OTIOSE D̲-	Yemen

'We came to Yemen'

This phenomenon, far from being language-specific, is made sense of if one interprets the item *d̲-* as a definite article: indeed, crosslinguistically the just mentioned properties of *d̲-* are in fact typical of the so-called PLEONASTIC ARTICLE (Lyons, 1999, p. 163ff. and references cited there). Familiar examples from Indo-european languages include English *The Hague* and Northern Italian *il Giovanni* for the proper name *Giovanni* etc... where clearly the items *the, il* generally functioning as definite articles in these cases have no semantic import, proper names being definite by their own nature.

6.4.4 The minimalist analysis

The analysis of the (putative) prenominal article made in the previous Section can be translated in minimalist / antisymmetric terms having it that the demonstrative sits in Spec, DP and *l-*

[20] The situation of linguistic contact between the Bedouin and the Hijazi branches might plausibly have favored spread of the 'otiose' *d̲-* from the former to the latter, since Rabin (1951, p. 76) reports that the 'otiose' *d̲-* was also heard in Hijaz.

in its Head, along the lines of Campbell (1996) and Giusti (2002) (see Section 5.2.4). Simply put, at least in structural terms *dā l-N* (525), say *had le-ktab* of (530) is the Arabic equivalent of Modern Greek *afto to vivlio* (333):

(537) $[_{DP}$ had $[_{D'}$ le [ktab ...
$[_{DP}$ afto $[_{D'}$ to [vivlio ...

As for *(0_{dha}) l-N* (527) the null demonstrative (as well as its overt counterpart) can be interpreted in the minimalist framework as an Operator, because it picks out one or more individual entities out from a generic set of entities. Accordingly Moroccan Arabic *le-ktab* etc.. ('the book') is represented as follows:

(538) $[_{DP}$ Op $[_{D'}$ le [ktab ...
$[_{DP}$ Op $[_{D'}$ the [book ...

In this respect, (Moroccan) Arabic is not different from English *the book*, Italian *il libro* etc... in synchrony (see (336) in Section 5.2.4), but *the former sensibly differs from the latter in diachrony*. While English / Italian etc... articles *the, il...* develop out of a demonstrative, possibly via reanalysis (from Spec, DP to its Head: Giusti 2001):

(539) the Arabic prenominal article *l-* arises as a disambiguating element intervening between the demonstrative *d-* and the noun

and

(540) its NP-initial, Indo-european-like position is a deceiving effect of a phonological phenomenon of dropping of the demonstrative *d-* because of the non-euphonic cluster *d - l*

In passing these intermediate conclusions answer the question (IX) The structural analogy between the Arabic article and its Indo-european counterpart in synchrony (*not* in diachrony) can be represented under a minimalist approach as follows:

(541) ... $[_{DP}$ Op $[_{D'}$ l- [NP...

More generally, *l-* heads either DP and CP (when introducing an RC) and this phonological identity is desirable because it is standardly assumed in MP that DP and CP at a deep level of representation are one and the same thing: a phase (see Section 5.2.4):

(542) $\boxed{[_{PhaseP}\ Op\ [_{Phase'}\ l\text{-}}$ $[_{ConjP}$ NP $[_{Conj'}$ $\boxed{[_{PhaseP}\ Op\ [_{Phase'}\ l\text{-}}$ [RC]]]]]]

6.5 Generalizing the copular analysis

It has been seen throughout these Chapters that in the Arabic RC whose Subject is an overt or covert anaphoric pronoun (e.g. Op, *huwa*), this latter sits in the topmost position available, namely Spec, CP. This position is made available by the Head signalling the RC: the complementizer (C), exemplified by *l-* in (512).
It ensues that the RC Predicate obligatorily follows the Subject and the C: it occurs in the portion of the RC lower than CP. From the vantage point of the syntax of Predication, this creates a predicational structure where C intervenes between the RC Subject and the RC Predicate, so that C acts as a copula. Notice that the C *l-*, albeit a copula of pronominal origin,

does not totally follow the twofold diagnostics of Section 6.3.1: one may concede that *l-* resists case-marking, but not that it undergoes *la-*prefixation.
This systematic gap is less problematic than it appears at first sight, if one calls into the play the dichotomy Amorite - Pre-semitic existing in OA: the diagnostics at issue is at work in the Amorite parastrate (it involves only the Amorite items *d̲ī, huwa*) and therefore it does not concern *l-*, belonging to the Pre-semitic parastrate. Also recall that *l-* is a copular *C*, not a copular *Conj*, because of the diagnostics worked out in Section 6.3.3.
However, the previous Section plausibly demonstrates that *l-* manifests itself not only as C, but also as a disambiguating element (D) intervening between a Demonstrative / Subject *d̲-* (later turned into an article) and a Noun / Predicate. This amounts to saying that in this morphosyntactic context too *l-* is a copula.
An overall view of the distribution of both the copular C / D and Conj unveils a 'Russian doll effect': Arabic copulae recursively manifest themselves in the modification structure. The OA (mixed) RC type *l*-N-*u*-*l*-RC[21], exemplified by (543) is illustrative of the Russian doll effect, since the Conj copula, C, D overtly co-occur in it (□ indicates the copula):

(543) Old Arabic (based on (49))

l-bāb-u l-ṣaġīr
the-gate-NOM the-little

'the little gate'

(544) [0_{DEM} ⇐ □l ⇒ [bāb ⇐ □u ⇒ [0_{PRN} ⇐ □l ⇒ ṣaġīr]]]

a. [*Subject:* 0_{DEM} ⇐ □l ⇒ [$_{\text{Noun Phrase}}$ *Predicate:* bāb-u 0_{PRN} l-ṣaġīr]]

b. [*Subject:* bāb ⇐ □u ⇒ [$_{\text{Relative Clause}}$ *Predicate:* 0_{PRN} l-ṣaġīr]]

c. [*Subject:* 0_{PRN} ⇐ □l ⇒ [$_{\text{Noun Phrase}}$ *Predicate:* ṣaġīr]]

6.6 Classifiers in the Arabic NP

A well established analysis of the North Arabic article corroborates the claim made in the previous Section that underlying the Arabic article *l-* are two items, as depicted in (525).
Moscati et al. (1980, p. 112) as well as Rubin (2005, p. 76) and Hasselbach (2007a, p. 20) contend that in North Arabic, the ancestor of CA (see Section 2.2.2.3), the prenominal article *hn-* is a *bimorphemic* constituent made up of the two demonstrative bases *h-, n-*:

(545) North Arabic (MacDonald, 2008, p. 208)

h-n-'lt
the-goddess

'the goddess'

[21]This is a mixed type because the phonological representation of this structure includes the demonstrative base *d̲-*, of Amorite origin: *d̲-l*-N-*u*-*l*-RC. All the other functional items (*l-, -u*) are Pre-semitic.

In particular, the demonstrative base *n-* also develops in OA / CA into the complementizer *'inna / 'anna* (Fleisch, 1961, vol. II, p. 62), which can be assigned this categorial status not only because of its ability to open a clause (see p. 198 for an example) but also because it undergoes *raddoppiamento sintattico* (see end of Section 6.3). The phase behavior (Head, CP) of *n-* in OA / CA therefore points to the same analysis of North Arabic *n-* (Head, DP) preceding the demonstrative *h-* (cp. in the previous Section the behavior of OA / CA *l-* as Head of both the *phases* DP, CP):

(546) $[_{DP}$ h- $[_{D'}$ n- [Noun ...
$[_{DP}$ (d̲-) $[_{D'}$ l- [Noun ...

Hasselbach (2007a, p. 21) observes that the North Arabic *-n* of (545) has the same phonological realization and performs the same function (ie adnominal marking) of the various manifestations of the *tanwīn* attested in OA / CA, MCV and Epigraphic South Arabian.
But while North Arabic *-n* (545) and OA / CA *'inna / 'anna* argue for the equation *-n* - complementizer, turning to OA / CA, MCV and Epigraphic South Arabian leads one to move to a weaker formulation of such a claim, since data from these languages resist a complementizer analysis of *-n*. Consider the following data:

(547) Epigraphic South Arabian (Moscati et al., 1980, p. 99)

ṣlm-n
statue-*n*

'the statue'

(548) Old Arabic / Classical Arabic (Veccia Vaglieri, 1937, p. 102)

rajul-u-n
man-COP-*n*

'one / a man'

(549) Najdi Arabic (Ingham, 1982a, p. 54)

sāyg-i-n
driver-COP-*n*

'one driver '

The *-n* found in (547, 548, 549) show a diagnostic property not detectable in the 'pure' phase *l-* in OA / CA: *-n* can follow the noun it refers to with no additional overt and transparent modifier (adjective, RC, complement of specification), *l-* cannot, to the effect that in such a syntactic context they are *not* commutable. This can be seen from (550),

(550) *rajul-u-l
man-COP-*l*
'the man'

The just outlined scenario strongly suggest that *-n* differs from *l-* in that originally it is *not* a complementizer and only later has been it attracted into the complementizer pattern of *l-*.
In order to clarify the exact nature of *-n* prior to its development into a complementizer,

it seems useful to look at its syntactic position. The item *-n* is postnominal and generally postnominal morphemic material in Arabic is a non-restrictive RC (recall in fact that restrictive RCs, adjectives and appositions too are at least originally non-restrictive RCs: see Table 3.12 in Section 3.2.2.1). This implies the same analysis for *-n*.

Intuitively, the main problem with this hypothesis is that *-n* alone appears to be too a small item to be interpreted in this way. A closer exam of the contexts of occurrence of *-n* however shows that it is part of a more complex modifier, thus solving this problem. The modifier *-n* belongs to is *covert or opaque*.

To begin with, consider postnominal *-n* in CA: it displays, among others, an opposition SG - SOUND PL that at first sight involves the morph *-n* for the SG and the morph *na* for the SOUND PL. Upon closer scrutiny, however, this opposition involves in both the SG and SOUND PL a combination of *-n* and a QUANTIFIER. In the SG, in fact, the morph *-n* is elliptical for the bimorphemic expression *n- wāḥid* '*-n* + one', with the quantifier *wāḥid* showing up when emphasis is put on it (Veccia Vaglieri, 1937, vol. I, p. 102). Remarkably, Ingham reports the same behavior for Najdi Arabic *-n*:

> "...where a noun occurs as the sole constituent of a nominal phrase. In this case it seems that the occurrence of *-in* signals something of the semantic value of 'a certain...', 'a particular', i.e. one single but undefined members of a class, whereas the absence of the marker means purely 'a member of the class in general'"
> (Ingham, 1982a, p. 54)

> "It also signifies to some extent incompleteness in the statement so that a noun with the *-in* ending is taken to imply some specification which has not been mentioned"
> (Ingham, 1994, p. 50)

This is exemplified in (551), (552), where capital italics indicate an elliptical (unpronounced) item:

(551) Classical Arabic (Veccia Vaglieri, 1937, p. 102)

rajul-u-n (wāḥid)
man-COP-*n* (one)

'one / a man'

(552) Najdi Arabic (Ingham, 1982a, p. 54)

sāyg-in WĀḤID
driver-COP-*n* ONE

'one / a driver'

In passing, *-n* (wāḥid) in the SG oscillates between an indefinite and a numeral interpretation, as also shown in the English translation. This is totally expected from a crosslinguistic and typological standpoint and Indo-european languages (Italian *un(o)*, French *un*) are but a sample (Dryer, 2005b) in this regard.

As for the SOUND PL, in OA and to a certain extent in CA it denotes a small set of entities, ranging from 3 to 10 and in this usage (PAUCAL PLURAL) it alternates with a kind of BROKEN PL characterized by the prefix *'a-* (see Section 3.2.2.1), as illustrated in (553a), (553b), respectively:

(553) OA (*Šarḥ al-Mufaṣṣal*, vol. III, *Faṣl jam‘ al-qillah wa-jam‘ al-kaṯrah*, pp. 224-5)

a. *’a-frās*: ‘few horses (from 3 to 10)’

b. *muslim-ū-na*: ‘few Muslims (M) (from 3 to 10)’

In this connection, Barth (1894, pp. 422-423, 440) and Fleisch (1961, vol. I, pp. 416-417) after him contend that the paucal meaning of (553b) is due to the presence of the prefix *’a-*, which they regard as no more than a (grammaticalized) quantifier *few* (‘petit nombre’).

The semantic identity between (553a) and (553b) leads to extend Barth’s / Fleisch’s analysis from the latter to the former, taking the *a* of the SOUND PL ending *-na* (553a) to be the same *’a-* found in the BROKEN PL (553b), simply having undergone deletion of the glottal stop ’ (or HAMZAH), a widespread phenomenon in OA. In this light the sequence *-na* is assigned the phonological representation *-n-’a*.

This move, grounded on the level of meaning, is corroborated on the level of form by the rule of *hamzah*-deletion that will be discussed in Section 7.6: this latter applies in the same phonological context C’V characterizing the phonological representation *-n-’a*. Cp. the OA alternation *(ya)s’a(lu) / (ya)sa(lu)* ‘he asks’[22].

It therefore emerges that in Arabic the opposition SG - SOUND PL actually calls into the picture - along with the item *-n*, whatever its nature - two quantifiers, both of them denoting a low number[23]:

(554) Noun + w + [n + wāḥid$_{\text{ONE}}$]
Noun + COP + [n + Q$_{\text{LOW NUMBER}}$]

(555) Noun + w + [n + (’) a$_{\text{FEW}}$]
Noun + COP + [n + Q$_{\text{LOW NUMBER}}$]

to the effect that adopting a non-restrictive RC analysis for the complex *-n cum* quantifier is plausible at least in principle.

To demonstrate that this is actually the case, another aspect of the behavior of *-n* should be taken into account, namely the meaning of the noun when not combined with it (in structural terms, a relation *in absentia*). The noun thus characterized in OA has both a SG and a PL meaning:

(556) a. *fulk*: ‘ship’/ ‘ships’ (*Šarḥ al-Kāfiyah*, vol. III, *al-Jam‘*, p. 368)

b. *tamr*: ‘date (fruit)’/ ‘dates’ (*Šarḥ al-Kāfiyah*, vol. III, *al-Jam‘*, pp. 366-7)

c. *rūm*: ‘Greek’/ ‘Greeks’ (*Šarḥ al-Kāfiyah*, vol. III, *al-Jam‘*, p. 366-7)

d. *taba‘*: ‘companion’/ ‘companions’ (Fleisch, 1961, vol. I, p. 365)

e. *walad*: ‘boy’/ ‘boys’ (*Lisān al-‘Arab*, p. 4914, s.v. *W L D*)

Western scholars traditionally analyze this form of the noun as COLL (= high number, more than 10 units: see Section 3.2.2.1), but the data just reported clearly show that this is not the whole story, COLL coexisting with SG and PL. Greenberg (1990b) must be credited as the first to rely upon primary sources to provide this accurate scenario. The relevant passage, also

[22] In the same Section it is also explained why the unattested forms are to be considered as phonological representations instead of reconstructed forms.

[23] The Long Vowel Shortening Rule described in Section 7.6 bridges the gap between the phonological representation *w* of Conj / NOM and the alternation in its phonological representation between *u* and *ū*.

quoted in Greenberg (1990b, p. 182) is from Al-'Astārābāḏī's *Šarḥ al-Kāfiyah* (on which, see Section 2.3.5):

> "...to which one may add that the COLL noun is applied to the low and the high number, since *l-tamr* [cp. (556)] stands for: one date, two dates, the dates [PAUCAL PL] and the same holds for *l-rūm*, so that if you eat a date or two dates, or deal with a Greek or two Greeks you may still say 'I ate *l-tamr*' and 'I dealt with *l-rūm*'...".
> (*Šarḥ al-Kāfiyah*, vol. III, *al-Jam'*, p. 367)[24]

Greenberg himself interestingly remarks that in this respect the Arabic *n*-less noun behaves as the noun found in the Numeral Classifier languages (on which, see Section 1.1.3), where the bare noun displays the same ambiguity between SG and PL:

> "Emeneau (1951: 85) describes the Vietnamese noun when unaccompanied by a classifier in terms quite reminiscent of Raḍīuddīn in regard to the generic noun of Arabic: "A non-numerated substantive phrase...lacks any indication of number or individuation; that is when there is no explicit indication of number, a number is entirely free of reference to the number category. For example *tôi mûon mua sàch* 'I want to buy book(s)'. There is absolutely no indication how many books are intended.".
> (Greenberg, 1990b, p. 183)

Shortly put:

(557) In Numeral Classifier languages, SG and PL nouns have one and the same form

Given that the counterpart of the bare noun in such languages is the noun associated with a complex Numeral Classifier *cum* quantifier (the so called Numeral Classifier Construction: see Section 1.1.3) and that the counterpart of the Arabic *n*-less noun is a noun associated with a complex *-n* cum quantifier, *-n* is very likely to be identified with a Numeral Classifier on the basis of typological considerations. If this reasoning is correct, instances of postnominal *-n* like those of (551, 552) are to be reinterpreted along the following lines:

(558) Old Arabic / Classical Arabic (Veccia Vaglieri, 1937, p. 102)

rajul-u-n (wāḥid)
man-COP-CLF (one)

'one / a man'

(559) Najdi Arabic (Ingham, 1982a, p. 54)

sāyg-i-n WĀḤID
driver-COP-CLF ONE

'one / a driver'

[24] *'...wa-tazīdu 'alay-hi 'anna-sma l-jinsi yaqa'u 'alà l-qalīli wa-l-kaṯīri fa-yaqa'u l-tamru 'alà-l-tamrati wa-l-tamratayni wa-l-tamrāti wa-ka-ḏā l-rūmu fa-'in 'akalta tamratan 'aw tamratayni wa-'āmalta rūmiyyan 'aw rumiyyayni jāza li-ka 'an taqūla 'akaltu l-tamra wa-'āmaltu l-rūma...'.* English translation of this text partly reproduces Greenberg's.

(557) is a semantic diagnostic property of OA corroborating this assumption. The next Section provides further syntactic diagnostics supporting this hypothesis.
It is important to bear in mind throughout reading of the next Section that in the Numeral Classifier Construction *no item* is marked for SG / PL, number being expressed by the quantifier itself (Greenberg, 1990a).

6.6.1 Diagnostics and minimalist analysis

The diagnostics proposed here concentrates on the syntax of the SG Numeral Classifier Construction[25] and relies upon three considerations made in the typological literature about this latter:

- the word order tends to be Noun > Classifier > Quantifier (Simpson, 2005, p. 822), Greenberg (1990a, p. 235)[26]
- the numeral 'one' can be deleted Greenberg (1990a, p. 227, fn. 2)[27]
- the remaining material Noun > Classifier subsequent to deletion turns into either an indefinite or a definite article Greenberg (1990a, p. 235)[28]

The scenario is summarized in (560), (561), (562) (563), where the Numeral Classifier languages are Thai (560) and two Indo-Aryan languages, Bengali and Oriya (561 - 563), belonging to the Magadhan subfamily and closely related each other, as per Greenberg (1990a, p. 235) and Ray (2003, p. 485). The contrast between (561) and (562) shows that the functional item involved in definiteness (*khana*) is a Classifier. The contrast between (563) and (564) shows that the functional item involved in indefiniteness (*jone*) is a Classifier:

(560) Thai (Simpson, 2005, p. 822)

dek khon nung
child CLF:HUMAN one

'one / a child'

(561) Bengali (Greenberg, 1990a, p. 235)

boĭ-khana
book-CLF:GENERAL

'<u>the</u> book'

(562) Bengali (Greenberg, 1990a, p. 235)

[25] For a richer diagnostics, see Grande (2011).

[26] 'The number one however is commonly placed following the classifier in Thai.'. 'In the closely related Assamese...Kakati (1941), however, gives with the numeral 'one' expressed and in the indefinite meaning both Q-Cl-N and N-Cl-Q.'

[27] 'There are, however, a few instances in which the classifier goes with the noun, but these usually involve deletion of the numeral 'one'. An example is the Bengali and Assamese 'definitive' mentioned later in the paper, e.g. Bengali *boĭ-khana* 'book-the',in which *khana* functions as a numeral classifier in other constructions.'

[28] 'Many classifier languages have constructions in which the classifier appears without a quantifier. The meaning here is invariably singular and is in some languages specified as definite, in others indefinite and still others neither.'

pāc-khana boĭ
five-CLF:GENERAL book

'five books'

(563) Oriya (Ray, 2003, p. 499)

ḍaktor jone
doctor CLF:HUMAN

'a doctor'

(564) Oriya (Ray, 2003, p. 493)

dvi jone ḍaktor
two CLF:HUMAN doctor

'two doctors'

To summarize, the diagnostic properties of the SG Numeral Classifier Construction can be stated as follows:

Syntactic Diagnostics 6.6 (Classifier status)**.**

a. If Quantifier = one, then the word order: Noun > Classifier > Quantifier, with Noun and Classifier unmarked for number
b. If Quantifier = one, then deletion of Quantifier and the word order: Noun > Classifier
c. If Quantifier = deleted one, then the same word order turns the Classifier into a definite or indefinite article.

The first two properties are also found in the just mentioned Arabic constructions featuring the SG *-n*. This lends further empirical support in favor the hypothesis entertained here that *-n* is a Numeral Classifier: see (558, 558) immediately above.
In addition, the third property of the numeral classifier construction featuring the numeral 'one' straightforwardly accounts for a so far puzzling behavior of *-n*, namely its ability to work both as definite and indefinite article (contrast OA, CA, Najdi Arabic (558), (559) with Epigraphic South Arabian (547)).
As just shown, in fact, this behavior is typical of the diachronic development (grammaticalization) of the Numeral Classifier and is a consequence of its intrinsically being neither definite or indefinite, since both these meanings develop out of the (understood) numeral 'one' combined with it (see Simpson 2005 and his subsequent work on this issue).
The just discussed parallelism between the properties of Numeral Classifier and that of *-n* is depicted in (565 - 567). The parallelism is even stronger if one calls into the picture the property that the Numeral Classifier and *-n* have *in absentia* (568), namely the pretheoretical fact that the Classifier-less and the *n*-less noun conveys both a SG and a PL meaning (see end of the previous Section):

(565) Thai: dek khon nung (= 560)
OA / CA: rajul (u) n wāḥid (= 558)

(566) Bengali: boĭ khana (= 561)
Epigraphic South Arabian: ṣlm n (= 547)

(567) Oriya: ḍaktor jone (= 563)
(OA / CA,) Najdi Arabic: sāyg (i) n (= 559)

(568) Thai: sach = book / books (= Greenberg 1990b, p. 183, 557)
OA: walad = boy / boys (= 556)

To this, one may add that the OA Numeral Classifier Construction traditionally referred to as SOUND PL has the same word order Noun > Classifier > Quantifier found in its SG counterpart. Greenberg (1990b) reports that this word order is also found in Kuna, an endangered Native American language spoken in Panama and Colombia:

(569) Kuna (Greenberg, 1990b, p. 186)

ome war po
woman CLF two

'two women'

(570) illustrates the parallelism at issue:

(570) Kuna: ome war po
OA: muslim (ū) n (')a^{FEW}

What precedes leads to rewrite (554), (555) as follows:

(571) Noun + w + [n + wāḥid$_{\text{ONE}}$]
Noun + COP + [CLF + Q$_{\text{LOW NUMBER}}$]

(572) Noun + w + [n + (') a$_{\text{FEW}}$]
Noun + COP + [CLF + Q$_{\text{LOW NUMBER}}$]

A non-trivial corollary of identifying *-n* as a Classifier is interpreting the entire complex Classifier plus quantifier as a non-restrictive RC, since it has been seen in Section 5.2.3 that in some languages the substitution test is successful in replacing the former with the latter (see the Gilbertese examples (318), (319) in Section 5.2.3). This is desirable under the current perspective, since the modifier featuring *-n* is realigned with the other modification structures, themselves interpreted as a non-restrictive RCs.
Therefore the syntactic structure assigned to the Arabic SG and SOUND PL marked for *-n*, actually Numeral Classifier Constructions, turns out to be identical with that of the Pre-semitic RC type (511):

(573) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ pro [Modifier ...
... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ pro [n + Quantifier ...

As a corollary, what has been said above in connection with (573.a) holds for (573.b) as well. The postnominal position of the Arabic SG and SOUND PL marked for *-n* (573.b), when displaying the Conj *w* (e.g. OA / CA (558)) derives from the *representational* neutralization of the blind POS *w - Op* (Abstract Incorporation, by analogy with the Amorite POS), while the same construction displaying the Conj *a, i* plausibly still has the syntactic structure in (573.b) but in this case the POS at issue is *derivationally* neutralized, to the effect that the the syntactic structure in (573.b) undergoes movement.
Finally, it is possible to provide a more accurate definition of the Numeral Classifier *-n* by looking at its semantics in OA. Sībawayhi (*apud* Fleisch 1961, vol. I, p. 445, fn. 1) reports

that at least in the SG, *-n* is combined with the formative *i* (as much as it happens in Najdi Arabic: cp. (559)) to denote a single *person* out of a 'COLL' of nationality (actually, a bare noun unspecified for SG, PL etc...: see Al-'Astārābāḏī's passage quoted at the end of Section 6.6):

(574) Old Arabic (Fleisch, 1961, vol. I, p. 445, fn. 1)

yamān-i-n (WĀḤID)
Yemen-COP-CLF ONE

'one / a man from Yemen'

One may raise the objection that (574) is not a reliable piece of data because this construction is not cited in its pausal form, contrary to what would be expected for OA (see Section 3.2.1): the pausal form at issue is *yamān-ī*, not *yamān-i-n* (Wright, 1896, vol. II, p. 370).
This move, however, is justified by the findings of this Section, where it has been shown on the basis of evidence from Najdi Arabic that *-n* preceded by the item *-i- is* word-final (cp. (559)). This evidences that *-n* is a Numeral Classifier (originally) referring to human beings, that is an ANIMATE CLASSIFIER - which from a typological standpoint does not rule out the possibility that it undergoes semantic changes.
Greenberg (1990a) finds that crosslinguistically any Numeral Classifier can lose its original function and turn into a different functional item and Simpson (2005) points out that among the new functions taken over by the General Classifier is that of definite or indefinite article. A crucial implication of these typological considerations is that *-n* can be plausibly interpreted in this way, since it has the ability to work either as a definite article (in South Arabian) or as indefinite article (in OA / CA). Such a characterization of *-n* will prove to be useful in due course.
To summarize the main finding of this Section:

(575) The tanwīn is an Animate Numeral Classifier turned into a General Classifier

(575) answers the question (VI), since the traditional opposition of pausal forms: SG bare noun vs. PL noun ending in *-na* (see Section 3.2.1) is to be abandoned in favor of the opposition: singulative *-n (wāḥid)* vs. paucal *-na*.
In the best case, the Greenbergian Numeral Classifier Hypothesis extended in this work from *-at* (on which, see Section 7.9.3 in the next Chapter) to *-Vn* should be able to derive its restrictions of cooccurrence with the (alleged) prenominal *l-* (see Section 3.2.1). This is left for future research.

6.6.2 The Possessive Classifier

In this Section the Arabic possessive constructions (Construct State, Free State: Section 3.2.2.2) are examined through the lens of the Numeral Classifier Hypothesis worked out in the previous chapters.
The interpretation of the Construct State and of the Free State proposed here heavily draws from Ouhalla (2010) in two main respects[29].
First, the particle intervening between the Possessee and the Possessor, Pennacchietti (1968)'s

[29]Thanks to Jamal Ouhalla for helpful discussion about the main ingredients of this Section: possessive construction, (in)alienability, kinship terms and pronouns in Arabic.

nota genitivi does not have the same syntactic status of the particle intervening between the Noun and the ('Adjectival') RC, Pennacchietti (1968)'s *nota relationis (/ designationis)*[30]. Thus, whatever the categorial status of the former, it is anyway different from the original Conj status of the latter ones.

Second, this work adopts the twofold typology of the possessive construction set up in Ouhalla (2010), departing from the positions of the received view (generally accepted in the generative paradigm) and being rather reminiscent of the unified account of Construct State and Free State given by Arab Grammarians (cp. (13)). This is illustrated in Table 6.2.

This section will concentrate on Shamaliya Arabic, a variety currently spoken in the northwest

Table 6.2: Arabic Possessive Constructions: Typology

Approach	Constructions	
Received view	PRONOMINAL CS, NON PRONOMINAL CS	PRONOMINAL FS, NON PRONOMINAL FS
Ouhalla's	PRONOMINAL CS	NON PRONOMINAL CS, PRONOMINAL FS, NON PRONOMINAL FS
FS = FREE STATE, CS = CONSTRUCT STATE, *pronominal* = whose complement is a pronoun		

region of Morocco and described in Ouhalla (2010), for a diachronic reason: Shamaliya Arabic shares with Akkadian a feature concerning the possessive construction lacking in OA / CA. Needless to say, such a feature-sharing points to an archaic state of affairs for Shamaliya Arabic, which therefore provides more useful information about the nature of Arabic (and Semitic) possessive construction than OA and CA, whose possessive construction, departing from Akkadian, is a later and less telling innovation.

This comes as no surprise under the approach à la Owens (2006) adopted here, where MCV are regarded as very conservative languages (see Section 2.2.2.6).

The feature at issue is the pretheoretical fact that both Akkadian (Blau, 1979b, pp. 32-33) and Shamaliya Arabic (Ouhalla, 2010) tend to avoid the *non-pronominal CS* if the Possessee is INALIENABLE (in the wake of Blau 1979a, pp. 31-32 an inalienable entity can be described as something whose ownership is not transferable or intrinsic). Instead, they prefer to instantiate a pronominal CS, where the Possessee is fused with a pronominal Possessor (the so-called PROLEPTIC PRONOUN), in its turn followed by a particle and the non-pronominal Possessor, as exemplifed in the following:

(576) Shamaliya Arabic (Ouhalla, 2010, p. 3)

xt-hu d (al-'ayl)
sister-his (of the-boy)

'his sister / the boy's sister'

(577) Akkadian (Blau, 1979a, p. 32)

aḫātī-šu ša S.
sister-his of S.

[30]Pennacchietti (1968) takes the phonological identity of *nota genitivi, relationis, designationis* found in Djidjelli Arabic *eddi* (see (1 - 3) in Section 1.1.1) as the main argument in favor of his unified account of these items. Recently, however, Ouhalla (2010) argues on the basis of diachronic and textual evidence that the *nota genitivi eddi* (along with its variants: *di, dyal* etc...) in the Maghrebi area is the result of language contact between the Moroccan Colloquial Varieties and the Indo-european preposition *di, de* etc...

'S.'s sister'

On the other hand OA (since its earlier stages: North Arabic) and CA combine the Possessee with the non-pronominal Possessor *tout court* (non-pronominal CS):

(578) Liḥyanitic (Pennacchietti, 1968, p. 36)

'nzh bn 'aws
'Anazah son 'Aws

'Anazah, son of 'Aws'

Returning to Shamaliya Arabic, (576) exemplifies the pronominal CS, (579) the (rarely used) non-pronominal CS, (580) the pronominal FS, (581) the non-pronominal FS in this Colloquial Variety of Arabic:

(579) Shamaliya Arabic (Ouhalla, 2010, p. 15)

wld al-bnt
son the-girl

'the girl's son'

(580) Shamaliya Arabic (Ouhalla, 2010, p. 3)

al-ktab dyal-u
the-book GM-him

'his sister'

(581) Shamaliya Arabic (Ouhalla, 2010, p. 3)

al-ktab d al-'ayl
the-book of the-boy

'the boy's book'

In Shamaliya Arabic, both the pronominal CS and the FS in general show some peculiar morphosyntactic and semantic properties.
Taking the meaning-side of the pronominal CS as departure point, it should be noted that the pronominal CS is possible iff and only if it has an INALIENABLE reading, as the contrast between (576) and (582) illustrates:

(582) Shamaliya Arabic (Ouhalla, 2010, p. 16)

*ktab-ha
the-book her

'her book'

There seems also to be a strict solidarity between the CS and the pronominal Possessor: as just seen, in (576, 577) it is this latter that is fused into a CS with the inalienable Possessee, not the non-pronominal Possessor.
As for its sound-side, it has been also remarked in Section 3.2.2.2 that the CS has at least two

properties, namely lack of article on the modified noun (the Possessee) and strong morphological cohesion between the Possessee and the Possessor: they make up a word-like unit, easily recognizeable by the phenomenon of phonological reduction of *-at* into *-t*. This is a property of pronominal CS not only in Shamaliya Arabic, but more generally in Moroccan Arabic and appears evident from the contrast between (583) and its CA equivalent *madrasatī*:

(583) Moroccan Arabic (Benmamoun, 2006, p. 481)

mədras-t-i
school-*at*-me

'my school'

Another diagnostic property characterizing a pronominal and (*unambiguously*) inalienable CS is observed in OA (/ CA) and consists of a lengthened V intervening between the Possessee and the Possessor. In this connection Ingham (1994, p. 16) remarks that in Najdi Arabic 'the final *-a*, when followed by a non-stressed particle or monosyllabic word with no long vowel, is often lengthened', so that a long V is another morphosyntactic cue that signals a word-like unity:

(584) Old / Classical Arabic (Wright, 1896, vol I, p. 249)

'ab-ū-ka
father-COP-you

'your father'

(585) Najdi Arabic (Ingham, 1994, p. 16)

hin<u>a</u> → (ma) hn<u>ā</u>-šin
here (not) here-thing

'there is nothing...'

To summarize, the (Shamaliya) Arabic pronominal CS is such that:

- it has an inalienable reading (576, 582, 584)
- the Possessor is a pronoun (576, 584)
- the Possessee is not marked by the article (576, 584)
- the Possessee and the Possessor are a word-like unit (583, 584)

Turning to the FS, and looking at its semantics, it is characterized by a particle preceding the Possessor and such a particle typically is a former inalienable noun meaning *property, wealth* or the alike, having undergone grammaticalization (Ouhalla, 2010, p. 3). A telling example is Moroccan Arabic *nta'*, preceding the Possessor in the FS in (587 - 588) and etymologically connected with the word *mta'* 'property' still found in Shamaliya Arabic (586), as per Ouhalla (2010):

(586) Shamaliya Arabic (Ouhalla, 2010, p. 25)

mta' al-makhzn
property the-State

'State property'

(587) Moroccan Arabic (varieties A) (Ouhalla, 2010, p. 10)

al-ktub ntaw‘ Nadia
the-book *nta‘*.BROKEN PL Nadia

'Nadia's books'

(588) Moroccan Arabic (varieties B) (Ouhalla, 2010, p. 2)

al-ktub nta‘ al-bnt
the-book *nta‘* the-girl

'the girl's books'

Bardeas (2009) and Ouhalla (2010) analyze the agreed-with manifestation of the particle of the FS (e.g. *nta‘* of (587) and see also Section 3.2.1) as a preposition resulting from grammaticalization of a former inalienable noun, but following Dryer (2005c, p. 347) this analysis appears to be problematic from a typological viewpoint: for items to be defined as grammaticalized, 'there must be some reason to believe that they have grammaticalized to some extent, that they are to some extent *grammatically distinct* from other nouns (or verbs)' (italics mine).
The point is that the particle *nta‘* in (587) is not distinct enough from a noun, since it is able to take the same markers usually taken by a noun or its dependencies (cp. the BROKEN PL on it), which rules out its grammaticalization and hence its prepositional status.
Nonetheless, Dryer (2005c)'s diagnostic criterion supports Ouhalla (2010)'s and Bardeas (2009)'s prepositional analysis of the same particle when it is not agreed with its Possessee: lack of agreement properties point to its non-nominal and hence grammaticalized / prepositional status. Thus, while in some varieties of Moroccan Arabic (587) the agreed-with *nta‘* is still an inalienable noun, in some others (588) the same item, *qua* not agreed-with, has evolved into a preposition.
Furthermore, the FS is preferably associated with an alienable reading, because among inalienable Possessees kinship terms cannot occur into the FS (Ouhalla, 2010, p. 16) and at the same time alienable Possessees are only allowed to occur in it, their instantiation into a CS being barred (see (582)). This semantic property is confirmed by Brustad (2000, p. 86)'s dialectological research: 'In most dialects, the genitive exponent [= *dyal*, *nta‘* etc...] may not modify an inalienable noun'. In sum, the (Shamaliya) Arabic FS can be characterized as follows:

- it (preferably) has an alienable reading (582)
- it is an inalienable noun (possibly grammaticalized into a preposition) (586, 588)

The scenario sketched out so far is typical of Numeral Classifier Languages, where the opposition inalienability vs. alienability is realized by means of the semantic, syntactic and phonological properties described above.
In these languages, in fact, inalienable possession (first property of the pronominal CS) is expressed *tout court* by juxtaposing the Possessee and the Possessor (second property), with no additional functional material on the Possessee (third property): (589) exemplifies this state of affairs.

Juxtaposition gives rise to a word-like unit: in (589) the item *nu-* is a suffix (cp. also Aikhenvald 2000, p. 128, which more generally says that 'inalienably possessed nouns can take affixes directly, while the alienably possessed nouns cannot') and is clearly a reduced form if compared with its non-suffixal variant *nuha* of (592). Data are drawn from Tariana, an (endangered) language spoken in the Amazonas (Brasil), and reported by Aikhenvald (2000), which explicitly states that (589) has an inalienable reading:

(589) Tariana (Aikhenvald, 2000, p. 131)

nu-pana
me-home

'my home (inalienable reading)'

As for the alienable possession, Aikhenvald (2000, p. 128) observes that in Numeral Classifier languages this piece of information is conveyed by a construction featuring a Possessee and a Possessor 'governed' by a *dedicated* Classifier, different from that co-occurring with quantifiers. This is the so-called POSSESSIVE CLASSIFIER.
Its diagnostic properties are defined by Aikhenvald (2000, p. 128) as follows: 'the [possessive] classifiers - which are used with alienably possessed nouns - are themselves in fact a subclass of inalienably possessed nouns'. This is tantamount to saying that there is a strict correlation between the alienable reading (first property of the FS) and manifestation of an inalienable possessed noun preceding the Possessor (second property of the FS). (590) exemplifies the Possessive Classifier Construction, where *ite* is to be likened to Arabic *ḥaj* etc...(Aikhenvald 2000, p. 132 explicitly states that (590) has an alienable reading):

(590) Tariana (Aikhenvald, 2000, p. 132)

nu-ite tshinu
me-POSSESSIVE CLF: ANIMATE dog

'my dog (alienable reading)'

There emerges then a strict parallelism between manifestation of possession in (Shamaliya) Arabic and in Numeral Classifier Languages: more accurately, between the pronominal CS and their inalienable possessive construction, on one side, and between the FS and their Possessive Classifier Construction. This parallelism has been also hinted at in Greenberg (1990a, p. 239) and Aikhenvald (2000, p. 138) on the basis of data from Egyptian Arabic and Hebrew, respectively, and is illustrated in the following:

(591)

Pronominal CS	(584) = Classifier-less type	(589)
FS	(580) = Possessive Classifier Construction	(590)

This typological scenario is flexible enough to allow a certain overlapping between inalienable and alienable possession under certain conditions, which once again are strikingly similar to some phenomena encountered in Arabic.
First, Aikhenvald (2000, p. 131) points out that the Classifier-less type (589) / pronominal CS is no longer a word-like unit if it loses its inalienable meaning in favor of the alienable one. Actually, (592), the alienable counterpart of the inalienable (589) replaces the reduced form *nu-*, an affix fused with its Possessee *tshinu* into a single word, with the non-reduced form *nuha*, a word detached from the word *tshinu*:

(592) Tariana (Aikhenvald, 2000, p. 131)

nuha tshinu
I dog

'my dog (alienable reading)'

In a likewise fashion, in OA (/ CA) the pronominal CS is used not only when the Possessee is intrinsically inalienable (e.g. the kinship term in (584)), but also if the Possessee is potentially alienable (e.g. (593)). Noteworthily, in this latter case the Vowel Lengthening signalling a word-like unit (cp. (584) with (585)) cannot apply. This entails that the Possessee and the Possessor involved in the pronominal CS having an alienable reading (e.g. *tijārat-u-hum* of (593)) are not a word-like unit, but two separate words:

(593) Old Arabic (*Koran*, 2:16)

fa-mā rabiḥat tijārat-u-hum
and-not profited commerce-COP-hum

'and their commerce (alienable reading) has not profited them'

An argument in favor of analyzing the pronominal Possessor (e.g. *hum* of (593)) as an independent word is the fact that in OA (/ CA) it also functions as a verbal object (594): in the same distributional context, Indoeuropean languages instantiate a clitic, which differs from a suffix in that it is not (obligatorily) attached onto the verb (cp. the English translation of (594))[31]:

(594) Old Arabic / Classical Arabic

ra'aytu-hum
I.saw them

'I saw them'

These data corroborate the claim that Arabic (593) can be likened to Tariana (592) in that their alienable meaning correlates with the inability for the pronominal Possessor to make up a word.
Second, Aikhenvald (2000, p. 131) reports that the Possessive Classifier Construction (590) / FS can be used regardless of its alienable / inalienable semantics for pragmatic reasons: 'Possessed classifiers can be used with any noun if the fact of a noun being possessed is to be focused on. Example 5.18 [= (595)] illustrates the use of a possessed classifier with an inalienably possessed noun *panisi* 'home' in Tariana':

(595) Tariana (Aikhenvald, 2000, p. 131)

nu-yadapana panisi
me-POSSESSIVE CLF: HOUSE home

'my home$_j$ (inalienable reading + focus$_j$)'

In much the same way, Brustad (2000, p. 76) observes that 'the genitive exponents fulfill specific pragmatics functions that the construct state does not' , among them 'contrastive focus' and she brings as evidence in favor of this latter function what follows:

[31] But see Shlonsky (1997, Ch. 9) for arguments in favor of the opposite view.

(596) Kuwaiti Arabic (Aikhenvald, 2000, p. 78)

bēt aḥmad yābir <u>māl</u> id-dīra
house Ahmed Jaber POSSESSIVE CLF the-old.city

'Ahmed Al-Jaber's house$_j$ [the one] of the old city (inalienable reading + focus$_j$)'

In (596), the particle *māl* puts the focus on the possessed noun, as explicitly stated by Brustad (2000, p. 78) 'here, one particular house of this member of the ruling family is contrasted to his other houses' . Arabic FS (596) hence parallels Tariana Possessive Classifier Construction (595) in that *māl*, like the Classifier *yapanisi*, is 'exceptionally' associated with an inalienable meaning in order to convey focus.

6.6.3 Deriving the non-pronominal Construct State

The arguments brought in the previous Section constitute evidence enough for identifying the twofold Arabic possessive construction along the lines of Numeral Classifier languages. Nonetheless, in spite of the converging evidence provided so far, the main difficulty with this account is that there is no room in it for the *non-pronominal* CS.

This problem evaporates as soon as this latter is subsumed under one of the two just discussed constructions: either the (pronominal) CS or the FS.

The received view opts for the former choice, OA native speakers as well as Ouhalla (2010) for the latter (see Section 3.3.6 and Table 6.2). If one follows native speakers' judgments, the non-pronominal CS is a non-pronominal FS, at least insofar as its syntactic structure prior to spell-out is concerned (*'aṣl*)[32], because the particle preceding the Possessor is dropped:

(597) NON PRONOMINAL FS = Possessee + X + Possessor, X: overt
NON PRONOMINAL CS = Possessee + X + Possessor, X: covert

What is left explained so far is why such a particle is dropped at all.

It will be argued here that dropping of the particle at issue follows from the Possessive Classifier analysis proposed here.

To begin with, consider the non-pronominal CS, exemplified in (579) and characterized by the item *l-*. This latter cannot be taken to be a 'prenominal article' referring to the Possessor to its right, because it is a well known fact that *l-* scopes over *both* the Possessor and the Possessee (see e.g. Benmamoun 2006): *both* of them, not only the Possessor, have a definite reading.

The only alternative left is to analyze it as a 'medial article' , ie as a complementizer of a RC (see Section 6.4.3), which in principle makes sense, given that Classifiers are embedded into an (underlyingly non-restrictive) RC (see Section 6.6.1 and (573)).

Nevertheless, Section 6.4.3 plausibly demonstrates that the complementizer *l-* can be inserted in a syntactic structure manifesting itself in secondary predication as a disambiguating element, in order to avoid that it be mistaken for an instance of primary predication (see (524)).

Intuitively, the *l-* found in the *non-pronominal CS* (itself a subtype of *non-pronominal FS*) arises from the fact that it can be ambiguously read also as some kind of possessive construction occurring in primary predication - ie as a POSSESSIVE SENTENCE. It should be added that the Possessive Sentence at issue must be sought in OA and / or in North Arabic because these

[32] See Section 2.3.3 for more details on Arab Grammarians' notion of *'aṣl* as well as on its modern equivalent(s) and more generally on Arab Grammarians' *'aṣl*-based model.

are the oldest Arabic varieties where the Arabic non-pronominal CS is first documented (in particular, for North Arabic, see (578)).
The Late Lexicographer 'Abū l-'Abbās al-Fayyūmī (d. 760 / 1358) in its dictionary *al-Miṣbāḥ al-Munīr* (*apud* Lane 1863, vol. III, p. 986) reports the following possessive construction in primary predication, found in a line of OA poetry attributed to the Arabic poet Nābiġah al-Ḏubyānī[33]:

(598) Old Arabic (Lane, 1863, vol. III, p. 986)

maḥall-at-u-hum ḏāt-u-l-'ilāh
abode-*at*-COP-their thing-COP-the-God

'Their abode is that of God '

Of particular relevance here is that Lane (1863, vol. III, pp. 985-986) follows Arab Grammarians and Lexicographers in glossing the noun *ḏāt* as 'possessum, thing, essence' and assigns the *ḏāt* preceding the Possessor in (598) a literal meaning ('it seems more reasonable to render this phrase agreeably with the primary signification of *ḏāt*').
Lane (1863, vol. III, p. 985) and Pennacchietti (1968, p. 39) after him also remark that another item listed by Arab Lexicographers under the same root of *ḏāt*, namely *ḏū*, at least in OA has a similar meaning: '*ḏū* and *ḏāt* ... are also used as prefixed nouns in various expressions here following, in several thereof as meaning *Something in possession* or *the like*; *not a possessor*' . Among the examples mentioned by Lane (599) is found:

(599) Old Arabic (Lane, 1863, vol. III, p. 985)

wada'at il-mar'at-u ḏā / ḏāt-a baṭn-i-hā
brought.forth-3FSG the-woman-NOM *ḏū*-ACC / *ḏāt*-ACC womb-OBL-her

'The woman brought forth [the fruit] of her womb'

Adopting the diagnostic criterion set by Dryer (2005c) *ḏ(ū)* in OA cannot be interpreted as a preposition because of its nominal properties: its ability to be marked for 'case' (actually, a copula intervening between it and its complement of specification) in (598 - 599).
Unluckily, as far as it known, no instance of *ḏū* and *ḏāt* meaning 'possessum, thing, essence' is documented in OA in a context of secondary predication (modification structure). Yet, Pennacchietti (1968, pp. 39-40) observes that (the cognate of OA) *ḏū* occurs in this context in North Arabic[34], as illustrated in (600):

(600) Thamudean Arabic (North Arabic) (Pennacchietti, 1968, p. 37)

'mḫm ḏ-tmm
'Ammaḫam *ḏ*-Tamīm

'Ammaḫam, one of the Tamīm tribe'

[33]Recall from Section 2.2.2.4 and 2.3.2 that data drawn from OA poetry can be regarded as reliable if transmitted by Arab Grammarians and Lexicographers.

[34]'Ciononostante, sono attestati vari costrutti con *ḏū* che documentano una fase della lingua in cui il pronome non veniva collegato ad alcun significato o in cui esso significa addirittura l'opposto di "possessore di", ossia "qualcosa in possesso di"...In ogni caso, l'uso della *nota genitivi* pronominale ha in arabo classico un carattere puramente locuzionale e rappresenta la specializzazione e il residuo fossile di un fenomeno sintattico che un tempo associava il nordarabico all'aramaico e alle lingue sudarabiche.'

The comparison between (598) and (600) reveals that they share one and the same syntactic structure, namely:

(601) Possessee + ḏ̲ū + Possessor

potentially ambiguous between a secondary predication interpretation (FS), and a *be*-sentence whose Predicate is introduced by *ḏ̲ū*. This is exemplified in the following:

(602) ‘mḫ̲m ḏ̲-tmm = Ammaḫ̲am, which is one of the Tamīm tribe
‘mḫ̲m ḏ̲-tmm = Ammaḫ̲am is one of the Tamīm tribe

It is therefore expected (cp. (525, 526)) that the disambiguating element *l-* is inserted in secondary predication (600), intervening between the Subject and its Predicate, in order to keep it distinct from its primary predication counterpart (599):

(603) Dem. Cstr.: Dem.$^{Subj.}$ + Noun$^{Pred.}$ > Dem. + l- + Noun
Non-pron. FS: Poss.ee$^{Subj.}$ + [ḏ̲ū + Poss.or]$^{Pred.}$ > Poss.ee + l- + [ḏ̲ū + Poss.or]

Insertion of *l-* thus characterized would result in an OA syntactic structure along the following lines:

(604) ḥ̣arr-u l-ḏ̲ū šams
heat-COP *l*-POSSESSIVE CLF sun
‘the heat of the sun’

True, as a matter of fact (604) is not attested in OA (/ CA) but nonetheless OA native speakers feel that something is dropped in the syntactic position preceding the Possessor, so that one may hypothesize that what is dropped is *ḏ̲ū*. This results in the familiar instance of non-pronominal CS, exemplified by (605)[35]:

(605) Old Arabic (/ Classical Arabic) (Wright, 1896, vol. II, p. 199)

ḥ̣arr-u l-šams
heat-COP *l*-sun

‘the heat of the sun’

and taken here actually to be:

(606) ḥ̣arr-u l-0^{DHU} šams
heat-COP *l*-POSSESSIVE CLF sun
‘the heat of the sun’

In this light, dropping of the element preceding the Possessor posited by Arab Grammarians is straightforwardly accounted for as a manifestation of the phonological rule of *ḏ̲*-deletion (*lḏ̲ū* > *l-*) triggered by the need of avoiding the non-euphonic cluster **l(V)ḏ̲, ḏ̲(V)l* (457), also responsible for reshaping of the RC marker *llaḏ̲ī* into *lli* (see Section 6.3) and of the article *ḏ̲ā-l* into *l-* (see Section 6.4.3).
Evidence in favor of this claim lies in two phenomena observed in the MCV.
The first piece of evidence is that Moroccan Arabic manifests the same situation of interpretive ambiguity between a possessive construction and a possessive sentence just reported for OA

[35]Recall that the OA forms generally correpond to the CA pausal forms and therefore there is no (putative) case-ending at the end of the non-pronominal CS exemplified below.

/ North Arabic - in this case the syntactic structures at stake involve the equivalent of *ḏū*, namely *dyal*.
Remarkably, in Moroccan Arabic such ambiguity is solved by insertion of a disambiguating element, which lends empirical support to the hypothesis entertained here that the same strategy is at work in the OA counterpart of Moroccan Arabic. Moroccan Arabic and OA minimally differ in that the disambiguating element in the former is the pronoun *huwwa* showing up in primary predication, along the lines of (461, 462), while the latter inserts *l-* in secondary predication. In Moroccan Arabic, in fact, the possessive construction (588) discussed in the previous Section and the possessive sentence (607):

(607) Moroccan Arabic (Ouhalla, 2000, p. 237)

had l-ktab dyal Nadia
this the-book *dyal* Nadia

'This book is Nadia's'

share one and the same syntactic structure (cp. (601)):

(608) Possessee + dyal + Possessor

so that (607) can be disambiguated via insertion of *huwa* (in the required agreed-with form):

(609) Moroccan Arabic (Ouhalla, 2010, p. 10)

had al-dar hiyya dyal-t Omar.
this the-house.F huwwa.F *dyal-at* Omar

'This house is Omar's'

The second piece of evidence is an interesting correlation found in the dialectal opposition Bedouin vs. Sedentary Varieties, already mentioned in the Table 2.3 of Ch. 2 and replicated here in the Table 6.3 (from Palva 2006). A simple glance at this Table clearly shows that in the Bedouin varieties manifestation of the FS is 'blocked' by the presence of the interdental fricatives *ṯ, ḏ...*, while in in the Sedentary Varieties, where the interdental fricatives *ṯ, ḏ...* are absent (having evolved into postdental stops), the FS does occur.
The lack of FS in the Bedouin varieties is a natural consequence of the fact that it (underlyingly) has a Classifier *ḏū* made up of an interdental fricative *ḏ*, which phonologically conflicts with the complementizer *l-* and is therefore deleted, turning the FS into a CS.
The conclusion is a sobering one:

Table 6.3: Syntax of the Noun Phrase: distinctive isoglosses #2

Variety	Bedouin	Sedentary
Interdental fricatives	retained	> postdental stops
Free State	very low frequency	high frequency

(610) The non-pronominal CS displaying a medial *l-* is no more than a non-pronominal FS (Possessive Classifier Construction) in which the Possessive Classifier *ḏū* is deleted because of the non-euphonic cluster *l(V)ḏ*. The Possessive Classifier *ḏū* in fact is immediately preceded by the complementizer *l-*

Notice that this interpretation also derives a well known property of the non-pronominal CS, namely

- lack of article on the Possessee

Consider again the non-pronominal FS featuring a non-prepositional particle introducing the Possessor: in diachrony, this is an old construction, already attested in Akkadian, a language where nouns are *not* adorned with the 'prenominal article' (see (6) in Section 1.1.3).
If the OA / CA non-pronominal CS is a subtype of non-pronominal FS, as plausibly demonstrated here, there is no need of assuming that the Possessee first gets a 'prenominal' article and hence deletes it, as usually done. More simply, the OA / CA non-pronominal CS is a relic construction, preserving the same state of affairs of the non-pronominal FS attested in Akkadian (6).
Finally, notice that OA / North Arabic *d̲ū* parallels Moroccan Arabic in another respect, namely its development into a preposition, because in CA, a later stage of language, the non-pronominal CS (606) is paraphrased as a non-pronominal FS where the particle preceding the Possessor is not a noun / *'ism* (grammaticalized into a Possessive Classifier) but a preposition / *ḥarf*: *li, min* etc... see (11 - 13) in Section 1.1.3 and Section 3.2.2.2.
What has been discussed so far is the non-pronominal CS displaying the complementizer *l-*. But this is not the only instance of non-pronominal CS: in OA / CA its *l*-less counterpart also occurs, as illustrated in (113), reproduced below as (611):

(611) Old Arabic (Wright, 1896, vol. II, p. 125))

qafīz-u burr-i-n
qafīz-NOM wheat-OBL-*n*

'one *qafīz* (unit of measurement) of wheat'

Notice that this construction is not cited in its pausal form (which would have a final *burr* instead of *burrin*), contrary to what would be expected for OA, because of the finding of Section 6.6.1: here it has been shown on the basis of evidence from Najdi Arabic that *-n* preceded by the item *-i-* is word-final (559) (see also discussion of (574)).
The Najdi Arabic piece of data, however, requires that the complex *-i-n* is a complex Conj copula plus Classifier with a numeral 'one' understood, an issue that will be clarified shortly.
Returning to the *l*-less non-pronominal CS as a whole, the discussion made in Section 3.2.2.2 has shown that contrary to what the received view holds, the opposition between non-pronominal CS displaying *l-* and its *l*-less counterpart is *not* one of definiteness vs. indefiniteness and that they instead differ, as hinted at by OA native speakers' judgments, *in their syntactic structure prior to spell-out* (*'aṣl*).
Indeed, Arab Grammarians report for the just discussed non-pronominal CS displaying *l-* the *'aṣl* in which some particle occurs (here taken to be *d̲ū*), while they assign its *l*-less counterpart a different *'aṣl*, which coincides with the so-called Measure Construction: thus (611) is said to arise from the Measure Construction illustrated in (612). It is worth observing in this connection that the core of such a construction is the item *-n* intervening between the unit of measurement and the material it measures and, furthermore, that it displays the invariable (putative) case-ending *-an*, occurring not only in CA non-pausal forms, but also in CA pausal forms and in its OA equivalent (see Section 3.2.1):

(612) Old / Classical Arabic (= 112)

qafīz-u-n burr-a-n
qafīz-NOM-*n* wheat-ACC-*n*

'one *qafīz* (unit of measurement) of wheat'

Interestingly, Greenberg (1990b, p. 228) takes the Measure Construction to be the more typical instance of a Numeral Classifier Construction - actually, the latter diachronically arises from the former ('in numeral classifier languages the numeral classifier construction is always almost identical with the measure construction').
This strongly suggests that syntactically the *l*-less non-pronominal CS is a Possessive Classifier Construction / FS featuring the Numeral Classifier *-n* instead of the Possessive Classifier *d̲ū*.
In spite of its appeal, this account raises two problems.
First, the Possessive Construction crosslinguistically instantiates a dedicated Possessive Classifier and therefore a Numeral Classifier taking over its function looks suspicious. Second, even admitting that this is possible, one may wonder where the numeral meaning typically associated with the Measure Construction (cp. English *1 kilo of bread*) is expressed, if *-n* of (612), being interpreted as a Possessive Classifier, is no longer seen as a Numeral Classifier. Indeed, in the normal case the numeral 'one' can be taken to be covert iff and only if *-n* is a Numeral Classifier (Section 6.6.1).
The former problem is solved if one considers that *-n* diachronically develops into a General Classifier (see end of Section 6.6.1 and (575)), which crosslinguistically has the ability to replace all the other Classifiers, Possessive Classifiers included (Greenberg 1990a, Aikhenvald 2000, pp. 127, 260, Aikhenvald 2003, Ch. 21). A telling example from Tariana is (613), where the General Classifier *yarupe* replaces the 'dedicated ' Possessive Classifiers *ite, yadapana* etc... of (590, 595).

(613) Tariana (Aikhenvald, 2003, p. 483)

Saba yarupe marie
Saba GENERAL CLF:THING knife

'Saba's knife'

The latter problem is solved simply by looking more carefully at the syntactic structure of the Measure Construction in Arabic (612): one of its distinctive properties is presence of the invariable ending *(-a)n*. Thus, the *-n* preceding the Possessor is a Possessive Classifier, signalling the (metaphoric) possession relationship between the unit of measurement / Possessee, e.g. *qafīz* of (612) and the material it belongs to / Possessor, say *burr* of (612).
On the other hand the *-n* found at the end of the entire construction is a Numeral Classifier introducing the quantity measured by the unit of measurement, such a quantity being a covert *WĀḤID*. In other words, *qafīz*, the Possessee of (612) has two modifiers, the former a Possessive Classifier Construction, the latter a Numeral Classifier Construction.
Precisely in the same way the Possessee is connected with the Possessive Classifier Construction *-n + Possessor* by means of the copula *-u-* (the putative NOM), so is it connected with the Numeral Classifier Construction *-n WĀḤID* via the copula *-a-* (the putative ACC). This is schematized in the following:

(614) Possessee + u + [n + Possessor] + a + [n + *WĀḤID*]
Noun + COP + [POSS. CLF + Noun] + COP + [NUM. CLF + Q]

A non-trivial corollary of the just motivated claim that *-n* shifts from a Numeral Classifier into a General Classifier (replacing the Possessive Classifier) is explaining why the Measure Construction (612) turns into the non-pronominal CS (611).
It has been recently pointed out that the semantic weakening of the General Classifier may result in its deletion. Thus Tao (2009) in his study about the Mandarin Chinese General Classifier *ge* brings evidence in favor of the assumption that not only it has an overt realization, as standardly assumed, but also a covert one. In (615) the General Classifier *ge* replaces the Numeral Classifier *liàng* when the numeral is *yí* 'one' and they form a complex modifying *jiǎotāchē* 'bycicle', while in (616) the same numeral directly modifies the noun *huoji* 'employee' without the (overt) General Classifier *ge*:

(615) Mandarin Chinese (Aikhenvald, 2000, p. 408)

...qǐ-zhe yí liàng jiǎotāchē uh shì yí ge hěn kěaì-de
ride-ing one Cl:VEHICLE bicycle uh be one Cl:GENERAL very cute-PARTICLE
xiǎo-de jiǎotāchē
small-PARTICLE bicycle

'(a child) riding a bycicle, a very cute small bicycle'

(616) Mandarin Chinese (Tao, 2009, p. 299)

yi-huoji tui men jin lai
one-employee pushed door enter come

'An employee pushed-opened the door and came in'

It ensues that interpreting (the leftmost) *-n* of the Measure Construction (612) as a General Classifier derives its dropping in the non-pronominal CS (611) with no additional machinery.
In this light, the non-pronominal CS is no more than a subtype of Measure Construction / non-pronominal FS, sharing with it the same syntactic structure:

(617) Possessee + u + [0^n + Possessor] + i + [n + *WĀḤID*]
Noun + COP + [POSS. CLF + Noun] + COP + [NUM. CLF + Q]

Notice that in (617) the item preceding the Numeral Classifier (the rightmost *-n*) has turned from *a* into *i* for reasons unclear at the current stage of research (but probably having to do with the covert realization of the General Classifier *-n*). This however does not affect the main point: both the items *a* and *i* being commutable with *u* (see end of Section 6.4.2) are copulae .
To sum up:

(618) The *l*-less non-pronominal CS is no more than a non-pronominal FS (Possessive Classifier Construction) in which the General Classifier *-n* taking over the function of the Possessive Classifier is deleted because of its semantic weakening.

Finally, it seems convenient to remark that a strict application of Bardeas (2009)'s / Ouhalla (2010)'s prepositional analysis of the Arabic non-pronominal FS (as revised in the light of Dryer 2005c: cp. previous Section) makes it possible to interpret the General Classifier *-n* of the Measure Construction (612) also as a preposition.

After all, the leftmost -*n* of the Measure Construction (612) does not have noun-like properties, because it is not marked for (putative) Case in a language (OA) where nouns usually are: cp. the discussion surrounding (598) and (599).
On these grounds, the fact that an item -*n* in Lybico-Berber, a Hamitic language, is usually interpreted as a preposition is taken here to be not accidental and to straightforwardly follow from an extension of Bardeas's / Ouhalla's account from *nta'*, *dyal* etc... (587 - 588) to -*n*:

(619) Lybico-Berber (Lipinski, 1997, p. 471)

awal n umaziġ
word of Amazigh

'the Amazigh's word'

In comparative terms, the Language Contact Hypothesis makes this theoretical move plausible iff and only if -*n* belongs to the Pre-semitic parastrate that Arabic shares with Hamitic. This is actually the case, because -*n* manifests itself within the Pre-semitic RC type, as shown by the fact that it co-occurs with the Pre-semitic copula *w*, *not* with the Amorite copula *d̲ī* - more accurately, -*n* manifests itself in the Pre-semitic RC type, not in the Amorite one: see (573) in Section 6.6.1.
The prepositional behavior of -*n* is interesting in two respects. First, it unveils an interesting parallelism with 0^{DHU}, since both of them turn from Possessive Classifiers into prepositions (see above on 0^{DHU}), the former in the Measure Construction (612), the latter in the non-pronominal CS displaying the medial *l*- (606):

(620) Non-pron. CS etc... l- + 0^{DHU} (0^{DHU} : Poss. Clf. > Prep.)
Measure Cstr. 0^{C} + n: (n : Poss. Clf. > Prep.)

Second, it derives the complementizer status of -*n* in North Arabic, discussed at the beginning of Section 6.6 and illustrated in (545, 546). Also recall from Section 6.4.4 that the complementizer status of North Arabic -*n* depends on the strict semantic and syntactic solidarity between D and C (phasehood: see (542)).
The prepositional -*n*, in fact, is the missing link between the (Numeral / Possessive / General) Classifier -*n* and the complementizer -*n*, because in the typological literature the shift of a preposition into a complementizer is a well known phenomenon. A familiar example from Indo-european languages is English *for* (Pesetsky & Torrego, 2004): contrast e.g. *this is for John* with *for John to be late / that John is late*.
However, if one assumes that -*n* in the Measure Construction (612) becomes a complementizer, one is also forced to assume that its prepositional function is taken over by a covert preposition, as illustrated in (621.b).
The presence of a covert preposition in this syntactic structure, far from jeopardizing the hypothesis of a shift of -*n* from complementizer into preposition, reinforces and evidences it, because the combination complementizer - covert preposition is precisely what is found in the counterpart of the Measure Construction, namely the non-pronominal CS displaying the medial *l*- (606), as illustrated in (621.a):

(621) a.Non-pron. CS etc... l- + 0^{DHU} (0^{DHU} : Preposition)
b.Measure Cstr. n- + 0 (0 : Preposition)

Simply put, analogical pressure exerted by the non-pronominal CS displaying the medial *l*- plays a key-role in the shift of *-n* into a complementizer and in the manifestation of a covert preposition in the Measure Construction (recall from (610) that the covert realization of the preposition in the source of analogy is due to phonological reasons).
To conclude this Section, the drift of the Animate Classifier *-n*, traditionally designated as *tanwīn*, is summarized in (622):

(622)

Anim. Num. Clf	> Gen. Clf.	> Poss. Clf.	> Preposition	> Complementizer
(574)	(547)	(612)	(612), (619), (620)	(612), (621)

The next Section derives the drift at issue in terms of Dynamic Antisymmetry.

6.6.4 The minimalist analysis

The findings of the previous Section are consistent with Ouhalla (2010)'s twofold typology of the Arabic Possessive Construction summarized in Table 6.2. In the minimalist implementation of this idea, the opposition pronominal CS vs. FS (including the non-pronominal CS) is translated into an opposition between the syntactic structure A and the syntactic structure B.
The core of the syntactic structure A that Ouhalla (2010, p. 15) assigns to the pronominal CS is (623) and is characterized by adjacency of the Possessee and its Possessor:

(623) ... $[_{N'}$ N $[_{DP}$ Pronoun...

While retaining the gist of Ouhalla (2010)'s analysis, the present work recasts it within a dynamic antisymmetric framework, as it will be seen in Ch. 8. For the time being, suffice it to say that the derivation involving (623) or the alike ends up with (624), since under standard antisymmetric assumptions this is the structure assigned to nouns not adorned with article (cp. Giusti 2002 and (337) in Section 5.2.4):

(624) ... $[_{DP}$ NP $[_{D'}$ D [...Pronoun...

To make the discussion more concrete, Shamaliya Arabic *xt-hu* 'his sister' (576) exemplifies the structure in (624)

(625) ... $[_{DP}$ xt $[_{D'}$ D [...hu...

Ouhalla (2010) assigns (623) also to the exceptional kind of non-pronominal CS exemplified by Shamaliya Arabic *mta' al-makhzn* 'State property' (586). This instance of non-pronominal CS cannot be taken to be a 'disguised' FS because its Possessee, unlike the Possessee of the latter, has an unambiguous inalienable reading. Accordingly (586) can be represented along the following lines at the end of its derivation:

(626) ... $[_{DP}$ NP $[_{D'}$ D [...DP...

(627) ... $[_{DP}$ mta' $[_{D'}$ D [...al-makhzn...

Turning to the FS / non-pronominal CS, the development of Ouhalla (2010)'s account put forward in the previous Section claims that this possessive construction has at its core a Possessive Classifier, making up with the Possessor a Possessive Classifier Construction.
In particular, in OA (612) the Possessive Classifier Construction is (asyndetically) coordinated

with a Numeral Classifier Construction (see (614)), an (underlyingly) non-restrictive RC. Applying the COORDINATION TEST to (612) in all likelihood points to extension of the non-restrictive RC analysis from the Numeral Classifier Construction to the Possessive Classifier Construction. Hence, the syntactic structure B assigned to the FS / non-pronominal CS is basically a modification structure made of a noun and of a non-restrictive RC.
Thus, the Moroccan Arabic FS etc... can be represented along the lines of the mixed RC type (513), where both the Conj and the C introducing the RC have a covert realization:

(628) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro [Modifier ...
... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro [nta‘ + Possessor ...

The OA / CA non-pronominal CS displaying the medial *l-*, exemplified in (606), is the Pre-semitic RC type (512), since its Conj and C are overt (*w, l-* respectively):

(629) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- [Modifier ...
... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- [(d̲ū) + Possessor ...

Finally, the *l*-less non-pronominal CS illustrated in (612) corresponds to the Pre-semitic RC type (511), characterized by the overt Conj *w* and by the null C (pro):

(630) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro [Modifier ...
... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro [-n + Possessor ...

Under this analysis, the internal structure of the RC is essentially a predication structure, where Op is the Subject and the complex Possessive Classifier plus Possessor its Predicate. In greater detail, the complex in question is the special kind of non-pronominal CS depicted in (627), because the Possessive Classifier is no more than a former inalienable Possessee (see Section 6.6.2).
Accordingly, a more fine-grained representation of (628 - 630) is the following:

(631) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{DP}}$ nta‘ [...Possessor...
... $[_{\text{ConjP}}$ l-ktub $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{DP}}$ nta‘ [...l-wld...

(632) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- $[_{\text{DP}}$ d̲ū [...Possessor...
... $[_{\text{ConjP}}$ ḥarr $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- $[_{\text{DP}}$ 0^{DHU} [...šams...

(633) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{DP}}$ -n [...Possessor...
... $[_{\text{ConjP}}$ qafīz $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{DP}}$ -n [...burr...

The Possessive Classifier sitting in Spec, DP in all of these structures is a Head-within-Spec adjacent with the null C and is clearly monosyllabic (*nta‘, d̲ū, -n*). Moreover, it cannot co-occur outside the non-pronominal CS with the Dependents usually combined with a noun: Ouhalla (2010) explicitly states that *nta‘* is grammatical in (586) but ungrammatical if combined with a Demonstrative etc...:

(634) Shamaliya Arabic (Ouhalla, 2010, p. 26)

*hada l-mta‘
this the-property

‘this property’

The same holds for OA *ḏū, -n*: the OA equivalent of (634) **hāḏā l-ḏū* is not possible on the ground that *ḏū* albeit described as a noun (*'ism*) by Arab Grammarians can be only combined with a Possessor (Lane, 1863, vol. III, p. 985), and *-n*, as discussed throughout this Chapter, is never found in contexts like (634). On the whole, the Possessive Classifier has clitic-like properties, so that its adjacency with the higher null C creates the blind POS *pro - nta'*, *pro - ḏū*, *pro - n*.

This POS cannot be representationally neutralized (Word Formation) via Abstract Incorporation, because pro and the Possessive Classifier are not coindexed for some common feature, and its derivational neutralization, ie movement, is then resorted to. Yet, such a strategy too is inhibited for a structural reason: *no landing site of movement is found, because the first Spec available, Spec, CP is at the same time the topmost Spec of CP and already fulfilled by Op.*

The only possible solution is adding more syntactic structure between Spec, CP and Spec, DP in order to allow the Possessive Classifier to raise and it seems plausible that the (portion of) syntactic structure at issue is to be identified with PP.

This is so because under a strict application of the Kaynean antisymmetry PP is the only projection that can select DP as its complement, even when a transitive VP appears to select an object DP: following Cinque (2006, p. 157, 165 fn. 38) this is a deceiving effect of the covert realization of the PP intervening between VP and the object DP (sometimes surfacing as *a, pe* in Spanish and Romanian, respectively).

After merging of PP, there obtains the following syntactic structure, exemplified by the OA *l*-less non-pronominal CS, but the same applies to the other instances of FS above:

(635) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{PP}}$ $[_{\text{P}'}$ P $[_{\text{DP}}$ $\boxed{\text{-n}}$ [...Possessor...
... $[_{\text{ConjP}}$ qafīz $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{PP}}$ $[_{\text{P}'}$ P $[_{\text{DP}}$ $\boxed{\text{-n}}$ [...burr...

A more careful examination of (635) has interesting implications for a minimalist approach incorporating the Chomskian phase theory (see Section 5.2.4), according to which any complete chunk of semantico-syntactic information is a phase and that therefore both DP and CP qualify as such. Of particular relevance here is that Chomsky contends that a phase, being a self-contained linguistic entity from a syntactic and semantic standpoint, is sent to C-I and A-P for interpretation and Spell-out respectively in order to become a linguistic entity not only in the speaker's mind but also in the 'real world'.

As such, a phase already shipped off to the interfaces is no longer in Syntax and hence no longer accessible to further syntactic operations (Merge, Move as well as, under the approach developed here, Word Formation) and relations (Spec-Head relation, asymmetric c-command) except for its EDGE, that is the topmost complex Spec - Head of a given phase: Spec, CP plus C and Spec, DP plus D.

The ontological status of the 'phase edge' is evidenced by the semantic phenomenon of co-referentiality, whereby a relative pronoun or a demonstrative pronoun / adjective refer back to a noun in some intuitive sense 'outside' of their domain, 'elsewhere' in the context (cp. *John$_i$, [who$_i$ had left yesterday]...* and *John$_i$, [this$_i$ guy]...*). But in minimalist terms, a relative pronoun sits in Spec, CP, a demonstrative pronoun in D, its adjectival counterpart in Spec, DP and their co-referentiality with a noun is the result, among others, of the *syntactic* relation of asymmetric c-command: this means that Syntax still operates on the Spec and on the Head of a phase and that therefore these latter are still 'located' in Syntax and have not yet been sent to C-I and A-P.

There emerges then a clear cut difference between, one the one hand, the phase edge (= Spec + Head), still located in Syntax and still accessible to its operations and, on the other hand, what is not a phase edge, ie its complement (say, Compl, CP / Compl, DP), which has been already shipped off to the interfaces and not further accessible to syntactic operations.
It follows that in (635) 'injecting' new syntactic structure (PP) into the already existing one (CP) must take place immediately above Compl, CP, ie immediately above the DP Op, because this is the only portion of the phase CP in which the syntactic operation of Merge can apply.
This phenomenon, if looked at from the vantage point of Dynamic Antisymmetry, strongly suggests that there might plausibly exist a kind of algorithm in instantiating movement (derivational neutralization of a POS) in natural languages. First, following Moro (2000), it takes place via insertion of a copula (cp. also (452) in Section 5.3.5). Second, if this strategy fails, via Domain Extending Head Movement (see (449) in Section 5.3.5). Third, if the second strategy too fails, Syntax deploys a new strategy, namely the just described instance of Merge, and seems that this latter has to do with Phasehood.
Further research is needed to get a deeper understanding of this phenomenon, which for the time being will be described in (636), replacing (452):

(636) In the derivational neutralization of a POS, the copula is a Head X merged into the syntactic structure to provide a landing site (Spec, XP) with the raised pole of a POS and hence to make the derivation converge at A-P. If the landing site is already fulfilled by another constituent, the raised pole of the POS is forced to move to the next available Spec via Domain Extending Movement. But if the copula X and its already occupied landing site Spec, XP are the edge of a phase (CP / DP), no next available Spec can be made available in order to raise the pole via Domain Extending Movement and to derivationally neutralize the POS. Accordingly, Merge applies, which injects new syntactic structure in the edge itself, ie immediately above Compl, CP / DP

Returning to the vicissitudes of the syntactic structure in (635), it seems convenient to observe that the very merger of the Head P solves one problem (it provides the landing site Spec, PP allowing neutralization of the blind POS *pro - nta'*, *pro - d̲ū*, *pro - n*) but at the same time it creates another. Indeed, a new blind POS *covert P - n* etc...arises because the clitic-like Possessive Classifier, of nominal origin, is the Head-within-Spec immediately below the null preposition. But given that the preposition crosslinguistically develops out of a noun (Dryer, 2005c), it is coindexed with the Possessive Classifier for the syntactic feature Noun or the alike and therefore this blind POS is representationally neutralized (Word Formation) via Abstract Incorporation, in a way highly reminiscent of the blind POS *covert Conj - d̲ī*, as schematized in the following:

(637) (= 476)

a. $[_{\text{Conj}'}\ 0 \qquad [_{\text{CP}}\ \underline{\underline{\text{dī}}}\ \rightarrow$
b. $[_{\text{Conj}'}\ 0 + \underline{\underline{\text{dī}}}\ [_{\text{CP}}\ \text{Op}$

(638) a. $[_{\text{P}'}\ 0 \qquad [_{\text{DP}}\ \underline{\text{-n}}\ \rightarrow$
b. $[_{\text{P}'}\ 0 + \underline{\text{-n}}\ [_{\text{DP}}\ 0^{CLF}$

As a consequence of the representational neutralization of the POS at issue, *-n* incorporates into P giving rise to a new word-like unit, the (composite) preposition *(P)n* hosted in the Head P. More generally, *-n* is illustrative also of the Abstract Incorporation involving *nta'* , *d̲ū*:

(639) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{C'}$ pro $[_{\text{PP}}$ $[_{\text{P}'}$ $\boxed{\text{-n}}$ $[_{\text{DP}}$ 0_{Clf} [...Possessor...
... $[_{\text{ConjP}}$ qafīz $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{C'}$ pro $[_{\text{PP}}$ $[_{\text{P}'}$ $\boxed{\text{-n}}$ $[_{\text{DP}}$ 0_{Clf} [...burr...

This derives the shift of *-n* etc... from Possessive Classifier into Preposition[36].

The new scenario, represented in (639) cannot gives rise to a new POS between the null C and the composite preposition *(P)n*, *(P)nta'*, *(P)d̲ū* because this latter, in spite of its being a monosyllabic Head adjacent to the null C, does not meet a condition for the POS to arise: namely, the inability to have Dependents.

By definition, in fact, a preposition without Dependents is ungrammatical. This explains why at least *(P)nta'* and *(P)d̲ū* after grammaticalizing into preposition do not undergo any further development.

Yet, the shift of *(P)n* into a complementizer (see end of previous Section) intuitively has to do with its adjacency to C and it is tempting to capture this insight using once again the phenomenon of POS, without invoking some *ad hoc* machinery.

The only obstacle to the application of the blind POS *pro - (P)n* in (639) is the Dependent-less status of the prepositional *-n*, but this problem is less serious than it seems to be, because in OA the prepositional *-n* in synchrony coexists with the Numeral Classifier *-n*, and Classifiers are by their own nature Dependent-less (see (634)).

Therefore it seems safe to maintain that the (Numeral) Classifier *-n* interferes by means of a *rapport associatif* with the prepositional *-n*, to the effect that this latter is felt by the speaker as Dependent-less.

This is even more so recalling that the Numeral Classifier *-n always* co-occurs with the prepositional *-n on the syntagmatic axis too*, ie in the same construction, because the *l*-less non-pronominal CS is actually a conjoined construction manifesting both the kinds of *-n* (see (614)). Shortly put, the blind POS *pro - (P)n* is motivated.

As for its neutralization, it applies representationally (Word Formation) via Abstract Incorporation, since the complementizer and the preposition do share a common feature. For the present purposes, it can be informally defined as the property of denoting space and location, as an informal comparison between English *where* and *in* demonstrates (Pesetsky & Torrego 2004 discuss this property for the preposition, Giorgi 2010 for the complementizer and the interested reader is referred to these works for more details). As a result, *-n* incorporates into pro, giving rise to the (complex) complementizer *(pro P)n* and thus yielding the expected shift from preposition into complementizer. This is schematized in the following:

(640) (= 476)

a. $[_{\text{Conj}'}$ 0 $[_{\text{CP}}$ d̲ī̲ →
b. $[_{\text{Conj}'}$ 0 + d̲ī̲ $[_{\text{CP}}$ Op

(641) a. $[_{C'}$ pro $[_{\text{P}'}$ -n̲ →
b. $[_{C'}$ pro + -n̲ $[_{\text{P}'}$ 0

(642) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{C'}$ $\boxed{\text{-n}}$ $[_{\text{PP}}$ $[_{\text{P}'}$ 0 $[_{\text{DP}}$ 0_{Clf} [...Possessor...
... $[_{\text{ConjP}}$ qafīz $[_{\text{Conj}'}$ w $[_{\text{CP}}$ Op $[_{C'}$ $\boxed{\text{-n}}$ $[_{\text{PP}}$ $[_{\text{P}'}$ 0 $[_{\text{DP}}$ 0_{Clf} [...burr...

[36]Notice that the just discussed blind POS: null C - Possessive Classifier does not occur if the Possessive Classifier is endowed with some adnominal marker (cp. (587)) probably because affixation of this latter makes the Possessive Classifier non-monosyllabic. As a consequence, the Possessive Classifier can be still analyzed as such and non-prepositional.

Once that -*n* grammaticalizes into a complementizer, the Head of a Phase, its merger can be generalized to Phases in general, so that it also manifests itself as D, as shown in (546), here repeated as (643):

(643) $[_{DP}$ h- $[_{D'}$ n- ['lt ...
$[_{DP}$ the $[_{D'}$ D [goddess ...
'the goddess'

The just given account of the diachronical drift of -*n* in part dovetails with Pesetsky & Torrego (2004), which propose a similar interpretation for English prepositions like *for* when they act as complementizers, arguing for their composite structure.
On their view, *for* is a complementizer plus a preposition, which they take to have to do with tense, so that they claim that their 'treatment of English *for*, however, suggests that such elements are actually instances of T whose presence in C is due to movement — a hypothesis that might be plausibly extended to similar phenomena in other languages.' (Pesetsky & Torrego, 2004, p. 510).
Movement from P to C à la Pesetsky & Torrego (2004) is replaced here by Abstract Incorporation, a theoretical move which makes it possible to dispense with the feature-checking framework (see Section 4.5.2) invoked by them.
Last but not least, -*n* when grammaticalized into C / D acts as a copula between a Subject (Demonstrative *h*- or Op respectively) and a Predicate (Noun and RC respectively) in secondary predication. This unveils a strong generalization in the Arabic NP:

(644) All Arabic adnominal markers, Classifiers included (after a process of grammaticalization), act as copula.

Focusing on the copular nature of -*n* (via a diachronical development), it is worth observing that the Classifier Hypothesis answers the question (VIII) by providing an unified account of the Demarcative Hypothesis and of the Copula Hypothesis.
As for the Demarcative Hypothesis, the 'demarcative' effect of -*n* (ie its ability to 'close' a word) à la Ayoub (1991, 2006) is a natural consequence of the fact that -*n* is a (Numeral) Classifier, which covertly manifests the modifier if this latter is the numeral 'one'. By the same token -*n* when followed by a Possessor / modifier is still interpreted as a (Possessive) Classifier but this latter can evolve into C, itself a copular element (between Op and the RC predicate), in line with the Copula Hypothesis.
The shift of -*n* into a copular C documented for the possessive construction opens the possibility that such an analysis (Copula Hypothesis) be generalized, as hinted at in Owens (2006), for the modification structure in which -*n* intervenes between the noun and its RC, this latter including the (Extended) Adjective (see e.g. (116) in Section 3.2.2.3). This is left for future research.

6.7 Summary

This Chapter has proposed a reduced typology of the Arabic RC types, (Extended) Adjectives included. This is based on the Language Contact Hypothesis à la Garbini & Durand (1994) defended in Ch. 2, which makes it possible to draw a clear dividing line between a pure Amorite RC type and a pure Pre-semitic RC type. All the other RC types one way or the other are

mixed types.
The Amorite RC type is characterized by the development of the demonstrative base *d̲ī* from a relative pronoun (documented in South Arabian) into a Conj copula (documented in OA and Nedroma Arabic), which points to its non-restrictive nature (at least originally).
A two-fold diagnostic tool, based on Arab Grammarians work, has been worked out in order to empirically ground such a claim: this is the ability for a copula of the Amorite parastrate to undergo *la*-prefixation and to resist case-marking, irrespective of whether it occurs in primary predication (*be*) or in secondary predication (Conj).
An account of the Amorite RC type resorting to Dynamic Antisymmetry derives not only the abovementioned shift, but also the postnominal position of the RC and the deletion of the sequence *w - relative pronoun* attested in some mixed RC types. The core of the explanation is the phenomenon of blind POS, which is representationally neutralized via Abstract Incorporation because its poles are coindexed for the Demonstrative feature. This is illustrated in the following (the blind POS is *0 - d̲ī*):

(645) $\begin{array}{l} ...[_{\text{ConjP}}\ \text{NP}\ [_{\text{Conj}'}\ 0_i \qquad [_{\text{CP}}\ \underline{\text{d}}\bar{\text{i}}_i ... \quad > \\ ...[_{\text{ConjP}}\ \text{NP}\ [_{\text{Conj}'}\ 0_i + \underline{\text{d}}\bar{\text{i}}_i\ [_{\text{CP}}\ \text{Op}_i ... \end{array}$

The Pre-semitic RC type too is originally non-restrictive, as shown by another diagnostic tool worked out on the basis of the behavior of non-restrictive RCs in Somali. This comparison is justified by the fact that Somali belongs to the Hamitic family, which under the Language Contact Hypothesis is conservative of a Pre-semitic state of affairs. The diagnostics at issue is the ability for a non-restrictive RC to get also an adverbial reading. The strict solidarity between the Somali RC and the Pre-semitic one, as found in OA, allows to identify the putative NOM of this latter - actually, a morpheme *u / w* of unclear function so far - with the conjunction *u / w* 'and' widespread in Semitic (cp. also its Somali equivalent *oo*). Moreover, this makes it possible to extend to the Pre-semitic (OA) RC the minimalist representation proposed for the Somali RC in Frascarelli & Puglielli (2006):

(646) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w- / oo $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro...

A dynamic antisymmetric account of (646) reveals the presence of the blind POS *w - Op*, which can be neutralized via movement only (ie derivationally) because its poles are not coindexed for the Demonstrative feature.
Nonetheless, a mixed RC type found in Levantine Arabic opens the possibility that the *Pre-semitic* blind POS *w - Op* can be also representationally neutralized by means of Abstract Incorporation if it gets the Demonstrative feature by analogy with the Amorite blind POS *0 - d̲ī*:

(647) $\begin{array}{lll} \text{a.} & ...[_{\text{Conj}'}\ \boxed{0_i}\ [_{\text{CP}}\ \boxed{\underline{\text{d}}\bar{\text{i}}_i} ... \Leftrightarrow ...[_{\text{Conj}'}\ \boxed{\text{w}_i}\ [_{\text{CP}}\ \boxed{\text{Op}_i} ... & > \text{Abstract Inc.} \\ \text{b.} & & ...[_{\text{Conj}'}\ \text{w}\ [_{\text{CP}}\ \text{Op}... & > \text{Movement} \end{array}$

More concretely, the Pre-semitic blind POS *w - Op* mimicking the Amorite blind POS *0 - d̲ī* (647.b) leaves the RC in situ, in the postnominal position and hence it is precisely this *analogical* POS that accounts for the pre-Semitic RC type displaying the Conj *w* (the putative NOM) exemplified in (500).
Intuitively, the 'pure' Pre-semitic POS is in some way responsible for the Conj other than *u* (the putative NOM), namely *a, i* (the putative ACC, OBL). This idea will be implemented in

Ch. 8.
A later Pre-semitic RC type exhibits an overt complementizer *l-* (medial *l-*), plausibly a former disambiguating element, corresponding to Frazier's morphological cue:

(648) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ l-...

The same medial *l-* occurs in the demonstrative construction, whose demonstrative base *d̲-* develops into an article:

(649) d̲-^{Dem} + *l-* + Noun

Dropping of the demonstrative base / article *d̲-* results in the covert realization of this latter and in the deceiving effect that the medial *l-* is prenominal, to the effect that the medial *l-* has been mistaken so far for the Arabic article:

(650) 0^{Dem} + *l-* + Noun

Dropping of the demonstrative base / article d̲- is due to a phonological rule of *d̲*-deletion triggered by the need of avoiding the non-euphonic cluster:

(651) *l(V)d̲, *d̲(V)l

also responsible for reshaping of the RC marker *llad̲ī* into *llī* in the MCV.
It has been also shown in the present Chapter that the Arabic nouns modified by *-n* at least in the SG and in the SOUND PL are Numeral Classifier Constructions, where *-n* is a Classifier followed by a quantifier: a (covert) numeral 'one' in the SG and a paucal morpheme *(')a* in the PL (Classifier Hypothesis):

(652) Noun + w + [n + wāḥid$_{\text{ONE}}$]
Noun + COP + [n + Q$_{\text{LOW NUMBER}}$]

(653) Noun + w + [n + (') a$_{\text{FEW}}$]
Noun + COP + [n + Q$_{\text{LOW NUMBER}}$]

The Numeral Classifier Construction has a syntactic structure corresponding in its essence to the Pre-Semitic RC type, as exemplified by the following OA example:

(654) ... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ pro [Modifier ...
... [$_{\text{ConjP}}$ NP [$_{\text{Conj}'}$ w- [$_{\text{CP}}$ Op [$_{\text{C}'}$ pro [n + Quantifier ...

By the same token, evidence has been brought in support of Arab Grammarians's hypothesis that the FS and the non-pronominal CS are one and the same construction and this idea has been combined with the Classifier Hypothesis. Under this perspective, the construction at issue is a Possessive Classifier Construction, characterized by the presence of a Possessive Classifier intervening between the Possessee and the Possessor, which can be realized either overtly (FS) or covertly (non-pronominal CS).
The covert realization of the Possessive Classifier is given either a phonological or a semantic rationale. The former is the constraint in (651), which involves the Possessive Classifier *d̲ū* when interacting with the medial *l-*, the latter is the semantic weakening of the Numeral Classifier *-n*, when taking over the function of the Possessive Classifier *d̲ū*. Once again, the Possessive Classifier construction is the Pre-semitic (non-restrictive) RC type, as exemplified by the OA non-pronominal CS displaying the medial *l-*:

(655) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ w- $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- [Modifier ...
... $[_{\text{ConjP}}$ Possessee $[_{\text{Conj}'}$ w- $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- [(d̲ū) + Possessor ...

It has been then shown that the several values ascribed to *-n* (*tanwīn*) are straightforwardly accounted for if the Classifier Hypothesis is adopted, since from this vantage point they are the result of the following grammaticalization process, which has been in part put forward for Arabic by Ouhalla (2010):

(656) Anim. Num. Clf > Gen. Clf. > Poss. Clf. > Preposition > Complementizer

Finally, a dynamic antisymmetric account has been proposed. It takes as its departure point Ouhalla (2010)'s minimalist account of the Arabic possessive construction and recasts it in terms of Dynamic Antisymmetry, thus deriving the drift in (656) as a recursive manifestation of a POS.

The unified interpretation of the Arabic RCs, adjectives and possessive constructions defended in this Chapter has the advantage of answering some questions unanswered in the received view.

Chapter 7

Towards a reconsideration of Arabic Case

7.1 The copula analysis of Case

Owens (2006) proposes that Arabic 'case-endings' *u, a, i* (NOM, ACC, OBL) originate from epenthetic vowels. Tosco (*apud* Owens 2006, p. 101, fn. 22) observes in this connection that vowels from which Arabic 'case-endings' plausibly arise can hardly be epenthetic in nature, since on a typological and diachronic viewpoint Case develops out of previous *morphemic* material (cp. also Blake 2001, Ch. 6).

Thus, Owens (2006, p. 83) concedes that 'it would not be surprising if such a system developed at the proto-stage Semitic out of markers of another type'.

The previous Chapter has plausibly shown that the morphemic value of the putative NOM *u* is a copula and the fact that the putative ACC, OBL *a, i* in OA are commutable with it strongly suggests the same analysis.

This scenario, however, does not explain why three copulae are needed instead of just one. A possible solution to this problem is assuming that *a, i* perform a copular function at least in part different from that of *u*.

Testing the validity of this assumption requires a more in-depth discussion of how Arabic 'Case' was thought of by native speakers and, to a certain extent, how 'Case' manifests itself in its original milieu, namely the pre-Semitic parastrate.

It is important to bear in mind throughout this Chapter that the linguistic phenomenon of Arabic Case subject to reinterpretation here corresponds to the (Arabic) Morphological Case of the generative paradigm[1].

[1]The Abstract Case will be not dealt with in this work. This is consistent with the dynamic antisymmetric approach. In some influential versions of the generative / minimalist paradigm, in fact, Abstract Case can be invoked to explain movement, being involved in the interplay between interpretable - uninterpretable features typical of the Checking Theory (on which, see Section 4.5.2), but Dynamic Antisymmetry does not need feature-checking to account for movement (*ibidem*), and therefore Abstract Case appears not to be a crucial ingredient of this framework.

7.2 Arabic Case and subjecthood

Arab Grammarians explain Arabic Case putting forward either the GOVERNMENT MODEL (‘AMAL) or the PREDICATION MODEL (’ISNĀD). Taking apart their technical implementation, which does not concern us here, both these models share a central idea, which was implemented in its fullest form by the Late Grammarian al-’Astārābāḏī in his *Šarḥ al-Kāfiyah* (on which, see Section 2.3.5) and can be very sketchily outlined as follows (see Bohas et al. 1990, Ch. 3 and Guillaume 1998 for details).

On the level of form, NOM realizes what is a predicational constituent (‘UMDAH) on the level of meaning, ie the Subject and the Predicate, while ACC, OBL are the phonological manifestation of constituents that semantically are non-predicational (or FAḌLAH: direct object, indirect object, adjuncts). Two points of this conception of Arabic Case are of special relevance for the purposes of the present work.

First, in modern terms, Arabic (putative) NOM on the Predicate is not a primitive entity, but a by-product of its relationship with the Subject carrying the same marker. The Predicate, in fact, inherits from the Subject whatever marker (e.g. NOM) by virtue of an agreement process, where the former is the TARGET and the latter the CONTROLLER[2].

Second, the notion of non-predicational constituent (*faḍlah*) neutralizes the difference between ACC and OBL. It seems convenient to remark in this connection that neutralization on the level of meaning at least in some cases is matched by neutralization on the level of form, to the effect that in CA the putative ACC-ending *a* can function as OBL as well and, the other way around, the putative OBL-ending *i* can function as ACC.

The former phenomenon is observed in the so-called DIPTOTIC DECLENSION, the latter in the SOUND PL of the Triptotic Declension, that is when the putative case-endings co-occur with the morphemic material *-na* (cp. Section 6.6). This state of affairs is illustrated in Table 7.1 (where *’usamat-* means ‘lion’).

Underlying Arab Grammarians’ conception of Case then is a *binary* system, where the *u*-

Table 7.1: Neutralization of ACC, OBL

‘Case’	Tript., SG	Tript., SOUND PL	Diptotic
NOM	l-mu’min- u	l-mu’min- ū-na	’usamat- u
ACC	l-mu’min- a	l-mu’min- ī-na	’usamat- a
OBL	l-mu’min- i		

ending, signalling the Subject, is opposed to both the *a*- and *i*-endings, signalling *any constituent other than the Subject*. This opposition is semantic but in some declensional paradigms it has a phonological correlate too and its distinctive feature is (presence vs. lack of) Subjecthood. Table 7.2 represents this scenario. To put it shortly:

(657) In CA the putative case-system is actually a binary opposition between *u* (NOM), as Subject vs. *a, i* (ACC, OBL), as non-Subject.

[2]Arab Grammarians were aware of the agreement process, but they did not regard it as involving an asymmetric relation between the Subject and the Predicate. Instead, they thought of it as symmetric: to describe this phenomenon, in fact, they coined the term TARĀFU‘, which literally means ‘reciprocal assignment of NOM’ (Bohas et al., 1990, p. 69).

Table 7.2: Arabic case and the Subjecthood-based opposition

Semantics	(Pseudo-)case endings	Received view
Subject	u	'NOM'
Non-subject	a, i	'ACC, OBL'

7.3 Arabic case and Hamitic

It is not necessary for the purposes of this dissertation to elucidate all the details of Hamitic case, but it is of the utmost importance to look at how it behaves when found in a predicational context.

In Hamitic an overt copula tends to mark its predicate with the ABS(OLUTIVE) case, which is typically realized as *a* (Lipinski, 1997, pp. 259-262):

(658) Copula + Predicate^{ABS}

An interesting case in point is Lybico-Berber, an archaic subgroup of languages (Lipinski, 1997, p. 36):

(659) Tarifit (Lipinski, 1997, p. 262)

Muḥnd ḏ a-mqrran
Muḥend is ABS-great

'Muḥend is great'

This example shows that presence of ABS on the Predicate creates a contrast between this latter, obligatorily marked for ABS and the Subject, *not obligatorily* marked for the same case:

(660) ABS^{Subject}: NO / YES + Copula + ABS^{Predicate}: *NO / YES

This holds not only for Lybico-Berber, but also for Somali: see Frascarelli & Puglielli 2006 and immediately below.

From a syntactic vantage point, it seems safe to maintain that ABS in Libyco-Berber arguably still keeps traces of an original prepositional status[3]. The argument is built as follows.

Adjectives in Lybico Berber are better understood as nouns (Bentolila, 1981, p. 44)[4] and hence in (659) the 'adjectival' Predicate is actually a noun, to the effect that ABS precedes a noun.

In so doing, Lybico-Berber ABS behaves as a preposition, since Lybico-Berber is a VSO language and its ABS is aligned with the general Head > Dependent pattern à la Hawkins (see Comrie 1981, Ch. 4 and references therein), where the word order V > O is paralleled by the word order Preposition > Noun. This is illustrated in the Table 7.3 (data from Lipinski 1997).

Thus, in Lybico-berber there occurs the following subpattern of (658):

(661) Copula + Prepositional ABS + Noun / Predicate

Remarkably, Najdi Arabic manifests a state of affairs strikingly similar to the Lybico-Berber subpattern (661).

[3] On the prepositional origin of case-endings, see Blake (2001, Ch. 6).

[4] 'Pour ma part, je n'ai rien trouvé dans leurs compatibilités ou dans leur comportement syntactique qui puisse les distinguer des autres noms.' Once again, an interesting connection between Hamitic and Arabic: see Section 3.2.2 on the equation Noun = Adjective in Arabic.

Table 7.3: The prepositional status of ABS *a*

Head	Dependent
Verb	Object
Preposition$_i$	Noun
Case$_i$	Noun

This is shown in (662). Although at first sight in this sentence no prepositional ABS precedes the noun / predicate, nonetheless the negative *copula ma* marks it with the *preposition b-*[5]:

(662) Najdi Arabic (Ingham, 1994, p. 45)

ma anta b-mirtāḥ
are.not you P-comfortable

'you are not comfortable'

Given that Najdi Arabic is a VSO language (Ingham, 1994, p. 37), and that its preposition precedes the noun, the same Head > Dependent pattern observed in Lybico-Berber (Table 7.3) also occurs in this Bedouin Variety: V > O and P > N, including ABS^P > N.
Notice also that the Somali RC, *qua* predicate of a Hamitic instantiation of secondary predication, does not make exception to the general pattern (660): Frascarelli & Puglielli (2006) explicitly state that the Somali non-restrictive RC signalled by the overt copula *oo* is marked with ABS, which in this construction has a null phonological realization (cp. (329) in Section 5.2.3).
In passing, when ABS in Somali has an overt realization, it is found on the right of the noun it refers to, in line with the general Head - Dependent pattern of this language: Somali, in fact is an (S)OV language (Frascarelli & Puglielli, 2006, p. 308 and references cited there) and this is why in Somali the general pattern (660) does not manifest itself as the subpattern (661).
Returning to Arabic, traces of ABS seem to be found not only in primary predication (cp. (662)), but also in secondary predication. Consider again the OA RC type (49) illustrated in Chapter 3 and repeated below as (663):

(663) Old Arabic (Wright, 1896, vol. II, p. 232)

bāb-u l-ṣaġīr
gate-NOM the-little

[5]The item *ma*, found in all the MCV (Ferguson, 1959a), as well as its OA / CA equivalent *mā* usually work as negation rather than as negative copula (= negation + copula). Nevertheless, evidence for interpreting *ma(:)* in (662) as a negative copula comes from OA, where this item is commutable with the negative copula *laysa*. This latter, in fact, along with the negation *lā* are felt by OA native speakers as items able to replace *laysa* by means of the Substitution test (*mušabbahatāni by-laysa*: Wright 1896, vol. II, p. 104) and, furthermore, *ma(:)* cannot be used as simple negation in the imperatives. Notice also that in OA *mā* marks its predicate not only with the cognate of Najdi Arabic *b-*, ie *bi-*, but also with OBL. 'Case'-marking, however, looks suspicious in this context. Rabin (1969) in fact remarks that:

> "In fact, in Semitic *-i* as case-marker is largely redundant, as the genitive relation is widely characterized by special forms of the *nomen regens* (construct state). Nor does it fulfil any function after prepositions, in contrast to Indo-European, where prepositions and cases interact. In fact, a case could be made out for treating the *-i* as a non-case."
> (Rabin, 1969, pp. 194-195)

'the little gate'

According to Arab Grammarians, the modifiee *bāb(-u)* of (663) can be marked for OBL iff and only if the appropriate semantico-syntactic context is met, while its modifier *l-ṣaġīr* is marked for OBL regardless of such a context and no convincing explanation has been found so far for this phenomenon (Pat-el, 2009). The behavior of the modifier in (663), however, straightforwardly follows from the Language Contact Hypothesis entertained here, which invokes a pre-Semitic parastrate in accounting for this OA RC type (see Chapter 6).
In this light, native speakers' feeling that the modifier of (663) is always marked for OBL amounts to saying that it is a predicate *obligatorily* carrying a marker that cannot be *obligatorily* carried by the modifiee / Subject *bāb(-u)*. This results precisely in the same contrast between the Subject and the Predicate depicted in (660), characterizing Lybico-Berber (659), Najd Arabic (662) and Somali (329).
On the whole, both Arabic primary and secondary predication (cp. (662) and (663) respectively) show some relics of pre-Semitic ABS, but with a crucial difference in this respect. While in primary predication (662) ABS is detectable on the level of both form (*b-*) and meaning (*b-not* having the locative, instrumental etc... value typically associated with it), in secondary predication (663), ABS can be inferred from semantic considerations only, and lacks a phonological correlate[6].
Given the deep link between Arabic predication and preposition, it seems convenient to provide a more in-depth discussion about their syntactic interplay.

7.4 Arabic case and prepositions

The Arabic preposition when occurring in the context of predication has an interesting behavior. Not only can it precede the predicate, as observable in the negative sentence (662), but it has also the ability to precede the copula, as revealed by a closer look to the possessive sentence in part discussed in Section 6.6.2.
The possessive sentence is considered to be the manifestation of a predicational relationship after Benveniste (1960), according to which two are the ways to express possession in primary predication crosslinguistically. The core of the former is the complex copula plus preposition, that of the latter the verb *to have*. In the wake of Benveniste (1960), they are referred to as *être-à*-construction and *avoir*-construction and exemplified in (664) and (665) respectively:

(664) French (Ouhalla, 2000, p. 221)

La voiture est à Jean
the car is to Jean

'the car is Jean's'

(665) French (Ouhalla, 2000, p. 221)

Jean a une voiture
Jean has a car

'Jean has a car'

[6] A refinement of this account will be offered throughout this Chapter.

Benveniste (1960) must be also credited as the first to offer a transformational account of the relationship between the *être-à*-construction and the *avoir*-construction, since he conceives the *avoir*-construction as an *être-à*-construction affected by Predicate Inversion and takes *avoir* to be made up - on the meaning side - of the same copula and preposition which in *être-à* are realized also on the sound side ('*avoir* n'est rien autre qu'un *être-à* inversé').
In this light, the Possessee (say *la / une voiture* of (664), (665)) is the real Subject of the possessive sentence, the Possessor its Predicate (e.g. *Jean* in (664), (665)), to the effect that the distinctive property of the *être-à*-construction vs. the *avoir*-construction is the initial position of the Subject in the former vs. its final position in the latter.
Heine (1997) and, within the generative tradition, Ouhalla (2000) improve Benveniste's account, remarking that crosslinguistically the *être-à*-construction and the *avoir*-construction differ in that the former requires an indefinite Possessee, the latter does not (contrast *une voiture* of (665) with *la voiture* of (664)). The diagnostic properties opposing the two kinds of possessive construction are summarized in Table 7.4.

In addition, theory-internal reasons lead Freeze (1992) and Kayne (1994, Ch. 8) to assign

Table 7.4: Possessive constructions: diagnostic properties

Type	Possessor	Possessee
être-à-Cstr.	sentence-final	definite
avoir-Cstr.	sentence-initial	indefinite

Benveniste's semantic equivalence *avoir* = *être-à* a syntactic (antisymmetric) representation where the preposition precedes the copula, to the effect that *avoir* = *être-à* is reformulated as *avoir* = *à-être*.
The exact syntactic characterization of the *être-à*-construction as a whole and of its derivational history within the Dynamic Antisymmetry framework falls beyond the scope of its work and is left for future research. But it must be pointed here that assuming the syntactic representation *avoir* = *à-être* (technically, P-incorporation into I) is in principle compatible with this framework: Ouhalla (2000, pp. 230-231) derives Moroccan Arabic possessive sentence along the lines of Moro (2000), albeit he sticks to the Chomskian checking theory in motivating movement.
For the purposes of the present work the just discussed (dynamic) antisymmetric implementation of Benveniste's idea can be schematized as in (666), (667):

(666) *être-à*-Cstr.: Possessee^{Definite} + Copula + P + Possessor

(667) *avoir*-Cstr.: Possessor + P + Copula + Possessee^{Indefinite}

Focusing on (Moroccan) Arabic, the diagnostic properties typical of the *être-à*-construction (666) are observed in a kind of *be*-sentence featuring the preposition *dyal* 'of' as illustrated in (668, 669) (cp. also Section 6.6.2):

(668) Moroccan Arabic (= 607)

had l-ktab dyal Nadia
this the-book of Nadia

'This book is Nadia's'

(669) Moroccan Arabic (Ouhalla, 2000, p. 239)

*shi ktab dyal Nadia
some-book-or-other of Nadia

'*Some book is Nadia's'

The presence of a (covert) copula in (668) is evidenced by its paraphrase in (670), which exhibits insertion of the disambiguating element *huwwa*. Recall in fact from Section 6.6.2 that (668) can be mistaken for a definite FS (Possessive Classifier Construction) and that therefore *huwwa* functions as a disambiguating *copula*, not as a pronoun (cp. also Ouhalla 2000, p. 240 for the same view):

(670) Moroccan Arabic (= 609)

had al-dar hiyya dyal-t Omar.
this the-house.F huwwa.F *dyal-at* Omar

'This house is Omar's'

On the other hand, the diagnostic properties typical of the *avoir*-construction (667) are observed in sentences featuring the preposition *'and* 'at' combined with the pronominal base *h-*, as exemplified in (671). Ouhalla (2000, pp. 227-228) reports that the same sentence, if displaying a *definite* sentence-final constituent, albeit grammatical, does *not* have a possessive reading: rather the entire expression conveys a meaning of temporary location. (672) exemplifies this state of affairs:

(671) Moroccan Arabic (Ouhalla, 2000, p. 228)

Nadia 'and-ha wld
Nadia at-her boy

'Nadia has a boy (= Nadia is a mother / *A boy is at Nadia's place)'

(672) Moroccan Arabic (Ouhalla, 2000, p. 228)

Nadia 'and-ha l-wld
Nadia at-her the-boy

'Nadia has the boy with her (= *Nadia is a mother / A boy is at Nadia's place)'

Or, more abstractly:

(673) *avoir*-Cstr.: Possessor + 'and-Pronoun + Possessee^{Indefinite}

Both the sentence-initial position of the Possessor (e.g. *Nadia* of (671)) and indefiniteness of the sentence-final Possessee (e.g. *wld* of (671)) thus qualify the Moroccan Arabic sentential type (671) as an instance of *avoir*-construction (667). This sounds a little bit odd, because one at this point expects to find either the equivalent of the verb *avoir* or the complex preposition plus copula (*à-être*), along the lines of (667), but what is actually found is the complex preposition plus pronominal base *h-* (673).
This problem evaporates as soon as *the pronominal base h- is taken to be a copula*. Evidence

for this claim is twofold. First, in Moroccan Arabic the *avoir*-construction and the *être-à*-construction are identical not only semantically, but also formally, in the sense that they share not only one and the same set of *substantive* constituents (Subject, Predicate), as already noticed in French, English etc... (cp. (664, 665)), but also *one and the same set of functional constituents*: the preposition and the pronominal base *h-*. Accordingly if this latter in the *être-à*-construction (*hiyya* of (670)) is regarded as a copula, so might plausibly be in the *avoir*-construction (671), where it surfaces as *-ha*.
Second, the pronominal base *h-* in (671) *cannot* be the anaphoric pronoun of an alleged sentence-initial Topic / Possessor (*Nadia* in (671)), because Moroccan Arabic *avoir*-construction cannot be derived by a basic word order via topicalization. The non-topicalized counterpart of the *avoir*-construction in Moroccan Arabic, in fact, is not grammatical if assigned a possessive reading, the only allowed interpretation being the locative one (Harrell 1962, p. 210 and Harrell et al. 1965, p. 210)[7]:

(674) Moroccan Arabic (Harrell et al., 1965, p. 210)

had l-bent ʕand-ha seṭṭašel weṛqa
this the-girl at-her sixteen tickets

'This girl has sixteen tickets'

(675) Moroccan Arabic (Harrell et al., 1965, p. 210)

*ʕand had l-bent seṭṭašel weṛqa
at this the-girl sixteen tickets

'*This girl has sixteen tickets'

Once that the sentence-initial Possessor of (671, 673) cannot be thought of as a Topic and *h-* as its (anaphoric) pronoun, the only alternative left is analyzing *h-* as a copula, so that in the Moroccan Arabic *avoir*-construction (673) *the preposition precedes a copula, not a pronoun.* Accordingly, (673) is more accurately restated as (676):

(676) *avoir*-Cstr.: Possessor + 'and-Copula + Possessee^{Indefinite}

Technically, this is an instance of P-incorporation into the copula, in the spirit of Baker (1988), Freeze (1992) and Kayne (1994, Ch. 8). It should be added that whatever the rationale behind this phenomenon, it has nothing to do with the possessive meaning of the *avoir*-construction (671, 673), because the *être-à*-construction (666, 668), which conveys the same meaning, is not affected by P-incorporation into the copula.
Rather, the 'precopular preposition' is a phenomenon having to do with what from Benveniste (1960) onwards is regarded as the distinctive property of the *avoir*-construction (671, 673) opposing it to the *être-à*-construction, namely Predicate Inversion (cp. (672)). In short:

(677) In Arabic, when the copula mediates the relationship between a Predicate and a Subject, the preposition, if needed, precedes the copula.

That is, (673) is rewritten as (678):

(678) P-incorporation: Predicate + P-Copula + Subject

This intermediate conclusion will get back into play in Section 7.6.

[7]More generally, inability of the *avoir*-construction to have a non-topicalized counterpart is also found in Egyptian Arabic (Woidich & Heinen-Nasr, 2004, p. 52).

7.5 Arabic Case and dialectal variation

Arab Grammarians' considerations about CA Case are deeply rooted in their view that CA is a mixture of the Hijazi and Tamimi (= Bedouin) branches attested in OA, already discussed in Section 2.2.2.6.

Thus, they account for different instances of CA case-marking in terms of the dialectal variation existing in OA between the Hijazi subgroup and the Bedouin one, as dealt with at the end of Section 2.2.2.4.

Given that the distinction between the Hijazi and Bedouin branches has been preserved till today (see Section 2.2.2.6 for an accurate list of their distinctive isoglosses throughout the time), it seems convenient to look at the distribution of the putative 'case-endings' within the present-day manifestation of this old dialectal dichotomy.

The sample of MCV taken into account by Owens (2006, pp. 102-104) shows that the *u*-ending ('NOM') occurs in the Hijazi Branch, while *a*- and *i*-endings ('ACC, OBL') in the Bedouin one. More accurately, the context in which they occur is that between a noun and *-n* (see Section 6.6): Hijazi N-*u*-*n* vs. Bedouin N-*a*-*n*, N-*i*-*n*.

It seems, however, that the same dialectal opposition holds valid for the syntactic structure that Owens (2006) takes to be the original context of occurrence of 'case-endings': the construction noun plus modifier (this is not surprising under the current approach, which takes *-n* as part and parcel of a more complex (opaque) modifier: see Section 6.6).

Some dialectological data corroborate the just made claim. In Tihami Arabic, an *Hijazi* Colloquial Variety spoken in the Tihama region (see Section 2.2.1.2) and by the Tihami migrant community of Obock (Gibuti), *u* still intervenes between the noun and the modifier, as reported by Mion (2009, and references therein):

(679) Tihami Arabic (Mion, 2009, p. 217)

nās-u katīr-u
people-*u* many-*u*

'many people'

Notice that this is *not* a NOM case-ending because it is found in all the syntactic positions, e.g. it marks the noun playing the object-role:

(680) Tihami Arabic (Mion, 2009, p. 218)

ḥidt biss-u
I.saw cat-*u*

'I saw a cat'

On the other hand, in Najdi Arabic, the *Bedouin* variety *par excellence*, Ingham (1986, p. 283) reports that the vowel intervening between a noun and an adjectival modifier adorned with the article *l-* is either *a* or *i*, but never *u*: the northern varieties (e.g. Dhafīr dialect) display the former, the southern varieties (e.g. Murra and Ajmī dialects) the latter[8] . (681) and (682), respectively, exemplify this scenario:

[8] An important source of knowledge for the the Ajmī dialect is Johnstone (1961), which somewhat misleadingly labels it as 'Dosiri': see Ingham (1986) for details.

(681) Najdi Arabic, Dhafīr dialect (= 44)

(al-)bēt al-ṭuwīl
(*al-*)house *al*-tall

'the tall house'

(682) Najdi Arabic, Ajmī dialect[9] (Johnstone, 1961, p. 266)

(il-)rajjāl-ēn il-zen-īn
(the-)man-DU the-good-PL

'the two good men'

This is not the only opposition involving the construction noun plus modifier, since Hijazi Arabic (e.g. Yemeni Arabic) also differs from Bedouin Arabic in that the former uses *llaḏī* as RC marker (Vicente, 2006, p. 71), while Bedouin Arabic follows the majority of MCV outside the Arabian peninsula (Ingham, 1994; Vicente, 2006; Ferguson, 1959a) in making use of the RC marker *illi*, with the variants *alli, halli, yalli* (see e.g. Ingham 1982b, Ingham 1986).
Focusing on the RC markers *yalli, halli*[10], they are bimorphemic, being made of an item *ya-*, *ha-* of demonstrative origin and of *-lli* (Barth, 1913, p. 159). As for their syntax, they do not occur in Nedroma Arabic, where *(e)lli* acts as a Conj copula (see Section 6.3.2):

(683) Nedroma Arabic (= 481)

l-mtirqae$_i$ elli b-āš$_i$
the-hammer *elli* with which

(684) ... [$_{\text{ConjP}}$ l-mtirqae [$_{\text{Conj}'}$ elli [$_{\text{CP}}$ b-āš ...

(685) *l-mtirqae ya-lli / ha-lli b-āš
the-hammer *ya-lli* *ha-lli* with which

nor are they found in OA / CA, where the putative NOM-ending-*u* too is a Conj copula (see Section 6.4.1):

(686) Old Arabic / Classical Arabic (= 455)

l-malik-u llaḏī
the-king-NOM *llaḏī*

(687) ... [$_{\text{ConjP}}$ l-malik [$_{\text{Conj}'}$ -u [$_{\text{CP}}$ llaḏī ...

(688) *l-malik-u ya-llaḏī / ha-llaḏī
the-king *ya-llaḏī* *ha-llaḏī*

Furthermore, *yalli, halli* are typical of those dialects in which the item *lli* is a C, because it *cannot* be immediately followed by the (constituent containing) the relative pronoun (cp. the diagnostics worked out in Section 6.3.3). Contrast Nedroma Arabic, where *ya-, ha-* do not occur and *lli* is a Conj copula immediately followed by the (constituent containing) the relative pronoun (684), with Levantine Arabic where *ya-, ha-* occur and *lli* is a C not immediately followed by it, as illustrated in (689):

[9] See fn. 8
[10] The variants *alli, illi* are discussed at the end of this Section.

(689) Levantine Arabic (Cowell, 1964, p. 415)

l-ḥādse$_i$ yə-lli b-aʕrəf-ha$_i$
the-incident-*at* ya-lli PRS-I.know-her

'the incident I know'

Impossibility of co-occurrence between the Conj copulae *-lli* (of Nedroma Arabic), *-u* on one side and *ya-, ha-* on the other is distributional evidence enough to analyze these latter as Conj copulae. Accordingly, a minimalist representation of (689) is (690)

(690) ... [$_{\text{ConjP}}$ l-ḥādse [$_{\text{Conj}'}$ yə [$_{\text{CP}}$ -lli ...

Indirect evidence for this claim is typological: demonstratives crosslinguistically share with conjunctions a key structural property, namely the ability to work as a copula mediating the relationship between a Subject and its Predicate (Stassen, 2005), and Arabic does not make exception to this claim: indeed, the ancestor of *(e)lli* acting as a Conj copula, namely the Hijazi Arabic *d̲ī*, is a demonstrative base performing this task. It ensues that the Bedouin *ya-, ha-* are Conj copula as much as the Hijazi *d̲ī* is.

It also ensues that the dialectal group Bedouin Arabic is opposed to the dialectal group Hijazi Arabic for the contrast in the selection both of the morphemic material *a, i* vs. *u* (whatever its nature) and of the Conj copula *ya-, ha-* vs. *d̲ī* (recalling from the previous Sections that this latter may evolve into C via *la-*prefixation and *d̲-*deletion). This is summarized in Table 7.5.

Before concluding, three remarks are in order. First, the just discussed opposition in the choice

Table 7.5: Morphemic material intervening between the noun and its modifier

Dialect group	(Pseudo-)case endings	Conj copula
Hijazi	u	(lla$_{Assev}$-) d̲ī
Bedouin	a, i	ha, ya (-lli$_C$)

of the RC marker has clearly to do with the ability for Hijazi Arabic to tolerate and then to preserve the non-euphonic cluster *l(V)d̲ / d̲(V)l* (457), and for the inability for Bedouin Arabic to do so. It is thus expected that Hijazi Arabic and Bedouin Arabic behave the same way in other sectors of grammar.

This expectation is met, as shown by their demonstrative systems: *already in OA*, as reported by the Early Grammarian al-Farrāʾ[11] (*apud* Fleisch 1961, vol. II, pp. 45-46) the Hijazi group makes use of the distal form *d̲alika* 'that', displaying the non-euphonic cluster*d̲(V)l*, while the Bedouin group employs the distal form *d̲āka*, where no such a cluster is found:

(691) Old Arabic (Fleisch, 1961, vol. II, pp. 45-46)

Hijazi: d̲alika 'that'
Bedouin: d̲āka 'that'

Second, there is a slight asymmetry in the syntax of the Conj copula *d̲ī, ha-, ya-*: while the latter works as a marker of Predicate Inversion, the two former cannot. The argument is built as follows. Ingham (1994, p. 30) reports that in the Najdi verbs of saying the direct object usually precedes the indirect object:

[11] On which, see Section 2.3.5.

(692) Najdi Arabic (Ingham, 1994, p. 30)

gilt-ih-li-hum
I.said-it-to-them

'I said it to them'

He adds that it is also possible to scramble this word order, making the indirect object precede the direct object and that in this case the item *iyyā* shows up between the indirect object and the direct object following it:

(693) Najdi Arabic (Ingham, 1994, p. 30)

gilt-li-hum-iyyā-h
I.said-to-them-*iyyā*-it

'I said it to them'

Recently, den Dikken (2006, pp. 20-22 and references therein) subsumes the complex direct object - indirect object *(V) X to Y* under the structure Subject - Predicate: his reasoning, briefly put, is that underlying this complex is the possessive construction *X (belongs) to Y*. From this viewpoint, (692) is an instance of Canonical Predication, while (693) is an instance of Inverted Predication. As for the item *iyyā*, following Barth (1913, p. 23) and Fleisch (1961, vol. II, p. 20) it is a variant of the demonstrative base *ya-*, inserted in the structure for phonological reasons, namely as a dummy support for the personal pronoun.
This latter in fact is a clitic in Arabic, by definition in need of a morphological item to be tacked onto (see Section 6.6.2).
Whatever the functional Head hosting the item *ya-* in Syntax (and later forced by A-P to *overtly* realize it), it is sufficiently clear that in (693) such a functional Head acts as a copula intervening between a Predicate and its Subject. On the other hand, Hijazi copular *d̲ī* (as well as the Bedouin *ha-*) cannot replace *ya-* in this function[12]:

(694) *gilt-li-hum-ha-h, *gilt-li-hum-d̲ī-h

This diagnostic property draws a sharp dividing line between *d̲ī, ha-* on one side and *ya-* on the other, which leads to revise the account given so far: while it is still possible to maintain that the two former are Conj copulae, the latter, albeit a copular element, cannot be longer seen as Conj.
This leads to wonder which is the exact nature of *ya-* between the Noun / Subject and the RC Predicate in Najdi Arabic etc... The answer provided here on the basis of Kern (2010)'s syntactic diagnostics mentioned at the beginning of Section 5.2.3 (p. 211) is that:

[12]It is temptatively suggested here that the morphological host of the clitic pronoun cannot surface as *d̲ī* or *ha-* for phonological reasons. As for *d̲ī*, in (694) it potentially gives rise to the non-euphonic cluster *l(V)d̲* (457), given that the indirect object is typically combined with the preposition *li-*. As for *ha-*, it potentially gives rise to a non-euphonic cluster *h(V)h* if combined with the third person pronominal base *h-*. Simplifying, this is due to the fact that two identical adjacent consonants are not tolerated in Arabic (see Greenberg 1950, p. 383). Notice that the base *ya-*, if combined with the first person pronoun *-ya* 'me' gives rise to the complex *iyyā-ya* and hence to the non-euphonic cluster *y(V)y*. Nonetheless, while the cluster *h(V)h* is not found in OA and CA, following the Greenbergian data pool, the cluster *y(V)y* does occur in them: the imperfective forms *yay'isu* 'he is in despair about' and *yaybisu* 'he dries up' are attested from Sībawayhi onwards (*apud* Fleisch 1961, vol. II, p. 260).

(695) *ya-* is a copula marking Predicate Inversion because it is a functional item at the same time able to intervene between a Predicate and a Subject and unable to introduce a sentential object.

Table 7.6 schematizes extension of Kern (2010)'s diagnostics from Thai, Burmese etc... to Arabic. Recall from Section 5.2.5 that a copula marking Predicate Inversion can also surface

Table 7.6: Diagnostic properties of the copula marking Predicate Inversion

Language	Item	Between Predicate > Subject	Cp.	Sentential C	Cp.
Thai, Burmese	thîi	YES	(311)	NO	wā (p. 224)
Arabic	ya-	YES	(693)	NO	'inna, 'anna... (p. 198)

between a Noun / Subject and a non-restrictive RC / Predicate in the so-called Reinstated Word Order, as first observed by den Dikken & Singhapreecha (2004). Therefore interpreting *ya-* as such a kind of copula forces the following intermediate conclusion:

(696) The Arabic (non-restrictive) RC displaying *ya-* is an instance of Reinstated Word Order

so that in all likelihood the Najdi Arabic RC depicted in (698) is the Arabic equivalent of the Thai RC displaying the copular item *thîi* (421), replicated below as (697):

(697) Thai (= 421)

[rôm] thîi [jàj săam khan nán]^{RC}
umbrella *thîi* big three CLF this

'those three big umbrellas'

(698) Najdi Arabic, Murra dialect (Ingham, 1986, p. 285)

[yā jazl al-'ata] yalli [b-labs il-kāmtīn adrāi]^{RC}
O abundant in.the-giving *yalli* of-hidding the-concealers aware

'O Generous, who are aware of the things hidden by the concealers'

as well as of French *une pizza de chaude* (as opposed to the non-Reinstated Word Order *une pizza chaude*: for *de* as a copula of Predicate Inversion, cp. *un imbécile de garçon*, as per den Dikken (2006); see also Sections 9.1.1 and 9.2.1).
Finally, a remark concerning the phonological realization of *ya-* is in order: Barth (1913, p. 90) lists as one of its variant the item *'i*. This is a morphologically conditioned form, which *ya-* takes when suffixed to *another demonstrative base*:

(699) Tripolitanian Arabic (Barth, 1913, p. 90)

hād-ulā-yä / hād-ulā-'i
this-these-*yä* / this-these-*i*

'these'

Such a variant of *ya-* will prove useful in due course[13].
Looking at the results of this Section from a global perspective reveals that the MCV, which

[13]In comparative terms, the account of the Arabic particle *ya-* given so far is neutral as to whether it is genetically connected with the Amharic particle *yä* dealt with in Section 5.3.1 or not. Nonetheless, the analysis of Arabic *ya-* as copula of Predicate Inversion proposed here owes much to den Dikken (2006)'s claim that Amharic *yä* performs the same function (his Linker).

represent an old state of affairs, exhibit the same putative case-endings found in CA, but they do not perform any semantic function. Rather their choice is a matter of dialectal variation and remarkably this latter is a *binary* opposition, where the Hijazi *u* (the so-called NOM) is opposed to the Bedouin *a, i* (the so-called ACC, OBL).
The same dialectal variation is characterized by another *binary* opposition that helps to clarify the precise function of the morphemic material *u, a, i*.
On one side, in fact, in the MCV there is found a binary opposition in terms of Predicate Inversion: while both the Hijazi branch and the Bedouin one manifest Canonical Predication, via the dedicated copulae *d̲ī, ha-* respectively, the Hijazi Branch does not allow Predicate Inversion, the Bedouin one does, via the dedicated copula *ya-* (see Table 7.5).
On the other side, lack of Predicate Inversion distributionally correlates with the marker *u* in the Hijazi branch, presence of Predicate Inversion with the markers *a, i* in the Bedouin one. As for the former correlation, this is totally expected in the light of the findings of the previous Chapter, where it has been shown that its phonological and syntactic properties qualify it as a copula of Canonical Predication (Conj). As for the latter correlation, it strongly suggests that the analysis of *ya-* as a copula of Predicate Inversion might be plausibly extended to *a, i*:

(700) the putative case-endings *a, i* (ACC, OBL) are copulae signalling Predicate Inversion and hence the construction Noun > a / i > [C - Modifier]^{RC} is an instance of Reinstated Word Order

where the construction:

(701) Noun > a / i > [C- Modifier]^{RC}

is that exemplified in (681, 682) (see previous Chapter for the equations *l-* = C and Adjectival modifier = RC).
Syntactic evidence in favor of this claim comes from Najdi Arabic, where at least the item *a* occurs in secondary predication in a clear context of Predicate Inversion (word order Predicate > Subject), thus matching one of the two diagnostic criteria set by Kern (2010) (cp. Table 7.6 above). This is the construction that Ingham dubs TRANSPOSED NP STRUCTURE and describes as follows:

> "Certain of the above structures can be transposed with no change of meaning for reasons which are stylistic in nature and difficult to give an actual semantic denotation of...It seems sensible to regard the second noun as the head although in a normal construct [state] the first noun would be regarded as head".
> (Ingham, 1994, p. 59)

The non-transposed NP structure is the modification structure in (702)

(702) Najdi Arabic (Ingham, 1994, p. 59)

wasm-i-n mbaccir
winter.rain-*i*-C early

'early winter rain'

Notice that the modification structure at issue exhibits an instance of *-n* plausibly to be interpreted as a (copular) complementizer because it is followed by an *overt* modifier, actually the

RC introduced by it (see end of previous Chapter).
The transposed NP structure simply inverts the order of constituents (and *l-* is plausibly an RC-final complementizer, replacing the complementizer *-n*):

(703) Najdi Arabic (Ingham, 1994, p. 59)

mbaccir a-l-wasim
early *a*-C-winter.rain

'early winter rain'

As just seen, Ingham explicitly states in his description of (703) that *mbaccir* is a modifier of the head noun *wasim* (in the Bloomfieldian sense, not in the X' one), so that the word order in which *a* intervenes is Predicate > Subject. In so doing, *a* is *in sharp contrast* with *u*, which in CA does not have the ability to intervene between a Modifier / Predicate and a Modifiee / Noun. Recall in fact from Section 3.2.2.1 and (53) that the CA equivalent of (703), ie the modification structure Adjective > Noun, is always interpreted by Arab Grammarians as an instance of CS whose head noun is the adjective, which they regard along the lines of the Indo-european substantivized adjective (*the best of...*):

(704) Classical Arabic (Wright, 1896, vol. II, p. 218)

ṣ̄aliḥ-u l-'axlāq
appropriate-u (of) the-manners

'the best of manners'

(705) illustrates the contrast at stake:

(705) mbaccir a l wasim = Modification = Dependent + a + l + Head
ṣ̄aliḥ u l 'axlāq = CS = Head + u + l + Dependent
'early winter rain'

Thus, the ability of *a* vs. the inability of *u* to manifest itself within the Inverted Predication structure in secondary predication (703), Ingham's 'transposed NP structure', is a diagnostic property convincing enough to prove that *a* and, generalizing, *i* are copulae signalling Predicate Inversion.
It is worth observing in this connection that the construction (701) by virtue of its structural make-up, namely copulae of Predicate Inversion *a, i* intervening between the Noun and the postnominal RC, is reinterpreted as an instance of Reinstated Word Order (cp. (698)).
Ingham (1994) in the just mentioned passage confirms this interpretation, because he observes that the construction (701) is *semantically equivalent* to its 'transposed' counterpart. The syntactic evidence discussed so far therefore demonstrates (700).
The Reinstated Word Order (701) includes not only (681, 682), but also other modification structures, whose complementizer is *-n* (e.g. (702)) and *lli*: in particular this latter when combined with *a, i*, gives rise to the RC marker *alli, illi* found in Najdi Arabic (see beginning of this Section).
It should be added that *a, i* of *a-lli, i-lli* in this Colloquial Variety have been interpreted as morphemic vowels (copulae) rather than as epenthetic ones because of their distributional correlation with the copula *ya-*, and therefore if this correlation is not found they cannot be

regarded as such, being epenthetic vowels instead. This is the case of the vowels *e, i,* ə preceding *lli, ddi* in Egyptian Arabic, Nedroma Arabic, Djidjelli Arabic etc...
Last but not least, a beneficial side-effect of identifying *a, i* with copulae of Predicate Inversion is the following:

(706) In the MCV the dialectal variation is actually a binary opposition between *u* ('NOM'), as copula of Canonical Predication vs. *a, i* ('ACC, OBL'), as copula of Predicate Inversion.

The careful reader will have noticed that the MCV opposition between *u* vs. *a, i* in terms of NO vs. YES Predicate Inversion (706) matches that found in CA between *u* vs. *a, i* in terms of YES vs. NO Subjecthood (657).
In greater detail, the CA Subject semantic role develops out of the (OA /) MCV Canonical Predication structure, the CA non-Subject semantic role out of the (OA /) MCV Reinstated Word Order (itself a subtype of Inverted Predication structure).
Given that (OA /) MCV reflect an old state of affairs, while CA stands out as an innovation (see Chapters 2, 3), the question arises of why and how the former binary opposition (706) turns into the latter (657). This is schematized in Table 7.7.
The next Section builds a phonological argument that further corroborates the hypothesis in

Table 7.7: Putative case-endings: from OA / MCV to CA

Form	Meaning				
	Revised account				Received view
	OA / MCV	Branch		CA	
u	copula (Can.Pred.)	Hijazi	(⇒)	Subject	'NOM'
a, i	copula (Inv. Pred.)	Bedouin	(⇒)	Non-subject	'ACC, OBL'

(700). Section 7.7 explains the development from the Predicate-Inversion based opposition to the Subjecthood-based opposition.

7.6 The fine-grained structure of the copulae *u, a, i*

Additional evidence for the hypothesis in (700) is phonological and is culled from OA.
It can be shown, in fact, that *a, i* occur in a phonological context that according to OA native speakers qualifies these vowels as a *phonetic* representation and implies for them a richer *phonological* representation, in which the copula of Predicate Inversion *ya-* is found. More generally, this will result in assigning the Arabic copulae *u, a, i* more fine-grained syntactic structures, which turn out to share in their core one and the same pattern. The argument is built as follows.
The Agglutination Hypothesis à la Petracek (1965), Zaborski (1976), Voigt (1997), Voigt (1999), Tropper (2004), Garbini & Durand (1994) has it that an abstract representation of *i*, when acting as PL putative case-ending, is decomposable into two distinct *morphs*, namely *w, i* (see (148) in Section 7.10).
The equivalence *w* = Conjunction Copula set in the previous Chapters leads to generalize their idea out of the phonological representation of PL to that of SG, and therefore to the putative ending *a* as well, because this latter occurs in the paradigm of SG along with *i*.

It now seems convenient to look at the phonological context the putative case-endings *a, i* occur in: it is made of a C(onsonant)[14] preceding them and of a sequence CCV following them.
This happens because (putative) case-endings originally arise in a modification structure, where they are tacked onto the OA pausal form of the modifiee, always C-final (see Owens 2006, Ch. 4 and Ch. 3) and the modifier is an RC introduced by the items *llaḏī, l-*, as discussed at length in the previous Chapter.
Notice in particular that prefixation of *l-* to the modifier gives rise to a sequence CCV since all words in OA must be CV-initial (CCV = *l*CV), CC-initial words being barred (a constraint dubbed by Arab Grammarians 'ADAM ILTIQĀ' AL-SĀKINAYN: see e.g. Fleisch 1961, vol. I, pp. 205-206). This is schematized in the following[15]:

(707) ...C $V_{a,i}$ lCV...
...C $V_{a,i}$ lla...

Or, more abstractly:

(708) ...C $V_{a,i}$ CCV...

This is precisely the phonological context that Arab Grammarians, from Sībawayhi onwards, take to be the result of (at least) two rules, namely GLIDE CONTRACTION and LONG VOWEL-SHORTENING, an account recently revived in Kouloughli (1979) and Angoujard (1990, p. 66ff.) (see Owens 2005 and fn. 17 for details).
On their view, *a*, when found in (708) is the phonetic representation (*far'*) of the phonological representation (*'aṣl*) *awi*, while *i*, in the same context, is the phonetic representation (*far'*) of the phonological representation (*'aṣl*) *awi* (see Kouloughli 1979, Veccia Vaglieri 1937, pp. 168-169 as well as Fleisch 1961, vol. II, pp. 364ff., Bohas et al. 1990, Ch. 4 and the primary sources cited there)[16].
. In both the cases, the phonological representation turns into the phonetic representation passing through (at least) one intermediate stage: a long vowel (*ā, ī*, respectively)[17]. This is exemplified by the so called HOLLOW VERB (or VERBUM MEDIAE RADICALIS WĀW, FI'L 'AJWAF WĀWĪ):

(709) (yan)qawidna > (yan)qādna > (yan)qadna 'we are resigned'
(CVC)CawiCCV > (CVC)CāCCV > (CVC)CaCCV

(710) quwiltu > qīltu > qiltu 'I said'
CawiCCV > CīCCV > CiCCV

It is worth observing in this connection that the phonological analysis made by Sībawayhi, a non-native speaker of OA, is already found in its embryonic stage in the work of his teacher al-Xalīl, an OA native speaker (see Ch. 2). This latter, in fact, at the end of his introduction to

[14]This abbreviation should not be confused with the C used in the previous Sections, where it stands for complementizer. By the same token, the abbreviation V, for Vowel, should not be confused with the V standing for Verb.

[15]The same phonological context obtains if OA *-lli, n-* are considered instead of *llaḏī, l-*, respectively.

[16]See also Section 2.3.3 for more details on Arab Grammarians' notion of *'aṣl* as well as on its modern equivalent(s) and more generally on Arab Grammarians' *'aṣl*-based model.

[17]More accurately, Arab Grammarians regard Glide Contraction as a more complex process, decomposable into further (sub)rules. See Kouloughli (1979), Angoujard (1990, p. 66ff.) and Bohas et al. (1990, Ch. 4) for details. This does not affect the point made here. For an overview of the phonology of (some) modern Colloquial Varieties, see also Watson (2002).

the *Kitāb al-ʿAyn*, explicitly states that the underlying form of a verb displaying a long vowel has a glide in it (see Haywood 1965, p. 36). In greater detail, the change *(yan)qawidna* > *(yan)qādna* (709), *quwiltu* > *qīltu* (710) is due to a rule of Glide Contraction observed in other cells of the paradigm, as illustrated in (711, 712):

(711) (yan)qawidu > (yan)qādu 'he is resigned'
(CVC)CawiCV > (CVC)CāCV

(712) quwila > qīla 'he said'
CawiCV > CīCV

Glide Contraction can be given the following formulation, where parentheses indicate that according to Arab Grammarians this process takes place both in a phonological environment C_CV and C_CCV (e.g. *quwila* and *quwiltu* above, respectively):

Phonological Rule 7.1 (Glide Contraction)**.**
awi ⟶ *ā* /*C_(C)CV*/
uwi ⟶ *ī* /*C_(C)CV*/

The change *(yan)qādna* > *yanqadna* (709), *qīltu* > *qiltu* (710) is due to a rule of Long Vowel Shortening, illustrated in (713, 714). The Late Grammarian al-Zamaxšarī (*apud* Fleisch 1961, vol. I, p. 165) exemplifies this rule as follows:

(713) (lan yaḍri) bālya (wma) > (lam yaḍri) balya (wma) 'he is not going to hit today'
(...) CāCCV (...) > (...) CaCCV (...)

(714) (lan taḍri) bībna (ki) > (lam taḍri) bibna (ki) 'you (F) will not hit your son'
(...) CīCCV (...) > (...) CiCCV (...)

(715) (lan yaḍri) būl'ā (na) > bul'ā (na) 'they are not going to hit now'
(...) CīCCV (...) > (...) CiCCV (...)

In these examples, Long Vowel Shortening takes place in the context C_CCV. Another instance of application of the same rule is the context C_C and (716) is illustrative of the behavior of all the long Vs *ā, ī, ū*:

(716) (lam ya) ṣīr > (lam ya) ṣir 'he did not become'
(...) CīC > (...) CiC

What these contexts have in common is that the C following the long V is not followed by a V. This rule is schematized in the following:

Phonological Rule 7.2 (Long Vowel Shortening)**.**
ā ⟶ *a* /*C_C(CV)*/
ī ⟶ *i* /*C_C(CV)*/
ū ⟶ *u* /*C_C(CV)*/

The discussion so far shows that at least in OA / CA both Glide Contraction and Long Vowel Shortening are changes conditioned by one and the same environment C_CCV, which is exactly the phonological context in which the putative case endings *a, i* occur (cp. (708)). This motivates the hypothesis entertained here that these latter derive from *awi, uwi* respectively. In particular, the contrast between the short V of the SG phonetic representation *a, i* and the

long V of its PL counterpart *ā, ī* is due to the just mentioned Long Vowel Shortening rule: the SG *a, i* are followed by a 'demarcative'-*n*, which create the context for application of this rule, while the PL *ā, ī* are followed by *-na*, a sequence *CV* inhibiting its application.
It should be remarked at this point that Arab Grammarians contend that the sequences *awi, uwi* (*'aṣl*) underlying *a, i* (*far'*) do exist in the speaker's mind *in synchrony*, while they deny that such forms represent an historical stage of the language (see Section 2.3.3).
Thus, native speakers' judgments about (709, 710) rule out the possibility that *awi, uwi* underlying *a, i* in the context C_CCV are reconstructed forms and evidence that they are phonological representations. Since the same sequences also underlie the putative case endings *a, i* in the same context (708), they are to be interpreted in the latter way, not in the former.
It is important to remind at this point that following Chekayri (2005) the phonological representations *awi, uwi* of (709, 710) are *not* the departure point of the phonological process yielding *a, i*, because they exhibit the glide *w*, which is the result of a process of Glide Creation (see Section 6.3.2).
Since one of the main and most reliable sources in which OA putative case endings *a, i* are documented is Koran (see Section 2.2.2.4) and that in this source the sequences *VwV* appear to be a strategy to eliminate the glottal stop (or HAMZAH, here transcribed as '), it seems reasonable to maintain that Chekayri's Glide Creation in this case targets an input phonological representation where the *hamzah* is found, in order to delete it.
It is a well known fact among Arabists that in OA 'the most celebrated feature of the Hijaz dialect is the disappearance of the *hamza* or glottal stop' (Rabin, 1951, p. 130). Rabin himself, however, remarks that elimination of *hamzah* is also found in the Tamimi (= Bedouin) branch: 'In the Tamim dialect **yar'ä* (Hebrew *yir'eh*) 'he sees' became *yarä* as elsewhere (Sībawaihi, ii, 37) and they dropped final *hamza* after a vowel in pause (*ibid.*, 311)' (Rabin, 1951, p. 131).
With this clarification, it is now possible to look at how sequences *VwV* are themselves outputs of input forms involving *hamzah*. Following (Early) Arab Grammarians OA speakers resorted to three strategies in order to eliminate *hamzah* from their speech, strategies on the whole designated by (Early) Arab Grammarians as TAXFĪF AL-HAMZAH (Fleisch, 1961, vol. I, pp. 102ff.).
The first is its WEAKENING (HAMZAT BAYNA BAYNA) and is of no concern here. The second is HAMZAH-MUTATION ('IBDĀL) into a glide and, interestingly, when affecting a sequence *u'i* turns it into the sequence *uwi* dealt with above, as explicitly stated in the *Kitāb al-Taysīr fī l-Qirā'āt al-sab'* (*apud* Fleisch 1961, vol. I, p. 100) by 'Abū 'Amr 'Uṯmān Ibn Sa'īd al-Dānī (d. 444 / 1053). This is a work in the variant reading tradition (see Section 2.2.2.4) reporting that underlying to the OA phonetic representation *sīla* is the phonological representation *su'ila*, via the intermediate form *suwila*:

(717) su'ila > suwila > sīla 'he was asked'
Cu'iCV > CuwiCV > CīCV

That is:

Phonological Rule 7.3 (Hamzah-mutation)**.**
' $\longrightarrow$ *w /u_i/*

Notice that the form *su'ila*, not documented in the Hijazi dialectal group, is reported by Early Arab Grammarians to have been heard among Tamimi (= Bedouin) speakers of OA

(Fleisch, 1961, vol. I, p. 105).
Finally, the third strategy is HAMZAH-DELETION and consists of erasing a *hamzah* preceded by a sequence CVC and followed by a V. Interestingly, the V and the C immediately preceding *hamzah* can also be the V *a* and the glide *w*, so that Sībawayhi (*apud* Rabin 1951, p. 133, Fleisch 1961, vol. I, p. 107) reports for the Hijazi dialectal group not only the change *yas'alu* > *yasalu* 'he asks', but also:

(718) ḥaw'ab > ḥawab 'large valley'
Caw'V(C) > CVwV(C)

The following rule describes the change at issue:

Phonological Rule 7.4 (Hamzah-deletion)**.**
'⟶ 0 /CVC_V/, e.g.
'⟶ 0 /Caw_V/

What precedes is taken to be evidence enough to assign the morphemic material *a, i* the phonological representation *aw'i*, *u'i*, respectively. Associating a semantic value to such phonological sequence straightforwardly follows from the findings of the previous Sections.
In fact, *u / w* found in *aw'i*, *u'i* is the same copula of Canonical Predication (Conj) also found in the putative NOM and its presence is motivated by the fact that *aw'i*, *u'i* (surfacing as *a, i*) intervene between a Subject and its Predicate in the modification structure, thus mediating the predicational relationship between them.
As for the *a* preceding the copula of Canonical Predication *u / w* in the phonological representation *aw'i*, it can be easily identified with the relic of ABS discussed in Section 7.4, for two reasons.
Consider again (701), rewritten as (719, 720), where the phonological representations *aw'i, w'i* replace its phonetic representations *a, i* respectively:

(719) Noun > *a* + *w* + *'i* > [C- Modifier]^{RC}

(720) Noun > *w* + *'i* > [C- Modifier]^{RC}

Section 7.4 has shown that in Arabic, given a context of Predication Inversion, the functional item immediately preceding the copula and fused with it is a preposition (cp. (678)) and the item *a* preceding the copula *w* in (719) manifests itself in the same context, since (719) is an instance of Reinstated Word Order, a subtype of Predicate Inversion (cp. the Najdi Arabic transposed NP structure (703)).
Accordingly, the *a* of *aw'i* in (719) is a preposition, but the ABS *a* documented in Hamitic has exactly the same categorial status (see Table 7.3), so that the former plausibly is to be identified with the latter.
The equation *a* (of *aw'i*) = ABS *a* also explains why the Reinstated Word Order (701) manifests the alternation *aw'i*, / *u'i* (surfacing as *a, i*): this is the Arabic reflex of a Pre-semitic state of affairs, where ABS is realized both as *a* and zero (see Section 7.3).
What is left is an item *'i* which can be easily identified with the suffixal variant *'i* of the demonstrative base *ya-*, provided that the appropriate morphosyntactic context is met: namely, the host of the suffix must be a demonstrative base (cp. (699)).
Although it is true that the item *u / w* cannot be immediately characterized as such, because it

is a Conj(unction), nonetheless conjunctions crosslinguistically overlap with demonstratives in a key structural property, namely the ability to work as a copula (Stassen, 2005). This situation of functional overlapping especially holds valid for Arabic, where both a demonstrative base (*d̲ī*) and a conjunction (*u / w*) perform the same copular function, as discussed at length in Ch. 6.

On such a basis, the Conj copula *u / w* provides an appropriate morphosyntactic context for the suffixal *'i* variant of *ya-* by virtue of its *rapport associatif* with the demonstrative *d̲ī*[18], to the effect that the item *'i* found in the phonological representation *aw'i, w'i* of *a, i* turns out to be a manifestation of *ya-*.

This comes as no surprise in the light of the fact that the construction displaying the (complex) copulae of Predicate Inversion *aw'i, w'i* (699) is an 'adjectival' RC type in complementary distribution in Najdi Arabic with the 'non-adjectival' RC type displaying *ya-* (in the sense that in this Colloquial Variety the complementizer *lli* is not combined with adjectives, and the complementizer *l* is not combined with the 'familiar'RCs).

The structural makeup of the phonological representation of *a, i* is summarized in (721, 722), while the parallelism between the Reinstated Word Order featuring *ya-* and that featuring *a, i* is depicted in (723):

(721) Noun + a + w + 'i + Modifier
Noun + ABS + Conj + *ya-* + RC

(722) Noun + 0 + w + 'i + Modifier
Noun + ABS + Conj + *ya-* + RC

(723) (= 698, 44)

a. [yā jazl al-'ata] ya- [lli b-labs il-kāmtīn adrāi]^{RC}
b. [al-bēt] a- [l-ṭuwīl]^{RC}

The alternation between the *non-adjectival* RC type featuring *ya-* and its 'adjectival' counterpart featuring *'i* once again can be traced back to the Language Contact Hypothesis, and it can be assumed that *ya-* belongs to the Amorite parastrate, *'i* to the Pre-semitic one.

Empirical support in favor of this hypothesis is twofold. First and foremost, Testen (1998, Ch. 3) considers the (complementizer) *l-* and the asseverative particle *la-* as etymologically connected, and in line with Garbini & Durand (1994), takes *l-* to be of (Pre-)semitic origin and to be older than the Amorite *la-* because of its phonological shape. The distinctive feature at issue is the lack vs. presence of vowel: the Pre-semitic *l-* is vowelless, the Amorite *la-* is vowelled.

Extending Testen's phonological argument from the alternation *l- / la-* to that *'i / ya-* reveals the Pre-semitic origin of *'i* and the Amorite origin of *ya-*.

The second piece of evidence lies in the domain of occurrence of *'i, ya-* and confirms the previous finding: the former co-occurs (in its phonological representation) with both of the Pre-semitic item ABS *a* and the Pre-semitic Conj copula *w*, the latter co-occurs with the Amorite item *lli* in the RC marker *ya-lli* (see e.g. (689), (698)). To put it shortly:

[18]More precisely, the dynamic antisymmetric account has it that the Pre-semitic POS in which *u / w* is involved gets the Demonstrative feature, via analogy with the Amorite POS where *d̲ī* occurs. See Section 6.4.2.

(724) The Bedouin Arabic RC is characterized by a copula of Predicate Inversion. When it instantiates a Pre-semitic RC type, the RC is 'adjectival' and the copula at issue is (underlyingly) *'i* (while surfacing along with other morphemic material as *a, i*). When it instantiates an Amorite RC type, the RC is non-adjectival and the copula at issue is *ya-*

Section 7.8 below further elaborates on this point.

Finally, it seems safe to maintain that the same alternation *a-* vs. *0* in realizing ABS is found not only in the phonological representation of the putative ACC, OBL, but also in that of NOM. Evidence for this hypothesis lies in the fact that OA native speakers feel that the modifier / Predicate of the archaic RC type in (663) is marked for an invariable 'OBL', highly reminiscent of the invariable Hamitic ABS marking the same kind of constituent (see Section 7.3). Their judgment is given irrespective of whether the copula intervening between the modifiee and the RC modifier is a copula of Canonical Predication (the putative NOM) or a copula of Inverted Predication (the putative ACC, OBL) (for example, in (663) itself the intervening copula is 'NOM').

Therefore, native speakers' semantic characterization of the OA RC type (663) evidences the presence of an ABS *a* in the phonological representation of the copula of Canonical Predication *u* ('NOM'). Since this latter is no more than ACC, OBL minus Predicate Inversion, its phonological representation is *aw* (see above and also fn. 21 for more evidence in favor of this claim).

The validity of this claim rests on a phonological argument: plausibility of the Glide Contraction $aw > \bar{u}$ (for the Long Vowel Shortening $\bar{u} > u$, see (715) above). As a matter of fact, this rule is documented in OA, at least in one of the two phonological environments it is expected to occur, namely C_CV. The Late Grammarian Ibn Mālik (d. / 1274, *apud* Rabin 1951, p. 165, fn. 15), in fact, states that the phonological representation of the OA form *'ūlād* 'boys' is *'awlād*[19].

On these grounds, the phonetic representation of the Conj copula *u / w* can be assigned the following phonological representation, as follows:

(725) Noun + a + u + Modifier
Noun + ABS + Conj + RC

For the contrast between the short V of the SG phonetic representation *u* and the long V of its PL counterpart *ū* the same remarks made above about the phonetic opposition *a, i* vs *ā, ī* apply.

To summarize this Section, it brings out three main results. First, the phonological analysis converges with the syntactic one in identifying *a, i* (ACC, OBL) with copulae of Predicate Inversion.

Second, the *morphosyntactic* opposition *u* vs. *a, i* in terms of lack vs. presence of Predicate Inversion has been given a *morpho-phonological* correlate, since lack of Predicate Inversion is lack of the morph *'i*, while presence of Predicate Inversion is presence of the morph *'i*. Third, the study of the phonological representation of the phonetic representations *u, a, i* unveils - in compliance with the Agglutination Hypothesis pursued in this work - that their syntactic structures is more fine-grained than assumed so far (cp. (721), (722), (725)) and, furthermore, that these phonological representation share one and the same functional complex, namely:

[19]Ibn Mālik also contends that such a form belongs to the Tamimi (= Bedouin) dialectal group, but the Early Grammarian al-Farrā', himself an OA native speaker, reports the word *mūtān* 'death' without ascribing it to any dialectal group (see *Lisān al-'Arab*, p. 4296, s.v. *M W T*, ll. 26-28)

(726) ABS: (a) + copula of Canonical Predication w

The parallelism between u and a,i is schematized in the following (where italics indicate the phonetic representation):

(727) u = (a) + w ... - 'i ⇔ Predicate Inversion: NO
a, i = (a) + w ... + 'i ⇔ Predicate Inversion: YES

The syntactic structure of the copulae u, a, i outlined in (727) will be derived within the framework of (Dynamic) Antisymmetry in the next Chapter.
Last but not least, notice that the presence of ABS in (727) is consistent with Owens (2006)'s claim that OA is caseless because syntax of ABS makes it clear that it still has a prepositional nature.

7.7 Deriving Arabic Case via Predicate Inversion

In Section 7.5 it has been hypothesized that the opposition lack of Predicate Inversion vs. presence of Predicate discussed in the previous Section evolves in CA into an opposition Subject vs. non-subject.
The question therefore has arisen of why and how such a development taks place. It is put forward here that the key of the development at issue is the different informational profile associated with the Canonical Predication structure (lack of Predicate Inversion) and with the Reinstated Word Order (presence of Predicate Inversion), as already discussed in Section 5.2.5.
Recall from Section 5.2.5 that in Canonical Predication (say, (351)) the noun qualifies the entity it refers to as old information, while its modifier is a piece of new information for the listener. On the other hand, in the Reinstated Word Order the entity the noun refers to is new information, while its modifier represents a piece of old information. This is due to the fact that the Reinstated Word Order passes through a derivational stage of Predicate Inversion, which is typically characterized by the just described informational profile (cp. (352) and (350)).
Given that the noun of the modification structure characterized by Canonical Predication represents old information and that denotation of old information is a crosslinguistic property of the sentential Subject (Comrie, 1981, Ch. 5), there is no doubt that the noun at issue can function as Subject.
On the other hand, the noun of the modification structure characterized by Reinstated Word Order represents new information, which in principle rules it out as a sentential Subject. Nonetheless, as discussed at length in Ch. 3, any modifier associated with a noun makes this latter specific, that is endowed with a certain degree of definiteness, if not even definite.
But definiteness, following Comrie (1981, Ch. 5), is another property enabling a noun to work as a Subject. This scenario leads to predict that the very presence of extra-material 'rescues' the noun denoting new information, allowing it to act as a Subject. The prediction is realized: in Tagallog, the official language of Philipine a modification structure displaying the Reinstated Word Order does work as Subject, as observed by Schachter & Otanes (1983, pp. 60, 75).
This is exemplified in the sentence (728), whose Subject is an instance of Reinstated Word Order because it freely alternates with the modification structure illustrated in (729), a clear instance of Inverted Predication, signalled by the particle *-ng* ('either the adjective-linker-noun order

or the noun-linker-adjective order may be used, without any apparent difference in meaning': Schachter & Otanes 1983, p. 122):

(728) Tagallog (Schachter & Otanes, 1983, p. 122)

magara ang [baro-ng pula]^{TOPIC}
attractive a / the dress-*ng* red

'the red dress is attractive '

(729) Tagallog (Schachter & Otanes, 1983, p. 122)

magara ang [pula-ng baro]^{TOPIC}
attractive a / the red-*ng* dress

'the red dress is attractive '

The overall scenario is that the most suitable candidate for Subjecthood is a modification structure instantiating Canonical Predication because of the oldness of its noun: C-I (the semantic component) simply needs to look at it to legitimate the syntactic structure.
On the other side, the modification structure exhibiting the Reinstated Word Order is a less suitable candidate for this role because of the newness of its noun. Thus C-I qualifies such a noun for Subjecthood only subsequent to an additional step: a repair strategy that legitimates the syntactic structure, looking not only at the noun, but also at the extra material combined with it.
Plainly, computation for C-I is more complex in the latter case than in the former. This is why it seems safe to maintain that C-I prefers Canonical Predication over Reinstated Word Order for the semantic role of Subject for reasons of economy.
Although a more accurate formulation of this semantic version of Economy Principle falls far beyond the purposes of this work, the strict solidarity between Syntax and C-I (see Section 4.3 and Chomsky 2008) strongly suggests that such a principle is to be understood here along the same lines of the syntactic Economy Principle discussed in Section 5.3.3: the lower the number of semantic operations that build a semantic structure, the more economical the semantic structure. This is depicted in the following:

(730) Canonical Predication: Subjecthood = Noun (step 1)
Reinstated Word Order: Subjecthood = Noun (step 1), Modifier (step 2)

Capitalizing on this observation, it is claimed here that when the OA dialectal variants *u* (Hijazi) and *a, i* (Bedouin) are mixed into a single language, ie CA, this latter has to cope with a redundant set of endings, which is reduced by specializing these variants in either a Subject- or Non-subject role according to their informational profile.
Therefore, the Hijazi modification structure marked by the copula *u*, having *not* undergone Predicate Inversion, represents old information and hence Subjecthood more transparently, prominently and easily for computation than the Bedouin one.
This latter, marked by the copula *a, i* although able to work as Subject *faute de mieux* (cp. the Tagallog example in (728)), has undergone Predicate Inversion and then has a certain degree of newness of information, which makes it a less suitable candidate for the function of Subject. This happens because C-I can enable the noun conveying new information to act as a Subject only carrying out an additional effort, namely looking at the modifier of the noun at issue.

In consequence of this, the Hijazi variants turn into Subject-markers and the Bedouin variants turn via polarization into non-Subject markers, as illustrated in Table 7.7. To sum up:

(731) Arabic case-endings *u(:), a(:), i(:)* are better understood as markers of an opposition Subject *u(:)* vs. non-Subject *a(:), i(:)*. This semantic opposition is originally a dialectal opposition between the Hijazi *u(:)* and the Bedouin *a(:), i(:)*, where these items are copulae intervening between a noun and its modifier. Development of Hijazi *u:* into a Subject role and of Bedouin *a(:), i(:)* into a non-Subject role is a consequence of their informational profile. While the Hijazi *u:* does not occur in a modification structure having undergone Predicate Inversion, the Bedouin *a(:), i(:)* do. But a modification structure thus characterized is less suitable to function as a Subject because of the newness of information encoded on its noun: crosslinguistically the Subject represents old information. Thus, when the Hijazi and the Bedouin branch fuse into a single language (CA), the modification structure containing *u:* is preferred for the Subject role over that containing *a(:), i(:)*, which accordingly polarize into the non-Subject role.

(731) answers the question (IV)(see beginning of previous Chapter).

7.8 Language Contact Hypothesis and dialectal variation

This Section focuses on the interplay between the two major 'dialectal' oppositions (in a broad sense) invoked in the present work: Hijazi vs. Bedouin and Pre-semitic vs. Amorite.
Previous sections have plausibly shown that the opposition Hijazi vs. Bedouin is based on lack vs. presence of Predicate Inversion.
However, this is not the only distinctive feature characterizing this dialectal opposition, because the Hijazi and the Bedouin branches also differ in in their treatment of the demonstrative base *d̲-*: Arab Grammarians report its loss of deictic force (pleonastic or 'otiose' *d̲-*) as typical of the Bedouin branch, not of the Hijazi one (see end of Section 6.4.3).
Accordingly, another distinctive feature of the opposition at issue is lack vs. presence of a desemantized, ie pleonastic or otiose, *d̲*[20].
Therefore a bird's eye view to the dialectal opposition Hijazi vs. Bedouin reveals that lack of the desemantized *d̲-* correlates with lack of Inverted Predication (Hijazi branch), and the other way around (Bedouin branch), as schematized in the following:

(732) Hijazi: desemantized (otiose) *d̲*: NO ⇔ Predicate Inversion: NO
Bedouin: desemantized (otiose) *d̲*: YES ⇔ Predicate Inversion: YES

As for the opposition Pre-semitic vs. Amorite, Section 7.6 has plausibly shown that the Bedouin branch preserves traces of it when manifesting a twofold RC system, one which opposes the 'adjectival' RC type (723.b) to its non-adjectival counterpart (723.a), as summarized in (724). The bulk of evidence lies in the morpho-phonological shape of their copulae: the adjectival and *vowelless 'i* (cp. the Pre-semitic *l-* co-occurring with it) vs. the non-adjectival and *vowelled ya-* (cp. the Amorite *la-* of *lli* co-occurring with it).
Turning to the Hijazi branch, it is likely to exhibit the same state of affairs, since it displays the same opposition adjectival vs. non-adjectival in its RC-system (contrast (679) at beginning of

[20]This statement admittedly involves a certain amount of abstraction and holds valid for the diachronic stage that precedes diffusion of this feature from the Bedouin branch to the Hijazi one: cp. fn. 20.

Section 7.5 with (466) at beginning of Section 6.3.2) and in this latter the adjectival RC type is signalled by the Pre-semitic RC marker *w* (cp. Somali *oo*) absent in the non-adjectival type, which manifests the Amorite RC marker *d̲ī* instead:

(733) The Hijazi Arabic RC is characterized by a copula of Canonical Predication. When it instantiates a Pre-semitic RC type, the RC is 'adjectival' and the copula at issue is (underlyingly) *w* (while surfacing along with other morphemic material as *u*). When it instantiates an Amorite RC type, the RC is non-adjectival and the copula at issue is covert (and later surfacing as *d̲ī* subsequent to reanalysis of the relative pronoun).

It ensues that each of the terms of the dialectal opposition Bedouin vs. Hijazi in its turn includes a further dialectal opposition Pre-semitic vs. Amorite, at least insofar as the RC system is concerned:

(734)
- Hijazi RC
  - Pre-semitic RC
  - Amorite RC
- Bedouin RC
  - Pre-semitic RC
  - Amorite RC

Underlying to these dialectal oppositions are the following grammatical oppositions:

(735)
- Pred. Inversion: NO + Desemantized (otiose) *d̲*: NO
  - Adjectival RC: YES
  - Adjectival RC: NO
- Pred. Inversion: YES + Desemantized (otiose) *d̲*: YES
  - Adjectival RC: YES
  - Adjectival RC: NO

These oppositions can be also schematized as in Table 7.8 (where CPI stands for copula of Predicate Inversion).

The just outlined scenario of oppositions necessarily involves a certain degree of idealization:

Table 7.8: Distinctive properties of the Hijazi and Bedouin RCs

Branch	RC marker		Desemantized *d̲*	CPI *ya- / 'i*
	Adjectival Pre-semitic	Non-adjectival Amorite		
Hijazi	(a) w	lla*d̲*ī	NO	NO
Bedouin	(a) w 'i	ya-lli	YES	YES

for example, in the data pool examined in this work, the non-adjectival Amorite RC type in Najdi Arabic is always adorned with the prenominal article, whose phonological and syntactic structure is Amorite only in part (the underlying *d̲* is, the overt *l-* is not: see Ch. 6).

Yet, in its essence and at least insofar as the modification structure is concerned, it turns out to be quite an accurate picture of what the Language Contact Hypothesis is if applied to the situation of dialectal variation typical of OA and still surviving nowadays.

7.9 Refining the Classifier Hypothesis

A better understanding of the phonological and semantic 'atoms' involved in the Arabic copula traditionally referred to as case-ending makes it possible to clarify some aspects of the Classifier Hypothesis left vague so far: they are the relationship between the 'SOUND PL' Numeral Classifier Construction and its 'DU, SOUND PL' counterparts, as well as the relationship between the 'M' Numeral Classifier Construction and its 'F' counterpart.

7.9.1 Deriving dual

Ch. 3 brings evidence (p. 103) in favor of the hypothesis that the phonological representation of the OA (/ CA) DU-ending *(N-ā)-ni* is *(N-ā)-na*, so that the DU-ending develops out of the SOUND PL-ending. In particular, the phonological form *(N-ā)-na* surfaces as such in some OA speeches that resist the phonological process responsible for the gap between the phonological and the phonetic representations.

For the sake of completeness, it should be added that Wright (1896, vol. I, p. 236) reports that the SOUND PL itself in some dialects of OA can undergo the same *a*-to-*i* change when preceded by the (putative) OBL ī, as shown by forms like *'axar-ī-ni* for the more common *'axar-ī-na* 'others'. Hence, the behavior of DU allows to assign the OA dialectal form *'axar-ī-ni* the phonological representation *'axar-ī-na* documented in other OA dialects.

Under the approach developed here, the pretheoretical fact that DU and SOUND PL *-na* turn into *-ni* when preceded by the copulae *ā, ī* points to a Bedouin origin of this phenomenon. This is in fact the dialectal group *a(:), i(:)* originate from (see Table 7.5).

Such a state of affairs is totally expected because the change *na* > *ni* is no more than the phonological rule of Low Vowel Raising applying in the context of an open syllable (C_), which Western scholars observed to be typical precisely of the Bedouin branch in both ancient and modern times, as already discussed in Section 2.2.2.6 and schematized in Table 2.8. To sum up:

(736) DU: (Noun + ā + na_i) > Noun + ā + ni (cp. *Zaydāna*$_i$)
SOUND PL: Noun + ī + na (> Noun + ī + ni_j) (cp. *'axar-ī-ni*$_j$)

It has been also remarked in Section 3.3.2 that semantically DU is no more than a SOUND PL ranging from *two* to ten, rather than from *three* to ten (Depluralization Hypothesis): this state of affairs is adamantly clear in the archaic usage of DU known as PSEUDO-DU.

Archaicity of OA SOUND PL in *-āni* over that in *-īna* is proved, following Retsoe (1995), by the existence in Akkadian of SOUND PL endings highly reminiscent of the former but not of the latter. This is exemplified by contrasts like *išd-um* 'fundament' vs. *išd-ān* 'fundaments' where *-ān* is the SOUND PL ending (Retsoe, 1995, p. 189). Thus, both sound-side and meaning-side considerations point to a development SOUND PL > DU.

Retsoe (1995), the proponent of the Depluralization Hypothesis, establishes a direct link between the semantic split of SOUND PL into SOUND PL and DU and the new phonological opposition *ā* (DU) vs. *ī* (SOUND PL), claiming that the long V *ī* is a phonological innovation of the system, introduced to distinguish the otherwise ambiguous reading of *-na* as both (PAUCAL) SOUND PL and DU.

He also contends that the *ī* at issue is a morpheme originally denoting abstractness and evolving

into a SOUND PL in this new context (Deverbal Abstract Noun Hypothesis: see Section 3.3.2), but as a matter of fact no such development is documented in Semitic (*ibidem*), which makes Retsoe (1995)'s account questionable in this respect. Under the copula analysis advocated throughout the present dissertation, this scenario naturally follows.

First, the OA system resorts to *ī* to replace *ā* because the former has the same structural make-up of the latter (Reinstated Word order: see e.g. (727)).

Second, *ā* stands for a phonological representation *aw'i*, *ī* for a phonological representation *0u'i* (see (727)), which share a (complex) copular core *w + 'i* and differ only in the overt vs. covert realization of ABS: *a* vs. 0.

Third, ABS is clearly a relic, because in Arabic its occurrence is confined to two copular constructions: the modification structure (end of Section 7.6) and the negative *be*-sentence (cp. (662) in Section 7.3).

On these grounds, introduction of *ī* as new marker of PAUCAL SOUND PL boils down to eliminating the fossilized and non-functional ABS *a* from the phonological representation *awi*, assigning it a null value, in order to create an opposition between the DU and the PAUCAL SOUND PL based on the distinctive feature overt vs. covert ABS and at the same time to preserve the functional copular core *w + 'i* of the modification structure.

To put it differently, deletion of *a* is triggered by the need to avoid a non-isomorphic form-meaning relationship: one morph *-ā(-na)* for two sememes (*PAUCAL PL, DUAL*), so that this phenomenon is a manifestation of the so-called ONE FORM - ONE MEANING PRINCIPLE (Thornton, 2005, p. 165 and references therein)[21]. It ensues that OA (/ CA) DU evolves into PAUCAL SOUND PL along the following lines

(737) a(:)na = a + w + 'i + [n + 'a]
ABS + COPULA + *ya* + [MODIFIER]

(738) *aw'i* → Delete *a* → *0w'i*

(739) i(:)na = 0 + w + 'i + [n + 'a]
ABS + COPULA + *ya-* + [MODIFIER]

As for the MCV, they usually display *ān* for OA *āni* and *īn* for *īna*, and interestingly both these forms often preserve a PSEUDO-DU (hence paucal) meaning. They are attested in Negev

[21] The other side of the token is that under this analysis, deletion of *a* from *aw'i* does not apply freely and can be admitted only under precise conditions of readjustment of the system. On this view, there is no compelling reason to justify the absence of *a* in the underlying phonological structure of the SG and SOUND PL nominative endings *-u*, *-ūna*. Therefore the ONE FORM - ONE MEANING PRINCIPLE derives on independent grounds the same analysis of *u(:)* (= *aw*) proposed at the end of Section 7.6. The same principle predicts that the alternation *aw'i / 0w'i*, ie the putative alternation ACC / OBL, is a strategy to solve an original situation of non-isomorphism between one morph and two sememes, in compliance with its need of one form - one meaning. Since in the Depluralization Hypothesis the morph associated with two sememes is the (complex) copula *a(:)*, that is the copula whose phonological representation *aw'i* features the overt ABS *a*, also in the case of the alternation ACC / OBL, the morph originally associated with two sememes and subsequently yielding the alternation at issue is plausibly the putative ACC *a*, since it has the same phonological representation of *a(:)*. The alternation ACC / OBL therefore would be the result of the putative ACC *a* 'splitting' into ACC *a* and OBL *i* via deletion of the ABS *a* from its phonological representation (*aw'i* > *0w'i*) in order to achieve a condition of isomorphism (two morphs for two sememes). On this assumption, the so-called Diptotic Declension would be the ancestor of the Triptotic Declension (see Table 7.1 for their illustration). Such an account is compatible with the Language Contact Hypothesis à la Garbini, according to which the Diptotic Declension is older than the Triptotic one (Garbini & Durand, 1994, p. 100). This is admittedly a sketchy outline, to be substantiated in future research.

Arabic, a Bedouin Colloquial Variety spoken in the Negev region of Israel, as well as in Tlemcen Arabic (see Sections 3.2.1 and 3.3.2):

(740) Negev Arabic (Blanc, 1970, p. 48)

rijl-ā-n = rijl-ā-n *FEW*
leg-COP-CLF = leg-COP-CLF *FEW*

'legs (from 2 to 10)'

(741) Tlemcen Arabic

dər'-ī-n = dər'-ī-n *FEW*
arm-COP-CLF = arm-COP-CLF *FEW*

'legs (from 2 to 10)'

The fact that such forms clearly have a paucal meaning evidence that the paucal quantifier *'a* is part and parcel of their phonological representation. Phonologically, its loss in the phonetic representation is a consequence of the Low Vowel Raising rule discussed at the beginning of this Section. In fact *-na*, when turned into *-ni* becomes a word-final sequence *Ci* (cp. (736)), which in OA is the phonological context in which a rule of *i*-deletion can apply.
This rule was first observed by Early Arab Grammarian, e.g. Sībawayhi (*apud* Rabin 1951, p. 120), which recorded changes like *'irmi* > *'irm* 'throw':

Phonological Rule 7.5 (*i*-Deletion)**.**
$i \longrightarrow 0$ */C_#/*

Thus, under the approach adopted here, the gap between OA (/ CA) SOUND PL / DU *(V̄)nV* and MCV *(V̄)n* is explained in terms of *hamzah*-Deletion, Low Vowel Raising and *i*-Deletion of the grammaticalized quantifier *'a*:

(742) DU: Noun + ā + n'a > Noun + ā + na > Noun + ā + ni > Noun + ā + n
SOUND PL: Noun + ī + n'a > Noun + ī + na > Noun + ī + ni > Noun + ī + n

More abstractly:

(743) Noun + V̄ + n'a > Noun + V̄ + na > Noun + V̄ + ni > Noun + V̄ + n
Phonol. repr. > *hamzah*-Del. > Low Vow. Raising > *i*-Del.

To summarize the main finding of this Section:

(744) splitting of SOUND PL into SOUND PL and DU boils down to the alternation covert ABS - overt ABS *a* in their phonological representation

(744) answers the question (VII) (see beginning of the previous Chapter), since DU, while arising from SOUND PL via preservation of the overt ABS *a* in its syntactic structure / phonological representation (deleted in the 'new' SOUND PL), still includes in its syntactic structure / phonological representation the same functional item *-w-* (Agglutination Hypothesis) and the complex Classifier plus quantifier *-n(')a* (Classifier Hypothesis) also found in the (old and new) SOUND PL.

7.9.2 Towards an unified account of Arabic plurals

The just described process of *i*-deletion yields a phonetic representation characterized by the word-final sequences *(C)ān, (C)īn*, that is *(C)V̄C*. In this connection, Arab Grammarians contend that in OA (/ CA) such a sequence at the end of a word is not tolerated, so that *(C)V̄C* is shortened into *(C)VC*. This is the rule of Long Vowel Shortening discussed in Section 7.6.
What is brought out is an opposition between the MCV, Najdi Arabic included, tolerating the word-final sequence *(C)V̄C* and OA, which does not.
The Najdi Arabic piece of data set forth in (745), from Abboud (1979, p. 481), occurs outside the nominal domain and confirms this statement.

(745) (yi)jīb 'he brings'
(...)CīC

Yet, the strong solidarity between Najdi Arabic and OA well documented on phonological level (Section 2.2.2.6 and Table 2.8) makes it plausible that Najdi Arabic can also behave as OA, allowing Long Vowel Shortening albeit in a non-generalized way.
One then expects to find a word-final ending *-an / -in* performing the same function of *ān / īn*, namely denoting a paucal PL. By definition, in fact, a given change on the sound side (here, Long Vowel Shortening) keeps the meaning side unaltered, and hence plausibility of the former rests not only on its constant application (cp. regularity of Long Vowel Shortening in OA) but also on the immutability of the latter: in componential terms, this is association of *-in* with a (covert) feature *FEW*, responsible for the paucal semantics (cp. (740, 741)).
In the Najdi Arabic data pool of this work, paucal PL *-an* does not occur, while paucal PL *-in* does. This latter is found in the Colloquial Arabic spoken by the Sunni community of Bahrain, a variety descended from Najdi Arabic (Holes, 2004, p. 96).
The ending at issue is suffixed to the noun *nās* 'people' and the complex *nāsin* thus formed has a clear paucal meaning, as syntactically evidenced in its context of occurrence (746) by the fact that *nāsin* triggers on its verb(s) plural agreement (Strict Agreement), not default agreement in *-at* (Deflected Agreement).
Blanc, in fact, after examining a large sample of MCV, including Najdi Arabic (Blanc, 1970, p. 42, fn. 3), concludes that 'it can be shown that in at least several of the dialects where the alternation PC - FSC does occur, PC [= plural concord, ie Strict Agreement] applies to other 'enumerative' [= from 3 to 10, ie paucal] expressions that go together with the dual, while FSC [= feminine singular concord, ie Deflected Agreement] applies to 'non-numerative' plurals' (Blanc, 1970, p. 49)[22].
The paucal meaning of *nāsin* appears evident also from Holes's English translation, where it is rendered as 'some [people]':

(746) Bahrani Arabic (Sunni community) (Holes, 2004, p. 93)

nās-i-n raḥa-w, nās-i-n inzila-w
people-COP-*n* went-PL people-COP-*n* stayed.put-PL

'Some (of the Dawasir) left (Saudi Arabia), others stayed put'

Accordingly, *nāsin* of (746) is assigned the following (informal) semantic representation:

[22]See also Section 3.2.1 on Strict Agreement and on its relationship with the enumerative / paucal meaning.

(747) nās-i-n *FEW*
people-COP-CLF *FEW*
'some people'

Syntactically, *-in* in (747) is no more than the Numeral Classifier Construction characterized by the complex Classifier *-n* plus the quantifier *FEW* acting as a modifier of the noun *nās* with the intervening copula *-i-* (see previous Chapter). Phonologically, *-in* is the phonetic representation of the phonological representation *īn'a*, so that it has the same derivation of its cognate endings *(C)ān, (C)īn* plus an additional step, namely Long Vowel Shortening - contrast (743) with (748):

(748) Noun + ī + na > Noun + ī + ni > Noun + ī + n > Noun + i + n
hamzah-Del. > Low Vow. Raising > *i*-Del. > Long Vow. Shortening

It is instructive at this point to look at *nās* from the perspective of Early Arab Grammarians (and of the received view). They regard this word as the BROKEN PL of *'insān* 'man'. In particular, *nās* is usually analyzed as a BROKEN PL following Sībawayhi's etc... account, according to which it is to be assigned the phonological representation (*'aṣl*) *'unās* (see e.g. *Lisān al-'Arab*, p. 147, s.v. *' N S*).
This Section thus evidences at least in part the hypothesis that:

(749) the difference between (PAUCAL) SOUND PL and (PAUCAL) BROKEN PL is only a matter of phonological variation: (PAUCAL) BROKEN PL is no more than a (PAUCAL) SOUND PL having undergone Long Vowel Shortening.

Nevertheless, it should be pointed out - in order to avoid a 'logical matrix' - that the term SOUND PL is used for the sake of commodity and it stands for Numeral Classifier Construction (see previous Chapter).
The (partially justified) hypothesis in (749) is consistent with the traditional claim that the BROKEN PL diachronically develops out of a COLL and that the former is no more than the latter having undergone some process of affixation, lengthening, gemination (p. 136).
If in fact the BROKEN PL is a Numeral Classifier Construction, the noun found in this construction is indeed a COLL, in the sense originally ascribed to such a notion by Al-'Astārābāḏī (see Section 6.6): a bare noun having the ability to function both as SG and PL.
As for the processes of affixation, lengthening, gemination responsible for turning the COLL noun into a BROKEN PL, they can be straightforwardly derived by the presence of the suffixal *-in* tacked onto the COLL noun at issue.
The comparison between Arabic BROKEN PL and German apophonic PL (on which, see p. 239) is enlightening in this respect.
In the German apophonic PL, e.g English *feet*, the *apophonic* infixal V marking the PL (*ee*) is a former *metaphonic* V triggered by the presence of a PL suffix *-i*. The subsequent dropping of this latter results in the deceiving effect that the PL-marker is the infix *ee*, not the suffix *-i*:
foot + $\underline{i}^{PL}$ > *f*$\underline{ee}$*t*$\underline{i}$ > *f*$\underline{ee}^{PL}$*t*.
Simply put, at least originally the real *locus* of German apophonic PL is a *suffix, not the infixal vowel*:

(750) ... X... Y^{PL}# > ... X^{PL} ... 0#

Consider now (159), repeated here as (751):

(751) Old / Classical Arabic (Fleisch, 1961, vol. I, p. 309)

naml COLL: swarm of ants FORM: NO length. MEANING: NO count.
nimāl BROKEN PL: ants FORM: YES length. MEANING: YES count.

It is a standard assumption that the BROKEN PL, say *nimāl* of (751), develops out of the COLL, say *naml*, via a 'process of plural formation by *ā*-infixation' (Ratcliffe, 2006, p. 446). But the infixal PL-marker *ā* is a long V found on a word-final syllable and a long V in this position in Najdi Arabic is the result of suffixation: e.g. *hinā* > *hnā-šīn* (see (585)).
Interestingly, the word-final syllable of the BROKEN PL in Najdi Arabic undergoes suffixation of the PL-marker *-in* (see (746)), so that the *infixal* long V of *nimāl* can be taken to arise as a consequence of the presence of the *suffixal* PL-marker *-in* tacked onto the COLL *nam(a)l*[23]. The parallelism is depicted in the following:

(752) hinā > hnā - šīn
nam(a)l > nimāl - in

More generally, the state of affairs of Arabic BROKEN PL is highly reminiscent of that of German apophonic PL, in the sense that in the former, like in the latter, the PL-meaning of the *infixal* V (*ā*) is actually encoded in the suffix *-in*, subsequently dropped (contrast *nimāl* with *nāsin*), along the same lines of (750). This parallelism is outlined in (753)

(753) foot > feet - i^{PL} > fee^{PL}t
nam(a)l > nimāl - in^{PL} > nimā^{PL}l

Before concluding this Section, a final remark, more speculative in its nature, should be made. Arab Grammarians report a particular instance of OA / CA BROKEN PL pattern, namely *CawāCin*, which they say to be typical of a noun having a glide as its final root C.
For example, the root Cs of the noun *jāriy(ah)* 'servant (F)' are *J R Y* and therefore its BROKEN PL is *jawārin* (Fleisch, 1961, vol. I, p. 490).
In the spirit of Bohas (1997), one may temptatively assume that the root Cs of *jāriy(ah)* are just *J R*, and that *Y* is originally a morph having some functional meaning. One may also assume that in the BROKEN PL *jawārin* the root Cs *J R* were originally adjacent, very much as they are in its SG *JāRiy(ah)*.
If one assumes that this scenario is correct, the *-w-* intervening between *J R* in *jawārin* is the result of some version of a Metathesis rule involving this glide, already known by Arab Grammarians (Bohas et al., 1990, p. 21): the phonological representation of *JawāRin*, prior to metathesis, is *JaRāwin*.
Recalling that the BROKEN PL ending *-in* is actually a complex Classifier *-n* plus a covert quantifier *FEW* (cp. (747)), it is possible to single out in *jarāwin* a stem *jar* modified by the complex Classifier *-n* plus quantifier. Accordingly, the sequence *a(:)wi* intervening between *jar* and *-n* (*FEW*) can be interpreted, abstracting away from its lengthening, as the complex copula *-awi-* discussed in Section 7.6 and schematized in (737):

[23]The phonetic representation *naml* can be assigned the phonological representation (*'aṣl*) on the basis of Arab Grammarians' analysis: for example, the Late Grammarian 'Ibn Jinnī (on which, see Section 2.3.3) interprets the lack of vowel in OA / CA (*sukūn*) as a 'micro-vowel'(*ṣuwayt*) to be likened to the vowel *a*: see Grande (2003, p. 64, fn. 19) for the primary sources and for details. In a similar vein, a recent study on CA verbal apophony by Guerssel and Lowenstamm lead them to set the equivalence *sukūn* = *a* (Guerssel & Lowenstamm, 1996, pp. 131-132).

(754) jar + a + w + 'i + [n + *FEW*]
Noun + ABS + COPULA + *ya-* + [Modifier]

It is worth observing in this connection that in Hamitic the same (morpho)syntactic structure Noun > *awi* > *n* is found in some BROKEN PL, as illustrated in (150), here repeated as (755):

(755) Amazigh (Boukhris et al., 2008, p. 47)
tudm-awi-n 'petits visages'

These speculations, if on the right track, lend indirect empirical support in favor of the phonological representation *awi* assigned to the Arabic copula *a* of Predicate Inversion (the putative ACC and DU) in the previous Sections: at least in the just mentioned Amazigh PL pattern, the composite nature of the copula at issue is still detectable.
The comparison between Arabic SOUND PL, BROKEN PL and the Amazigh (Hamitic) PL, far from being implausible, calls into the picture not only the Language Contact Hypothesis (see Section 2.2.2.2 on the 'isogloss' BROKEN PL shared by both Hamitic and the Pre-semitic parastrate of Arabic), but also a strict application of the Agglutination Hypothesis à la Petracek (1965), Zaborski (1976), Voigt (1997), Voigt (1999), Tropper (2004), Garbini & Durand (1994), already discussed at the beginning of Section 3.3.2.
These scholars establish a direct link between the *-w-* of Hamitic PL forms like *tudmawin* in (755) and the long vowel / glide *-w-* found in the OA / CA putative NOM *ū(na)* of SOUND PL. Crucially the present work, while retaining the idea of such a link and of the *morphemic* value of *-w-*, departs from Garbini's etc... account in that the feature associated with the morph *-w-* is not PL, but Conj.
This shift of perspective has the advantage of explaining why *-w-* occurs as 'NOM' in the OA / CA SG too, a phenomenon left unexplained by the partisans of the Agglutination Hypothesis. Other advantages of the revised Agglutination Hypothesis are dealt with in Section 7.9.4 below.

7.9.3 The bound *-t* as a Numeral Classifier

Recent research by Owens (2006) has convincingly argued that the form of the OA noun is closer to that of the MCV noun than to that of the CA noun. On this view, the CA noun develops out of the pausal form typical of the OA / MCV noun.
This holds particularly valid from the noun marked by the bound *-t*, which Bauer (1912) and Greenberg (1990b) interpret as an INANIMATE Numeral Classifier because of its tendency to be combined with entities endowed with a low degree of animacy and numerals in OA, CA and MCV (see Section 1.1.3).
The bound *-t*, in fact, surfaces as *-ah / -at* in both the OA and MCV SG, and as *ā̄t* in their PL, while CA displays more complex forms: *-at-un* and *-ā̄t-un*, respectively (see Section 3.2.1). It should be added that following Retsoe (1995) the *-at* ending of the SG is likely to be older than its variant *-ah*, because the former but not the latter is documented in Akkadian and accordingly *-at* will be taken as a departure point for this discussion. We are thus left with the alternation SG *-at* vs. PL *ā̄t*, exemplified in (756):

(756) Old Arabic (based on Wright 1896, vol. II, p. 369)

SG: baqar-at 'cow'
PL: baqar-ā̄t 'cows'

On the level of form, this alternation is identical with that (N)-*in* vs. (N)-*īn* characterizing the MCV Numeral Classifier construction. This is much more so if one espouses Fleisch (1961, vol. I, p. 314, fn. 1)'s view that the *-a-* of *-at* is morphemic (thematic), not epenthetic[24].
The same holds for the level of meaning. On the one hand, there is a growing consensus among arabists that the OA SG *-at* has a singulative reading, ie it implies the numeral 'one' when suffixed to a noun, all its other values being one way or the other derived from the singulative one, feminine included (see Section 3.3.2 for details).
On the other hand, Early Arab Grammarians report that at least in OA the PL *-āt* is used to convey a paucal meaning (Ferrando, 2006).
Both phonological and semantic considerations therefore allow to extend the conclusions arrived at in the previous Chapter concerning *-in / -īn(a)* to *-at / -āt*. Thus, SG *-at* is better understood as underlying a covert *WĀḤID* and PL *-āt* as paralleling *-īn* in its phonological representation: it has as its input form *-āt'a*, turning into *-āt* via the same rules of *hamzah*-deletion, Low Vowel raising, *i*-deletion that turn *-īn(a)* into *-īn* (Section 7.9.1 and (743)):

(757) Noun + ā + t'a > Noun + ā + ta > Noun + ā + ti > Noun + ā + t
Noun + ī + n'a > Noun + ī + na > Noun + ī + ni > Noun + ī + n

In particular Fleisch's remark that the *a* intervening between the noun and *-t* is *not* epenthetic reinforces the parallelism between *-in* and *-at*, in the sense that precisely the same way the *i* preceding the *-n* of *-in* is a copula of Predicate Inversion, so is the *a* preceding the *-t* of *-at*. (756) is therefore rewritten as (758):

(758) Old Arabic (= 756)

a. baqar-a-t *WĀḤID*
cow-COP-CLF one

b. baqar-ā-t *FEW*
cow-COP-CLF few

'cow, cows (from 3 to 10)'

This is tantamount to saying that the phonological representation of the *a* preceding the *-t* of *-at* is *aw'i*, in line with the findings of Section 7.6 (see e.g. (727)). This results in a fine-grained structure of *-at / -āt*, schematized in (759):

(759) a(:)t = a + w + 'i + [t + *FEW*]
ABS + COPULA + *ya* + [MODIFIER]

The parallelism between the fine-grained structure of *-at / -āt* and that of *-in / -īn(a)* is illustrated in what follows:

(760) in = 0 + w + 'i + [n + *WĀḤID*]
at = a + w + 'i + [n + *WĀḤID*]

(761) īn = 0 + w + 'i + [n + ('a)]
āt = a + w + 'i + [n + ('a)]

It is worth observing in this connection that in Syriac, OA and Najdi Arabic there are documented some instances of noun marked by bound *-t* where the former and the latter are separated by a sequence displaying a glide, highly reminiscent of the phonological representation *aw'i* here assigned to the *a* usually intervening between them. As for Syriac and OA, they

[24] See Fleisch (1961, *Ibidem*) for phonological arguments in favor of his claim and references.

are the forms *lebb-awā-ṯā* 'hearts' and *san-awā-t* 'years' respectively, dealt with in Section 3.3.2 (see (151) and (152)): in such forms, a sequence *-awā-* similar to *aw'i* can be singled out.
Turning to Najdi Arabic, Ingham (1982a, p. 70) reports that the bound *-t* surfaces both as *-at* (in pause) and *-eih* (otherwise) in one of its dialects (Shammar):

(762) Najdi Arabic, Shammar dialect (Ingham, 1982a, p. 70)
NON-PAUSAL: sayyār-at / PAUSAL: sayyār-eih 'car'

In particular in this case, the *i* of the variant *-eih* seems to phonetically realize the item *'i* posited in the phonological representation *aw'i* on independent grounds.
The behavior of the bound *-t* in Shammar raises again the issue of its *t*-less phonetical realization in pause, also found in OA (and CA): Shammar *-eih* in this respect parallels OA (/ CA) *-ah*.
Recalling from Section 3.2.1 that the *-h* found in *-eih, -ah* is only a graphic notational device signalling the allomorphy of *-eih, -ah* with *-at*, there arises the question of why and how diachronically the older *-at* turns into the more recent *-eih, -ah* - or, more accurately, *-ei, -a*.
A partial answer to this question capitalizes on the strong solidarity between *-in* and *-at* discussed so far. It has been discussed at the end of Section 6.6.1 that the OA Animate Numeral Classifier *-n* of *-in* is dropped in CA and that only the *i(:)* preceding it is left in presence of a *covert* modifier (*WĀḤID*) - what is more traditionally referred to as the *-n* which 'closes' the word or 'demarcative'-*n*: e.g. *yamān-i-n* > *yamān-ī* (574).
It therefore comes as no surprise that *-at* in the same morpho-phonological context ('demarcative' *-at*) behaves the same way and is dropped, so that only its *-a* surfaces in the phonetic representation. (763) captures the parallelism under discussion:

(763) ANIMATE CLF: yamān-i-n > yamān-i(:)
INANIMATE CLF: baqar-a-t- > baqar-a

Whatever the ultimate phonological rationale behind it, the covert realization of the modifier seems to be the morpho-phonological context responsible for dropping of the Classifiers *-n, -t*, because in the archaic types of Arabic modification structure the *-t* followed by an *overt* modifier is preserved: cp. the Levantine Arabic RC type in (456) (beginning of Section 6.3) and the Moroccan non-pronominal CS (102) (beginning of Section 3.2.2.2). Recall that in this work both these types are regarded as archaic because they lack the so-called prenominal 'article'.
The strict parallelism between *-in* and *-at* on both phonological and semantic levels brings good evidence in favor of Bauer's / Greenberg's Classifier Hypothesis, but implements their idea in a different way, since the quantifier the Classifier *-t* is combined with is not an overt numeral (see (16) in Section 1.1.3), but a *covert* numeral *WĀḤID* or a grammaticalized quantifier *(')a* (evolving either into a PL or a DU).
This account seems preferable because it does not run into the interpretive problems encountered in Bauer's and Greenberg's, e.g. the anomalous word order of constituents (contrast (15) with (17) in Section 1.1.3). Adopting this approach the complex syntax of numerals responsible for the just mentioned word order, whatever its explanation, is an issue orthogonal to the Classifier Hypothesis and does not affect it.
Another advantage of interpreting the OA / MCV alternation *-at / āt* through the lens of the version of Classifier Hypothesis argued for here is its ability to derive the behavior of *t* in the

diachronically later CA. Following the received view (see Section 3.3.2), what distinguishes CA from OA is that :

- on the level of form *-un* tends to replace *-at* in its function of SG / singulative and PL / paucal marker
- on the level of meaning *-a(:)t* no longer functions as SG / singulative and PL / paucal marker turning into a F-marker

The former point is exemplified by two phenomena. In the SG / singulative, the CA nouns ending in *-at(u)*, e.g. *'usam-at-(u)* of Table 7.1, occupy a peripheral sector of grammar, the Diptotic Declension, where this suffix seems not to be productive. As for the PL / paucal meaning, Arab Grammarians maintain that in the form *-āt-un* the suffix *-un* is not original, being rather the result of a process of analogy (MUQĀBALAH), by virtue of which the ending *-āt* is attracted into the ending *-ūna*. Consequently *-n* replaces *-āt* as PL ending (see Ayoub 1991 for details).
This scenario can be schematized in the following[25]:

(764) OA: 'usām-at > CA: 'asad-u-n

Notice that both the nouns in (764) mean 'lion' and are M.
The latter point has been already discussed at length in Section 3.3.2 and then the following example will suffice to illustrate it:

(765) OA: 'usām-at > CA: 'asad-<u>at</u>-u-n
lion-CLF lion-<u>F</u>-COP-CLF
'a lion', 'a lionness'

The *-at* found on the CA noun *'asad-<u>at</u>-u-n* in (765) clearly differs from the *-at* of OA *'usām-at* because it turns the noun *'asad-u-n* of (764) from M to F, hence it is a F-marker.
Plainly, evolution of *-t* from a Classifier into a F-marker correlates with its inability to function as a Classifier: when *-n* takes over its Classifier function, *-t* specializes into a new function (F). What has been left explained so far is why *-n* replaces *-t*.
This scenario is totally expected under the assumption, motivated in the previous Chapter, that *-n* is an Animate Numeral Classifier evolving into a General Classifier: Greenberg (1990b, p. 190) in fact observes that crosslinguistically the General Classifer replaces all the other classifiers in their capacity to combine with a quantifier to express number ('may be used as an alternative to almost any other classifier').
This process of diachronic expansion of *-n* has been already documented for Arabic in the previous Chapter throughout the study of the Possessive Classifier. Remarkably, the same process is *a fortiori* plausible in the case of the Inanimate Numeral Classifier, since Greenberg (1990b, *Ibidem*) has found out that 'often the spread of such classifiers [= General Classifiers] is confined to inanimates'.
A case in point is the piece of data (615) dealt with at the end of Section 6.6.2: in this example the Mandarin Chinese General Classifier *ge* replaces the Inanimate Numeral Classifier *liàng*

[25] In the following diagram the form *'usām-at-(u)* is regarded as older than the form *'asad-u-n* because in CA the former belongs to the Diptotic Declension, the latter to the Triptotic Declension and under the Language Contact Hypothesis, the Diptotic declension is older than the Triptotic one: see fn. 21.

combined with the numeral *yí* 'one', and the same holds for numerals other than *yí*.
In a likewise fashion, the Arabic General Classifier *-n* replaces the Inanimate Numeral Classifier *-t* in its ability to get combined with the quantifiers *WĀḤID* and *FEW* (what in more traditional terms would be referred to as the fact that *-n* takes over the singulative / paucal function of *-t*). This parallelism is depicted in (766)

(766) Chinese: (yí) liàng (jiǎotāchē) > (yí) ge (jiǎotāchē) (= 615)
OA / CA: ('usām) at (WĀḤID) > ('asad) (u) n (WĀḤID) (= 764)

To conclude:

(767) The bound *-t* is an Inanimate Numeral Classifier replaced by the General Classifier *-n*

7.9.4 Numeral Classifiers: an overall view

In Section 3.3.2 Arab Grammarians' tripartite division of the nominal paradigm:

(768) Old / Classical Arabic (= 160)

baqar	/	baqar-at	/	baqar-āt
cow.COLL		cow-SG		cow-PL

(769) Old / Classical Arabic (= 161)

šajar	/	šajar-at	/	'ašjār-
tree.COLL		tree-SG		tree.PL

has been analyzed through the lens of the Greenbergian Classifier Hypothesis, and this vantage point has made it possible to set a parallelism between the tripartite division at issue and conceptualization of number in Numeral Classifier Languages, where this latter is built upon the three-fold opposition COLL - PL - SG. More precisely, the COLL is a noun unmarked for number, in the sense that it works both as SG and PL (see Section 6.6).
This *crosslinguistic* parallelism, however, is not the only one unveiled by Greenberg's original formulation of the Classifier Hypothesis.
Indeed, implicit in this latter is another parallelism, *language-internal*: that between SOUND PL and BROKEN PL, since both of them are organized into the tripartite division under discussion, as exemplified in (768 - 769). This parallelism is cued by a morpho-phonological piece of data first observed by the proponents of the Agglutination Hypothesis: the fact that both the SOUND PL ending with *-n*:

(770) (cp. Section 3.2.1)

- (???)	/	mu'min	/	mu'minūna
believer.COLL		believer		believer.PL

and some instances of BROKEN PL share one and the same morpheme *w*.
But while the BROKEN PL seems to arise, among others, from a combination of this morpheme with a COLL N, no such state of affairs is traditionally recognized for the SOUND PL ending with *-n* in literature, where the substantive item to which *w* is suffixed is simply designated as stem or the alike (cp. (770)). Given this, interpreting the stem of the SOUND PL ending with *-n* as COLL in order to realign it with its BROKEN PL counterpart is merely conjectural, as illustrated in (771), where □ indicates the conjectural cell of the paradigm:

(771) (= 164)

COLL > BROKEN PL: ⇳, -*w*-$_i$
[COLL] > SOUND PL: ⇳, -*w*-$_i$

The paradigmatic gap in (770) clearly threatens viability of the Greenbergian Classifier Hypothesis.
True, returning to (768 - 769) seems to solve this problem, because in this paradigm, which leaves -*n* out of the picture, there exists a clear connection between COLL and PL regardless of the BROKEN or SOUND status of this latter.
Nevertheless, taking apart COLL is not enough to save Greenberg's Classifier Hypothesis from criticism: if -*t* is a Numeral Classifier, as he claims, why does not the quantifier typically combined with it occur?
Sticking to the Greenbergian original formulation of this hypothesis, things seem to get even worse if (769) is taken into consideration, since the SG ending with -*t* has a *t*-less BROKEN PL: in this case, the Classifier too is lacking.
These problems are summarized in the question (V) at the beginning of the previous Chapter and schematized in the following:

(772) Old / Classical Arabic (= 160)

baqar	/	baqar-at [Q?]	/	baqar-āt [Q?]
cow.COLL		cow-CLF		cow-CLF

(773) Old / Classical Arabic (= 161)

šajar	/	šajar-at [Q?]	/	'ašjār- [CLF?] [Q?]
tree.COLL		tree-SG		tree.PL

(774) (= 770)

[???]	/	mu'min [CLF?] [Q?]	/	mu'minūna [CLF?] [Q?]
believer.COLL		believer		believer.PL

The findings of the previous and present Chapters fill up the empty cells of the just mentioned paradigm, in particular:

- the quantifier 'one' is generally omitted in the Numeral Classifier languages
- in OA the quantifier *(')a* occurs
- in Najdi Arabic, the Animate Numeral Classifier -*in* is not dropped in the SG and PL when conveying a singulative / paucal meaning respectively
- the Animate Numeral Classifier, *qua* evolving into a General Classifier -*Vn*, replaces the Inanimate Numeral Classifier -*at*

Generalizing the last point from -*in* to -*Vn*, the following picture obtains (abstracting away from the copulae *u,a,i* for the sake of simplicity):

(775) Old / Classical Arabic (= 160)

baqar / baqar-at WĀḤID / baqar-āt ('a)
cow.COLL cow-CLF Q cow-CLF Q

(776) Old / Classical Arabic (= 161)

šajar / šajar-at WĀḤID / ('a) šjār- un
tree.COLL tree-CLF Q Q tree CLF

Notice that the BROKEN PL of (776) 'deviates' from the general pattern of word order Classifier > Quantifier: this naturally follows from the General Classifier status of *-Vn*,whose presence may trigger fronting of the Quantifier in prenominal position (Greenberg, 1990a, p. 229)[26].
The noun having the SOUND PL ending with *-n* is realigned with (775 - 776) if one considers the SG noun *taba'* 'companion, follower', which in OA is characterized by two interesting properties. On one side it also works as PL (ie it is a COLL: see above) and on the other side it is interchangeable with the variant *tābi'*, whose SOUND PL is *tābi'ūna*, as reported by Lane (1863, vol. I, p. 295):

(777) Old Arabic (Lane, 1863, vol. I, p. 295)

taba' / taba'- un WĀḤID / tābi' ūn (')a
companion.COLL companion CLF Q companion CLF Q

Arab Grammarians' tripartite division of the nominal paradigm, if combined with the findings of this dissertation, yields the revised scenario in (775 - 776), which turns out to be no more than a *synoptic view of a Numeral Classifier Construction*.
In passing, this revised paradigm also substantiates the conjectural scenario posited in (771) and answers the question (V).

7.10 Summary

This Chapter has revived Arab Grammarians' view that the vowels *a, i* found in the twofold phonological environment consisting of the alternation C_VC(CV) / C_V̄CV have the *phonological representation awi, uwi*.
Furthermore, it has combined this idea with the Agglutination Hypothesis on the Arabic 'Case' à la Tosco, Garbini, Petracek, Tropper etc... according to which these vowels, when acting as 'case-endings' have:

- an abstract representation *decomposable into distinct morphs*
- a semantic value other than that traditionally ascribed to Indo-european Case.

As for the phonological representation of Case, this Chapter has clarified that the 'missing link' between Arab Grammarians' phonological analysis and the Agglutination Hypothesis on the Arabic 'Case' is the (pretheoretical) fact that 'case-endings' *a, i* occur in the same twofold phonological environment in which *awi, uwi* turn into *a, i*.
Turning to the semantic value of Case, the syntactic diagnostics worked out by Kern (2010)

[26] 'We may conjecture that in later stages with the spread of the general classifier to the point where the language is close to or in some cases reaches a situation in which there is only a single 'classifier', the construction come sto be treated more and more as simply a quantifier and hence moves to the more normal order Q-N...'

support Nyberg (1920)'s claim that (*awi, uwi* phonetically realized as) *a, i* are copulae, showing that more accurately they are copulae found in a context of Predicate Inversion.
The study of the syntax of Hamitic ABS, of the behavior of *hamzah* in OA, as well as of the alternation of the Semitic demonstrative base *ya- / 'i* brings good evidence in favor of identifying the morphemic material *awi, uwi* underlying *a, i* as an agglutinative complex, made of:

- the morph *a*, a relic of ABS marking the Predicate / Modifier
- the morph *w*, the Conj copula of Canonical Predication dealt with in the previous Chapter
- the morph *'i* acting as copula of Predicate Inversion

The same arguments corroborate the claim that the phonological representation of the putative case-ending *u* is better understood as *aw*, which minimally yet crucially differ from that of *a, i* in that the latter ones include the copula of Predicate Inversion *'i*, the former does not:

(778) *u* = (a) + w ... - 'i ⇔ Predicate Inversion: NO
a, i = (a) + w ... + 'i ⇔ Predicate Inversion: YES

As such, the Agglutination Hypothesis unveils that Case has an internal and fine-grained structure, and that the putative case-endings *u, a, i* are organized into a binary opposition *u* vs. *a, i* based on the lack vs. presence of Predicate Inversion. But, interestingly, this opposition is not the only one in which *u, a, i* are involved. Arab Grammarians in fact argue that *u* is opposed to *a,i* in CA in terms of presence vs. lack of ability to act as Subject. In addition to that, *u* is opposed to *a,i* in terms of dialectal variation, the former being found in the Hijazi branch, the latter in the Bedouin one and remarkably in these dialectal groups the opposition at issue is typically found in a context of modification structure where the modifier is an 'adjectival' RC modifier.
If, on the other hand, the RC modifier is a 'pure' (non-adjectival) RC, lack vs. presence of Predicate Inversion is signalled not by the opposition *u (aw)* vs. *a,i (awi, uwi)* but by the opposition lack vs. presence of the copula of Predicate Inversion *ya-* combined with the RC marker *lli*.
In so doing, both the Hijazi and Bedouin branches seem to preserve two distinct RC types, the Amorite and the Pre-semitic one.
Indeed, if one calls again into the picture the Language Contact Hypothesis, one finds that in both these branches the adjectival RC type is the Pre-semitic RC type because of the Pre-semitic nature of the putative case-endings, whereas in its non-adjectival counterpart the very occurrence of the RC marker *lli* points to its Amorite origin.
Concentrating on the multi-level parallelism characterizing the opposition *u* vs *a, i*, it has the advantage of explaining the nature, the birth and the development of Arabic case-system.
This latter in origin is a dialectal opposition between the Hijazi and the Bedouin branch (surviving in OA and still documented in the modern peninsular MCV) in terms of lack vs. presence of Predicate Inversion. When these two branches fuse into CA, their endings compete for the role of Subject, and the Hijazi *u* prevails over the Bedouin *a, i* because the former is associated with the NP informational profile of Canonical Predication, felt as a more suitable candidate for the role of Subject than the NP informational profile of Predicate Inversion associated with the latter. Hence the Bedouin *a, i* polarize into non-Subject markers.
In the Hijazi and Bedouin branches, the binary opposition *u* vs *a,i* correlates with another

opposition, that between the presence (Hijazi branch) vs. the lack (Bedouin branch) of a desemantized demonstrative base *d̲-*.
On the whole, investigation of the fine-grained structure of Arabic Case has a positive side-effect, namely providing a more precise description of the modification structure of the Arabic NP:

(779)
- Hijazi branch: Pred. Inversion: NO + Desemantized (otiose) *d̲*: NO
  - Amorite parastrate: Adjectival RC: YES
  - Pre-semitic parastrate: Adjectival RC: NO
- Bedouin branch: Pred. Inversion: YES + Desemantized (otiose) *d̲*: YES
  - Amorite parastrate: Adjectival RC: YES
  - Pre-semitic parastrate: Adjectival RC: NO

Another implication of this analysis is a refinement of the Classifier Hypothesis. Nouns marked for DU are assigned the same syntactic structure (Numeral Classifier Construction) of those marked for SOUND PL, the former differing from the latter only in the covert vs. overt realization of the relic ABS marking the modifier (the complex Classifier plus quantifier) in the phonological representation (*awi* vs. *uwi*), so that the DU meaning is the result of a later polarization. So are nouns marked for OBL relative to those marked for ACC (albeit the process of polarization undergone by them needs more inquiry).
Finally, combining the copula analysis of Case with the Classifier Hypothesis provides a better understanding of BROKEN PL and of bound *-t*. On this view, BROKEN PL might plausibly be no more than a SOUND PL, whose copula *V:(n)* has undergone shortening subsequent to either fronting or phonological deletion of the quantifier *(') a*, whereas the thematic *a* preceding the bound *-t* is rethought of as the same copula *a* traditionally referred to as ACC.

Chapter 8

Transformations

8.1 Phrase structure: an overview

The previous Chapter has set out a twofold typology of the Arabic (NONRE)RC (see Section 7.8 and Table 7.8 for details):

- Hijazi branch: (NONRE)RC: Canonical Predication
- Bedouin branch: (NONRE)RC: Reinstated Word Order

Both the Hijazi and Bedouin (NONRE)RC type in its turn can be split into two subtypes, an adjectival and a non-adjectival one, in complementary distribution and specializing the older typology Amorite - Pre-semitic (NONRE)RC (Language Contact Hypothesis):

- Pre-semitic parastrate: (NONRE)RC: Adjectival
- Amorite parastrate: (NONRE)RC: Non-adjectival

It seems convenient to focus on the Bedouin (NONRE)RC type because the Reinstated Word Order, *regardless of whether it manifests itself into an adjectival (Pre-semitic) or a non-adjectival (Amorite) RC*, in its derivational history passes through both the modification structures characterized by Canonical Predication and Inverted Predication (see Ch. 5). For example, in Najdi Arabic the copula of Predicate Inversion (CPI) *ya-* found in the following instance of Inverted Predication:

(780) Najdi Arabic (= 692 , 693)

gilt-ih-li-hum	⟶	gilt-li-hum-iyyā-h
I.said-it-to-them		I.said-to-them-CPI-it

'I said it to them'

also occurs in the non-adjectival Amorite (NONRE)RC (NONADJAMOR henceforth):

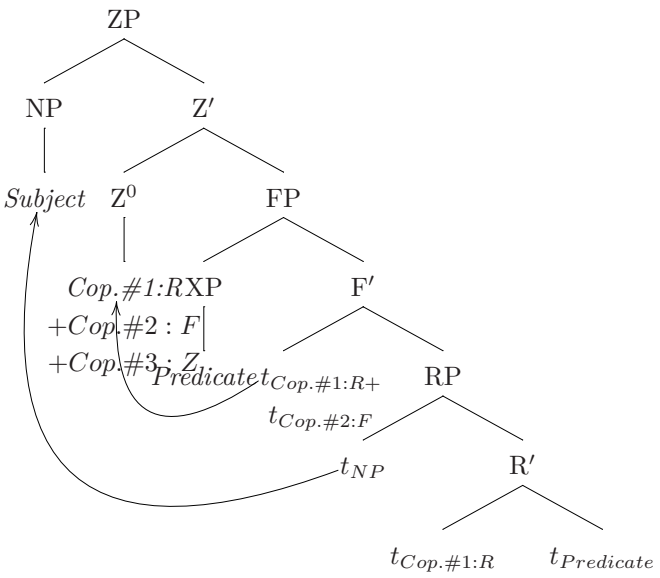
(781) Najdi Arabic, Murra dialect (= 698)

[yā jazl	al-'ata]	ya	lli	[b-labs	il-kāmtīn	adrāi]^{RC}
O abundant	in.the-giving	CPI	C	of-hidding	the-concealers	aware

'O Generous, who are aware of the things hidden by the concealers'

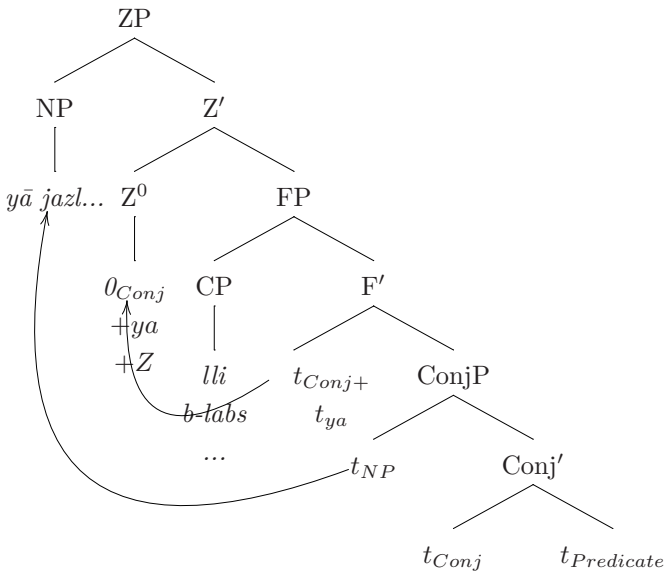
It ensues that the Bedouin Reinstated Word Order includes all the functional projections involved in the Arabic NP and then can be taken to be illustrative of them.
Furthermore, it represents the (Extended) Adjective and the non-pronominal CS too, because it has been evidenced throughout Ch. 6 and Ch. 7 that the two latter are no more than a subtype of the former (in particular, the non-pronominal CS is a non-restrictive RC because it is a Possessive Classifier Construction).
Last but not least, the Arabic Reinstated Word Order (781) also describes what is traditionally referred to as SG and (SOUND / BROKEN) PL displaying *-n* and bound *-t*, because these are Numeral Classifier Constructions and therefore (underlying) non-restrictive RCs.
The minimalist account developed in Ch. 5 assigns the Reinstated Word Order the syntactic structure in (360), replicated here as (782), which also takes into account the refinements of the *syntactic* predication theory argued for in Section 5.3. They basically are lack of distinction between Relators and Linkers, and the consequent generalization of the notion of copula, applying not only to den Dikken (2006)'s R (Canonical Predication), but also to his F (Inverted Predication) as well as to his Z (Reinstated Word Order).
On these grounds the Arabic manifestation of (782) can be represented as (783), which is based on the NONADJAMOR found in the Bedouin branch (781), traditionally designated as RC and discussed in Section 7.8[1].
A desirable side-effect of identifying NONADJAMOR with the Reinstated Word Order is that it allows to unveil at least the exact nature of R, namely Conj (see next Section on F, Z).
As already discussed in Section 7.5, (783) parallels Thai (359):

(782) Predicate Inversion *cum* Reinstated Word Order



[1]Recall from the end of this Section that this definition is meant to capture the functional core of the Arabic RC, abstracting away for later developments like 'injection' of Pre-semitic elements into the non-adjectival Amorite RC type and the other way around.

(783) Predicate Inversion *cum* Reinstated Word Order



In particular, by virtue of Baker (1988)'s Mirror Principle (see Section 4.5.3) the complex Head hosted in Z', in the left periphery of (783), namely:

(784) ...[$_{Z'}$ Conj: 0 + F: *ya-* + Z: 0...

is the result of recursive head movement through the following functional layer:

(785) ...[$_{Z'}$ Z: 0... [$_{F'}$ F: *ya-*... [$_{Conj'}$ Conj: 0...

In (783) all these Heads are *covertly* realized but F, surfacing as *ya-*.
On the other hand, the 'adjectival Pre-semitic RC type' found in the Bedouin branch (what is traditionally designated as Adjective, ADJPRESEM henceforth) assigns all of these Heads (nonterminals), except Z, an *overt* phonological correlate (terminals), at least in their phonological representation.
This is due to the fact that the putative case-ending *a,i* in Najdi Arabic works not only as CPI (F):

(786) Najdi Arabic (= 703)

mbaccir a / *u -l-wasim
early CPI / CONJ -C-winter.rain

'early winter rain'

but also introduces ADJPRESEM (Reinstated Word Order):

(787) Najdi Arabic, Dhafīr dialect (= 681)

(al-)bēt a-l-ṭuwīl
(*al-*)house CPI-C-tall

'the tall house'

But a combination of Arab Grammarians' analysis of *a, i* as having the phonological representation *aw'i, u'i* with the Agglutination Hypothesis à la Garbini, Petracek, Tropper etc..., according to which the phonological representation of *i*, when acting as putative case-ending, is *decomposable into two distinct morphs w, i* (see (148) in Section 7.10) has led to associate to the phonetic representation *a,i* of the putative case-endings ACC, OBL the following pair phonological representation - syntactic structure:

(788) (= 721)

Noun + (a) + w + 'i + Modifier
Noun + ABS + Conj + *ya-* + RC

In particular, the morphemic value assigned in (788) to *aw'i, u'i* is a consequence of extending to Arabic Kern (2010)'s diagnostics to detect copulae of Predicate Inversion as well as of the syntactic behavior of (the relics) of ABS in Hamitic and Arabic. The phonetic representation *u* of the putative case-endings NOM, *qua* copula of Canonical Predication, is assigned the same sound - meaning pair, minus the CPI *'i* (see (727)).
The minimalist representation of the CPI having the phonetic representation *a* and the phonological representation *aw'i*, if occurring in the Reinstated Word Order in (787) is a complex Head hosted in Z':

(789) ...[$_{Z'}$ Abs: *a* + Conj: *w* + F: *'i* + Z: 0...

that by virtue of the Mirror Principle arises subsequent to Head Movement through the following functional layer:

(790) ...[$_{Z'}$ Z: 0... [$_{F'}$ F: *'i*... [$_{Conj'}$ Conj: *w*... [$_{Abs'}$ Abs: *a*...

Remarkably, (790) not only parallels the functiona layer in (785) insofar as Z, F, Conj are concerned, thus confirming it, but also provides an additional piece of information about the lowest portion of syntactic structure of the Arabic NP, where Conj selects the Head Abs, in its turn selecting the substantive Head N. To put it differently, NP projects a functional layer beginning with AbsP and ending up with DP (on which, see Section 6.4.4):

(791) DP > ZP > FP ... ConjP > AbsP > NP

It seems desirable to define more precisely the nature of the just discussed functional projections F, Z. This issue is tackled in the next Section.

8.2 The universal hierarchy of functional projections and the Arabic NP

Upon closer scrutiny, the functional layer worked out for Arabic throughout Ch. 6 and Ch. 7 relying upon language-internal cues and some syntactic diagnostics closely matches that proposed for this language by Cinque (2003) and further refined by Ihsane (2003) within the antisymmetric framework[2].

[2]On independent grounds, Shlonsky (2004) proposes an antisymmetric account of the Arabic NP converging in many respects with Cinque (2003)'s. Shlonsky however carries out a different analysis for GenP, which will be not be adopted here.

The advantage of Cinque (2003)'s and Ihsane (2003)'s antisymmetric account is their cartographic perspective: the Arabic functional layer is not language-specific, being rather the manifestation of an universal hierarchy of functional projections. It should be added in this connection that at least Ihsane (2003)'s account is based on Moroccan and Levantine Arabic, which avoids the potential criticism that the data used for the analysis are not reliable, as is the case for MSA (see fn. 4 in Ch. 1).
The cartographic functional layer is characterized by WP > GenP in its right periphery, as per Cinque (2003) (cp. also Ouhalla 2004) and by DP > FocP > DefP in its left periphery, following Ihsane (2003):

(792) DP > FocP > DefP ... WP > GenP > NP

Starting with the right periphery, GenP of (792) could be easily identified with AbsP of (791), since both of them are the *locus* of the 'oblique constituent' / Dependent.
Furthermore, WP corresponds to ConjP because the key property of both these projections is *setting the informational profile Old - New*. This is evident for ConjP: its Spec is the antecedent of a non-restrictive RC, hence already known to the listener, while its Compl is the non-restrictive RC itself, adding new information to its antecedent (see fn. 10 in Ch. 5).
As for WP, this claim is based on the antisymmetric analysis of English sentences exhibiting a final FOCUS (piece of new information) developed by Kayne (1998, pp. 149-150) and den Dikken (2006, p. 28), according to which a sentence final focus (e.g. *only biscuits* in *Imogen eats only biscuits*) fulfills this position only subsequent to raising of its PRESUPPOSITION (old piece of information) *Imogen eats* to Spec, WP.
Without going into technicalities, the same account has it that *only* head-moves to the covert Head W while the focused constituent (*biscuits*) sits in Compl, WP:

(793) ...[$_{\text{WP}}$ [$_{\text{XP}}$ Imogen eats]$_j$ [$_{\text{W}'}$ only$_i$ + W [$_{\text{FocP}}$ biscuits [$_{\text{Foc}}$ t$_i$ t$_j$...

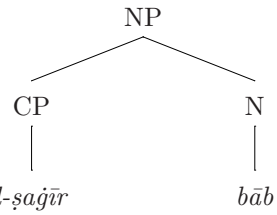
This means that WP has the same informational profile Old - New (presupposition - focus) of ConjP and on these grounds the former is identified with the latter.
Turning to the left periphery, DefP is the secondary predication counterpart of Rizzi (1997)'s FinP occurring in primary predication and as such denotes the time and the place a given entity is located in. Since in Arabic the functional item deployed to signal Predicate Inversion (F) is a former demonstrative base *ya-* (Amorite) / *'i* (Pre-semitic), as exemplified in (699) and the function of a demonstrative is by definition locating a given entity in time and space, F heading FP is taken here to be a (grammaticalized) Def heading DefP. Finally, ZP is FocP, because the NP sitting in its Spec carries new information[3]. (792) is rewritten as (791) accordingly.
The just sketched out minimalist analysis is refined in what follows in dynamic antisymmetric terms, especially insofar as displacement phenomena are concerned.
The present work, in fact, departs from the antisymmetric / cartographic approach in that it dispenses with the feature-checking account in explaining movement (see Section 4.5.2), and then all the displacement phenomena posited in Cinque (2003)'s and Ihsane (2003)'s to bridge the gap between the universal word order and their surface manifestation are to be reformulated along the lines of Dynamic Antisymmetry. This is done in the remainder of this Chapter.

[3]Jamal Ouhalla (pc) interprets ZP in the same way.

8.3 The POS DP - N: from Spec, NP to Spec, AbsP

Under standard antisymmetric assumptions, the noun and its non-restrictive RC (NONRERC henceforth) are merged into the syntactic structure fulfilling N and Spec, NP respectively, as already discussed in Section 5.3.1 (recall in particular Ouhalla 2004's unified account of the Arabic RC and of the non-pronominal CS).

(794)



(794) is exemplified by OA *bāb-u l-ṣaġīr* 'the little gate' (49), an ADJPRESEM type , but it holds for the NONADJAMOR type too because they have the same structural makeup (see Ch. 6).

In Section 5.3.1 it has been also remarked that lack of Compl in (794) yields a syntactic configuration that Dynamic Antisymmetry can regard as symmetric in some intuitive sense, as a simple glance to the geometry of (794) shows (this insight is captured in formal terms immediately below).

Developing the dynamic antisymmetric line of reasoning, one may also point out that the NONRERC thus merged appears to show clitic-like properties.

If in fact the noun the NONRERC modifies is marked for the *tā' marbūṭah*, this latter in Levantine Arabic surfaces as *-t*, but this the form that the *tā' marbūṭah* takes when its noun is combined with a *pronominal clitic* in the CS (see Section 6.6.2 on pronominal clitics). Compare (583) with (456) replicated here for the sake of commodity as (795):

(795) Mor. Ar. / Lev. Ar. (= 583, 456)

a. mədras-t-i
school-*at*-me

b. baTT-it illi...
the-duck-*at* that...

'my school, the duck that...'

Notice that the *tā' marbūṭah* is phonologically realized as *-t* also when the modifier of the CS is non-pronominal. This is not surprising under the approach adopted here, given that the non-pronominal modifier of the CS, *qua* underlying a Possessive Classifier, is actually a NONRERC (see above).

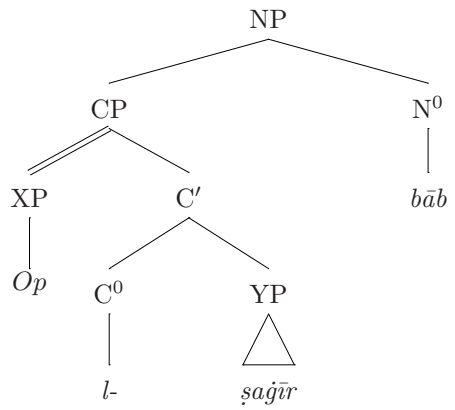
Dynamic Antisymmetry can capitalize on the parallelism in (795) to explain the difference between the base generation position of the Arabic NONRERC (Spec, NP) and its position after spell-out assuming the following working hypothesis:

(796) The difference between the base-generation position of the Arabic NONRERC and its postnominal position after spell-out is a function of its clitic-like status

It should be remarked in this connection that the Arabic NONRERC shows a paradoxical behavior, because on one side it morphosyntactically behaves as a clitic (795), on the other side, from a morpho-phonological standpoint, it is too a long phonetic string to be described as

such.
The solution of the conundrum lies in the c-command relationship dealt with in Section 4.5.3. Consider a more detailed representation of (794), namely (797)

(797)



In (797), the topmost constituent of the NONRERC, namely the nonterminal DP sitting in Spec, CP, whose terminal is Op, c-commands N and N c-commands it (*symmetric* c-command). This happens because the first *category* dominating DP is not CP, but NP, which dominates N as well.
More accurately, only one *segment* of CP dominates DP (CP being made of two segments: CP and C') and this is not enough for c-command: an entire *category* is required to dominate DP in order to set such a relationship (see Moro 2000, p. 55 and Section 4.5.3). The segment status of the portion of syntactic structure dominating DP is indicated by the double line in (797).
It should be reminded at this point that in the NONADJAMOR type the DP hosting *d̲ī* is doubtlessly a clitic-like element and that the DP belonging to the ADJPRESEM, hosting Op, indirectly has the same status because of the analogical pressure exerted by the Amorite parastrate (see Section 6.4.2). In other words, in both the cases *the DP hosting the relative pronoun is the real clitic-like element of the NONRERC*, which as such behaves as a Head and forms a Head - Head construction with N, thus creating the (blind) POS *d̲ī - N / Op - N*.
The two poles DP (*d̲ī*, Op) and N (*bāb*) cannot be neutralized representationally because they are not adjacent (e.g. *l-ṣaġīr* intervenes between Op and *bāb*).
Alternatively, the POS at stake is neutralized derivationally, not representationally, and then movement of DP to the first available landing site, namely Spec, AbsP is expected to apply.
Although both movement of DP and N in principle are possible, the Economy Principle favors the former over the latter (see Section 5.3.2 and (801) in the next Section).
Raising of DP alone, however, is barred for semantic reasons. Indeed DP is an Op coreferential with the antecedent N and as such must satisfy the PRINCIPLE A, which (slightly simplifying) imposes that any coreferential expression occur within a domain including a given Head governing it (the governor) and another Head agreeing with it (the ACCESSIBLE SUBJECT). A common formulation of the Principle A is the following (from Haegeman 1991, pp. 223):

Minimalist definition 8.1 (Principle A)**.**

An anaphor X must be bound in the minimal domain containing X, X's governor and an accessible subject/ SUBJECT.

If the domain at issue is designated as GOVERNING CATEGORY, the Principle A can be also restated as follows (Haegeman, 1991, pp. 223) :

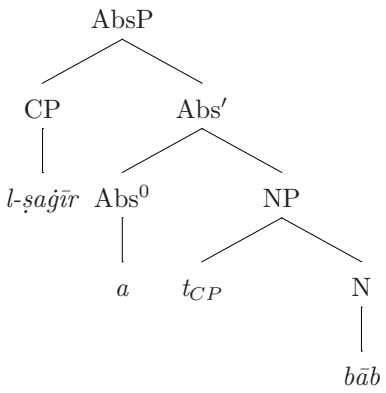
Minimalist definition 8.2 (Principle A)**.**
An anaphor must be bound in its governing category.

In (797), prior to movement, the DP Op is bound in its governing category since the internal structure of the RC modifier (CP) that contains it also contains, under standard (pre-)minimalist assumption (see (236) at the end of Section 4.5.3), the functional projection IP:

(798) $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- $[_{\text{IP}}$ t$_{Op}$ $[_{\text{I}'}$ ṣaġīr

Without going into details, this latter is projected by the Head I responsible for assigning the RC its tense and / or *agreement* etc... so that I works as the accessible SUBJECT of Op. Moreover, I is generally conceived as the Head governing the constituent sitting in Spec, Ip, which is precisely the position occupied by Op before moving to Spec (see (798)).
In consequences of this it is the entire (NONRE)RC (CP) that raises to the first available landing site Spec, AbsP, because CP includes the governing category IP required by Op (/ *d̲ī*) to comply with the Principle A[4]. This is schematized in (799):

(799)



This is precisely what Cinque (2003) and Ouhalla (2004) take to be the first derivational step of the Arabic NP, which is therefore recast in dynamic antisymmetric terms.

8.4 The POS Abs - NP: Remnant movement of N to Spec, ConjP

After movement of the NONRERC to Spec, AbsP, the nonterminals NP and N are left in situ along with their terminal, say *bāb* in (799).
This results in a scenario such that Abs, whose nonterminal is *a*, c-commands NP via the node AbsP and NP c-commands Abs via the same node, to the effect that the c-command is *symmetric*. Moreover, NP is Dependent-less, since it is base-generated without (the Dependent sitting in its) Compl and it has 'lost' the Dependent sitting in its Spec subsequent to movement

[4]Alternatively, one may invoke the Left Branch Condition (see Appendix 5.3.6) rather than the Principle A to block movement of the DP Op / *d̲ī* in (799). Section 8.6.2 shows that the Principle A is preferable over the Left Branch Condition because the former derives some displacement phenomena of the Arabic NP that the latter does not: see fn. 9. .

of this latter to Spec, AbsP.
If it was shown that NP is monosyllabic, it would have the two properties typically ascribed to a clitic, and this behavior would create a Head - Head construction creating the POS Abs - NP, e.g. *a - bāb*.
It is worth observing in this connection that the monosyllabic nature of the Arabic noun is evidenced by the semantic arguments of the so-called BILITERAL THEORY.
The core of this latter is that several Arabic words have a common semantic nucleus encoded in *two* consonants: e.g. the CA verbs *falà, falaqa, falaḍa, falaḥa, falaja* share not only the *meaning* of 'to make an opening by cutting' but also the consonantal *form f, l* (call it the BICONSONANTAL NUCLEUS). This is an example reported by Arab Grammarians themselves, which were the first to find out this property of the Arabic lexicon (see Fleisch 1961, vol. I, p. 256 for details).
The most recent and exhaustive formulation of the Biliteral Theory has been given in Bohas (1997), where the biconsonantal nucleus is thought of as a sequence *CVC* and *as the real stem ('radical') of all Arabic words (substantive items) prior to affixation processes (functional items)*, at least insofar as their phonological representation is concerned[5].
In this light, any Arabic substantive item is monosyllabic because other consonants / vowels making it more than monosyllabic are functional items and, remarkably, this state of affairs is very likely to be found *in synchrony in the speaker's mind*.
Bohas (1997) makes this point very clear, by stating that the biconsonantal nucleus:

> "est tout à fait accessible et ne constitue pas un «stade antérieur hypothétique» mais bien le facteur constant ou «panchronique» d'organisation phonético-sémantique de ce lexique".
> (Bohas, 1997, p. 166)

Sometimes, the monosyllabic stem CVC surfaces as well in the phonetic representation of the Arabic nouns: *'ab-, yad-, dam-* 'father, hand, blood' (cp. Garbini & Durand 1994, p. 92).
Empirical support in favor of the claim that Arabic substantive items (= stems) are monosyllabic is lent by an alternative argument, of *phonological* nature, namely that even a *bisyllabic* and triconsonantal phonetic representation alternates with a *monosyllabic* counterpart in one and the same stem.
Garbini & Durand (1994, p. 89) in fact observes that in what Arab Grammarians analyze as the verbal paradigm, there is documented an alternation stem CV-CV̄C (bisyllabic) vs. stem CCVC (monosyllabic): e.g. *du-xūl - (ya)dxul* 'entering, he enters'.
To this, one may add that the OA / CA stem CV̄-CVC (bysillabic) in the MCV surfaces as a stem CVCC (monoyllabic): contrast OA / CA *kātib(ah)* 'writing (F)',*(mu)sāmaḥ(ah)* 'forgiveness' , *(mu)ẓāhar(ah)* 'protest march' with Egyptian Arabic *katb(ah)*, *(mo)samḥ(ah)*, *(mo)ẓahr(ah)* 'id.'. (800) schematizes the alternation at issue:

[5]See in particular Bohas (1997, p. 1): 'nous entendons par ce terme une base dérivationnelle comportant des consonnes et au moins une voyelle, la valeur de cette dernière restant à préciser '.
In so doing, Bohas follows Chekayri's positions, according to which the phonological representation of, say, *qawala* before Glide Creation is *qal*: see Bohas (1997, p. 97) and, more recently, Chekayri (2005) for further developments of this idea. Cp. also fn. 42 in Section 2.3.3 on the notion of stem etc... and Section 6.3.2 on the phonological representation à la Chekayri as well as Bohas (1997, Ch. IV) for arguments against a vowelless and / or a triconsonantal phonological representation of Arabic substantive items.

(800) OA / CA: kā + tib ⇔ MCV: katb
Bisyllabic: $C\bar{V}$ + CVC ⇔ Monosyllabic: CVCC

The discussion so far evidences that the Arabic N is actually monosyllabic[6] and this observation, if coupled with that about its Dependent-less status, entails for a dynamic antisymmetric account that NP behaves as a Head and then that NP and Abs create the *special* POS Abs - NP (e.g. *a* - *bāb*), identical to the Dutch POS *heeft* - *wat* dealt with in Section 5.3.2. More abstractly, the special POS at issue can be referred to as the POS X - XP and, as usually, it can be broken either representationally or derivationally.
Assuming that Abs and NP are not coindexed for any feature, neutralization of their symmetry can be only derivational, ie movement, and in principle either Head movement of Abs *a* or remnant movement of NP *a* - *bāb* are viable options.
Whatever the kind of movement actually opted for, in both the cases a higher functional projection immediately above AbsP is needed, in order to provide a landing site for the raised constituent and then ConjP (Cinque's WP) is merged into the syntactic structure.
At this point, as discussed in Section 5.3.2 in connection with the POS *heeft* - *wat*, the derivational neutralization of the special POS X - XP (*a* - *bāb*) avoids X-raising (raising of the Head *a* to the Conj *w*) because *in this step of derivation* a more economical alternative is available, namely XP-raising (remanant movement of *bāb* to Spec, ConjP).
More precisely, it is the Economy Principle (in a technical sense) that bars raising of X (*a*) in favor of XP (*bāb*) because XP instantiates just one syntactic operation, X two, so that the former is more economical than the latter:

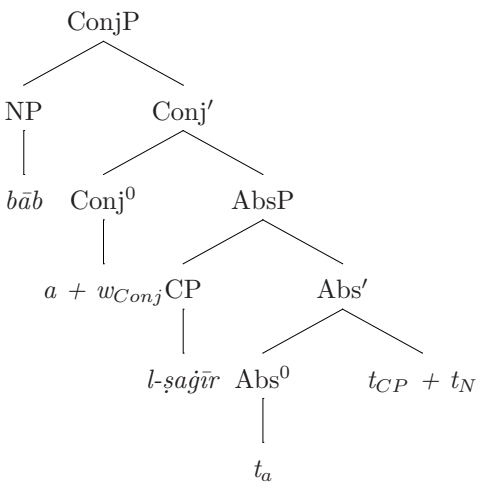
(801) (= 413)

bāb: POS → nonhead movement$_1$ → neutr.
a: POS → head movement$_1$ → POS → Word Formation$_2$ → neutr.

Once the XP *bāb* has been selected for movement, the Minimal Link Condition intervenes, prohibiting it because Spec, ConjP has not *bāb* as its closest constituent, this latter being the NONRERC *l-ṣaġīr* (see Section 5.2.2).
The only way to solve the problem is raising Abs *a* to Conj *w* (Domain Extending Head Movement): this is no doubt a costly solution, because the Head *a* is forced to raise to Conj *w*, thus giving rise with this latter to a further Head - Head construction / POS in need of further neutralization via Word Formation.
But since *in this derivational step* alternative solutions are not found, the same Head movement rejected in the previous derivational step is now forced to apply, *faute de mieux*, in order to satisfy the Minimal Link Condition and then to make the derivation converge (Last Resort Condition: cp. (414) in Section 5.3.2).
Domain Extending Movement thus takes place, to the effect that Abs *a* is left attached to Conj *w*, yielding the phonological representation *aw* of (788), and the NP *bāb* raises to Spec, AbsP (remnant movement). This is depicted in (802):

[6] An additional argument in favor of this assumption is brought in Section 8.10.2 below.

(802)



(802) is the second derivational step that Cinque (2003) assumes for the Arabic NP, which is now reformulated in dynamic antisymmetric terms.

An advantage of the dynamic antisymmetric account proposed here is that *it relates the syntactic phenomenon of XP-raising to the phonological property of monosyllabicity of the nominal stem:*

(803) XP-raising of the noun is a function of its monosyllabic stem

Generality of this statement is confirmed by recent work by Cinque (2010), which brings semantic evidence in favor of an unified treatment of the Indo-european (Germanic and Romance) NP in terms of XP-raising (reconsidering his own earlier analysis invoking X-raising, e.g. Cinque 1994).

Indeed, in Indo-european languages XP-raising of the noun à la Cinque (2010) co-occurs with another property, well known among Indo-europeanists: its being (at some level of representation) a stem *CVC* expanded through affixation of morphemic material (the so-called DETERMINATIVE etc...).

As Bohas (1997, pp. 41-44) himself recognizes, this conception of the Indo-european noun, championed by Benveniste (1935) (see Szemerényi 1996, Ch. VI for more references) makes it possible to set a strict morpho-phonological parallelism between the Arabic stem and the Indo-european one.

Under a dynamic antisymmetric approach, XP-raising of the noun shared by Arabic and Indo-european is a natural consequence of Bohas's remark that they share a common CVC stem pattern[7].

8.5 The POS Conj - DP: representational and derivational neutralization

Ch. 6 has plausibly shown that the derivational step schematized in (802) contains the POS Conj - DP, DP for Op or *d̲ī* depending on whether the NONRERC is the ADJPRESEM or the NONADJAMOR one. The POS at issue is a consequence of the fact that DP, *qua* clitic-like, works as a Head and then it is a Head-within-Spec left-adjacent to the Head Conj.

(802), in particular, improves the antisymmetric account given so far, since the NONRERC

[7]Needless to say, this parallelism is structural, *not* genetic.

modifier (be it the ADJPRESEM or the NONADJAMOR one) is the Compl of ConjP *only subsequent to movement*, in compliance with the architectural settings of the Kaynean Anti-symmetry (see Section 4.5.3). Accordingly, all the previous tree representations describing the same derivational step are rewritten as (802).

Practically speaking, the difference between the former ones and the latter is that the functional projection to which ConjP is merged is no longer the NONRERC (CP), but AbsP, with the NONRERC (CP) embedded into AbsP via movement.

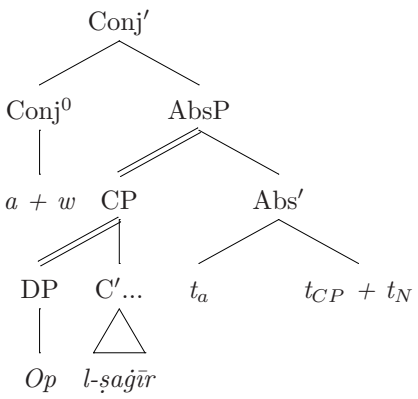
Rethinking phrase structure, however, does not affect the main point: in the POS Conj - DP, Conj c-commands the Head-like DP because the node ConjP dominates both Conj and DP and DP c-commands Conj via the same node (symmetric c-command).

This is so because the first *category* dominating DP is neither CP nor AbsP, but ConjP, which dominates Conj as well: only one *segment* of CP dominates DP (CP being made of two segments: CP and C') and, by the same token, only one *segment* of AbsP dominates DP (AbsP being made of two segments: AbsP and Abs').

But, as already observed in the previous Section, a segment is not enough for c-command, given that in order for such a relationship to be set, an entire *category* is required to dominate DP.

The portion of syntactic structure under discussion is represented in (804), where the segment status of CP and AbsP is indicated by the double line:

(804)



(804) exemplifies the scenario of symmetric c-command using the ADJPRESEM type, but the same applies for its NONADJAMOR counterpart as well, to the effect that in both the cases the symmetric c-command induces the POS Conj - DP and then the need for Syntax to break it, either representationally or derivationally.

It has been suggested in Section 6.4.2, as a first approximation, that both the strategies are at work in the Arabic NP and that the former is found (at least) in the ADJPRESEM type displaying the putative NOM *u*, the latter in the same type when displaying the putative ACC, OBL *a, i*.

The same Section has also pointed out the rationale behind this behavior of the ADJPRESEM type displaying the putative NOM *u* is to be sought *outside* it, in its NONADJAMOR counterpart, the real *locus* of Abstract Incorporation (analogy).

On this view, however, it is not clear why the NONADJAMOR impacts on the ADJPRESEM displaying the putative NOM *u*, while it does not on the ADJPRESEM displaying the putative ACC, OBL *a, i*:

(805) (= 516)

a. ...[$_{\text{Conj}'}$ $\boxed{0_i}$ [$_{\text{CP}}$ $\boxed{\underline{d}\bar{\imath}_i}$... ⇔ ...[$_{\text{Conj}'}$ $\boxed{\text{w}_i}$ [$_{\text{CP}}$ $\boxed{\text{Op}_i}$... > Abstract Inc.
b. ...[$_{\text{Conj}'}$ w [$_{\text{CP}}$ Op... > Movement

The results of Ch. 7, summarized in Section 7.8, allow a more precise formulation of this hypothesis.

First, the alleged opposition NOM *u* vs. ACC, OBL *a, i* found in (the OA manifestation of) the ADJPRESEM is better understood as an opposition Subject vs. non-Subject, which in its turn develops out of a *predicational* opposition NO vs. YES Predicate Inversion.

Second, the opposition NO vs. YES Predicate Inversion is a matter of *dialectal* variation: the Hijazi branch instantiates the ADJPRESEM where no Predicate Inversion occurs, the Bedouin branch the same RC type where Predicate Inversion does[8].

Third, the contrast NO vs. YES Predicate Inversion opposing the Hijazi branch to the Bedouin one has a correlate on the sound-side: the contrast (in the phonological representation) NO vs. YES CPI *'i*. This amounts to saying that the alleged opposition of case-endings NOM *u* vs. ACC, OBL *a, i* on *phonetic* level is to be rethought of as an opposition *w* + NO *'i* (surfacing as *u*) vs. *w* + YES *'i* (surfacing as *a, i*), as per (788).

Fourth, the same contrast NO vs. YES Predicate Inversion opposing the Hijazi branch to the Bedouin one is also found in their NONADJAMOR type, which exhibits the CPI *ya-* (etymologically connected with the Pre-semitic CPI *'i*) in the Bedouin branch, but not in Hijazi.

Fifth, the contrast NO vs. YES Predicate Inversion opposing the Hijazi branch to the Bedouin one (regardless of whether the NONRERC is ADJPRESEM or NONADJAMOR) correlates with another contrast, namely NO vs. YES desemantized demonstrative *d-*:

(806) (= 732)

Hijazi: desemantized (otiose) *d*: NO ⇔ Predicate Inversion: NO
Bedouin: desemantized (otiose) *d*: YES ⇔ Predicate Inversion: YES

On these grounds, ability of the NONADJAMOR to trigger Abstract Incorporation on its ADJPRESEM counterpart (in traditional terms, that exhibiting the putative NOM *u*), to the effect that both their POS Conj - DP are representationally neutralized, intuitively has to do with their Hijazi origin, and in particular with the *unaltered meaning* the demonstrative base *d-*.

On the other side, inability of the NONADJAMOR to trigger Abstract Incorporation on its ADJPRESEM counterpart (in traditional terms, that exhibiting the putative ACC, OBL *a, i*), to the effect that both their POS Conj - DP are *not* representationally neutralized, intuitively has to do with their Bedouin origin, and in particular with the desemantized (pleonastic, 'otiose') usage of the demonstrative base *d-*. As a consequence, both the types of NONRERC in the Bedouin branch are forced to resort to the derivational neutralization of their POS, namely movement (Predicate Inversion).

(805) is therefore rewritten as (807):

[8]Recall from Section 7.7 that in OA the Bedouin ADJPRESEM type evolves into a non-Subject basically because OA is a mixture of the Hijazi and Bedouin branches. When these two branches fuse into a single language, in fact, the *Bedouin* ADJPRESEM type, characterized by *Predicate Inversion*, enters into competition with its *Hijazi* equivalent, where *no Predicate Inversion* is found, for the role of Subject and loses the competition because the informational profile of this latter is more suitable for the role of Subject than its. Consequently, the Bedouin ADJPRESEM type polarizes into a non-Subject (cp. (731)).

(807) a. Hij. ...$[_{\text{Conj}'}$ $\boxed{0_i}$ $[_{\text{CP}}$ $\boxed{\underline{d}\bar{\text{i}}_i}$... $\Leftrightarrow$ ...$[_{\text{Conj}'}$ $\boxed{\text{w}_i}$ $[_{\text{CP}}$ $\boxed{\text{Op}_i}$... > Abstract Inc.
b. Bed. desemantized (otiose) *d* $\Leftrightarrow$ ...$[_{\text{Conj}'}$ $\boxed{\text{w}}$ $[_{\text{CP}}$ Op... > Pred. Inv.

This account is given a more accurate characterization in the next two Sections.

8.5.1 Representation: Abstract Incorporation in situ

Ch. 6 has plausibly shown that in the Hijazi branch it is coindexing between the null Conj and the relative pronoun *d̲ī* that constitutes the ultimate source of representational neutralization of the POS Conj - DP in the NONADJAMOR (*0* - *d̲ī*) and then, by analogy, in its ADJPRESEM (*w* - *Op*):

(808) (= 807.a)
a. Hij. ...$[_{\text{Conj}'}$ $\boxed{0_i}$ $[_{\text{CP}}$ $\boxed{\underline{d}\bar{\text{i}}_i}$... $\Leftrightarrow$ ...$[_{\text{Conj}'}$ $\boxed{\text{w}_i}$ $[_{\text{CP}}$ $\boxed{\text{Op}_i}$... > Abstract Inc.

In particular, coindexing between the two poles of the Amorite POS *0* - *d̲ī* is the result of their sharing a Demonstrative feature: the Conj 0, by virtue of its *rapport associatif* with the Amorite Conj *t̲umma* of demonstrative nature, the relative pronoun *d̲ī*, because of its clear demonstrative origin, still documented in South Arabian.
The POS at stake is thus neutralized and the derivational step they occur in can be sent to Spell-out, ending the derivation.
Spell-out of the NONADJAMOR type has been already dealt with in Section 6.3.2: a modern example is Egyptian Arabic *il-king...illi huwa kān malik...* 'the king..., who was king of' (477). In this case, the POS Conj - DP involves not the old morphemic material *0* - *d̲ī* but the more recent one *illi* - *huwa* (cp. the word-like unit wa_{Conj} - *huwa* > *wahu*).
It should be added that another development of the POS *0* - *d̲ī* consists of *la*-affixation, where *la*- is an Assev(erative) particle fulfilling a dedicated functional projection merged immediately above that hosting the Subject and the Predicate, à la Zanuttini (1997):

(809) ... $[_{\text{AssevP}}$ NP $[_{\text{Assev}'}$ (l)la- $[_{\text{ConjP}}$ t_{NP} $[_{\text{Conj}'}$ 0d̲ī...

In (809) NP raises to Spec, AssevP once again to break the special POS Assev - NP: Assev c-commands NP via the node AssevP, and the other way around (ConjP is a segment and hence plays no role in this relationship). Assuming that Assev and NP are not coindexed, the POS is derivationally neutralized raising to Spec, AssevP XP (NP) instead of X (Assev) for reasons of Economy (see (801)). Notice also that in CA and in many MCV *lla d̲ī* and its variant *lli* etc... are no longer felt as bimorphemic and turn into a monomorphemic Conj (see the just mentioned Egyptian Arabic example) or C.
As for its ADJPRESEM counterpart, its Spell-out also includes manipulation of its phonological representation *a* + *w*, converted into the phonetic representation *u* via contraction etc... as discussed at the end of Section 7.6. This results in a modification structure like *bāb-u l-ṣaġīr* (663).

8.5.2 Derivation: From Spec, AbsP to Spec, DefP

The phenomenon of Abstract Incorporation à la Baker (1988) requires coindexing of the two Heads involved in it (see above).
The dynamic antisymmetric account developed in the present work then predicts that in the

Bedouin branch representational neutralization of the POS Conj - DP in the NONADJAMOR (*0* - *d̲ī*) and then, by analogy, in its ADJPRESEM counterpart (*w* - *Op*) does not take place because no coindexing occurs between the two Heads acting as its poles.
The prediction is realized, if one considers that while the null Conj carries the Demonstrative feature, the relative pronoun *d̲ī* no longer does since in the Bedouin branch the desemantized, that is the *non-demonstrative* *d̲* interferes with it:

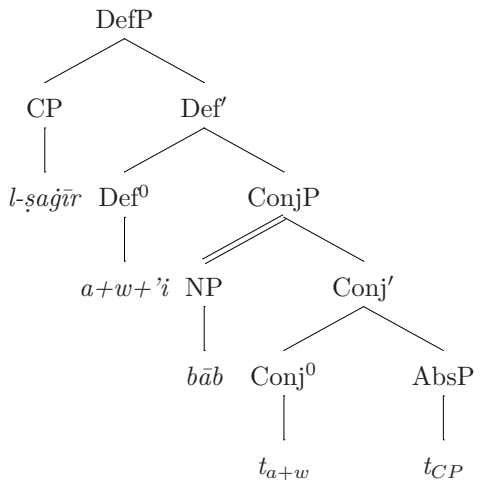
(810) (= 807.b)
b. Bed. desemantized (otiose) *d̲* ⇔ ...[$_{\text{Conj}'}$ w [$_{\text{CP}}$ Op... > Pred. Inv.

Representational neutralization cannot then be invoked to break the POS, and its derivational neutralization is resorted to instead. In other words, *the ultimate source of movement is the otiose* *d̲* or, more generally:

(811) A pleonastic pronoun can act as an (indirect) trigger of movement

Interference of the otiose *d̲* explains why the Bedouin branch, differently from the Hijazi one, does not end its derivation with the step schematized in (802) and requires addition of more syntactic structure (more derivational steps).
In greater detail, Syntax merges into the syntactic structure DefP, headed by the (former) demonstrative item *ʾi* and since the POS Conj - DP is a special POS X - XP, the same scenario discussed in Section 8.4 *recursively applies*.
XP-raising is preferred over X-raising and XP is raised not alone but as part of CP because it is an anaphoric element in need of the governing category provided by the IP contained in CP (Principle A: end of Section 8.3). Finally, X-raising takes place only to help XP-raising (Minimal Link Condition and Domain Extending Head Movement).
This means that the entire NONRERC (CP) moves across the noun it modifies, and the complex Head *a* + *w* attaches to the left of *ʾi*, yielding the phonological representation *awʾi* of (788). The result is the Inverted Predication structure illustrated in (812):

(812)



Yet, a more careful look to (812) shows that the very merger of Def *ʾi* into the syntactic structure solves one problem but creates another: true, it breaks the POS Conj - DP, but it creates the new POS Def - NP too.
It has been seen in Section 8.4, in fact, that NP is a Head-like (clitic-like) element because it is

Dependent-less (subsequent to raising of its NONRERC modifier) and (underlyingly) monosyllabic. Furthermore in the derivational step depicted in (812) NP c-commands Def via the node DefP and Def c-commands NP via the same node (symmetric c-command), because ConjP, the first node immediately above NP, is a segment unable to dominate it (see Section 8.5 and (812), where the segment at issue is indicated by the double line). Hence, the (special) POS Def - NP arises, exemplified in (812) by the complex *aw'i - bāb*.
It is precisely presence of the POS Def - NP that inhibits this derivational step to be shipped off to A-P for Spell-out: as far as it is known, (812) is a construction not documented in Arabic (cp. the 'adjectival' constructions listed in Section 3.2.2.1). Neutralization of this POS is dicussed in the next Section.

8.6 The POS Def - NP: representational and derivational neutralization

It is put forward here that the choice between representational or derivational neutralization of the POS Def - NP should take into consideration a key-property of Def, namely its *grammaticalized* nature: actually the nonterminal *'i* merged as Def is a former demonstrative base (still in usage in Tripolitanian Arabic: see (699) in Section 7.5).
The argument is built as follows. Recent work in typological research by Hopper & Closs Traugott (2003) evidences that grammaticalization is a *clinal* rather than an abrupt phenomenon, so that a given item undergoing it in diachrony passes through an intermediate stage of development. A particularly telling example of this in Arabic itself is the diachronic shift of *-n* from Animate Classifier into C, a process involving some gradual steps (cp. (622) at the end of Section 6.6.3).
Typology therefore lends indirect support in favor of the claim made here that the Demonstrative *'i* evolves into the Def *'i* through an intermediate diachronic stage in which it is a Def but still endowed with a Demonstrative feature:

(813) $\text{Demonstrative}_{Def:NO} > \boxed{\text{Def}_{Demonstrative:YES}} > \text{Def}_{Demonstrative:NO}$

It is thus expected that in some earlier stage of language Def *'i* carries the Demonstrative feature, while losing it later on.
Turning to another phenomenon of diachronic development observable in the Arabic NP, Gensler (2004) convincingly argues that in the Arabic RC the non-restrictive reading is older than the restrictive one, and in Section 6.3.2 it has been demonstrated that the syntactic correlate of this piece of information is raising of one or many constituents, *if not even of an entire sentence*, over the C introducing the NONRERC (cp. the Basque examples in (483)): shortly put, the word order RC > C is associated with the non-restrictive meaning.
As a corollary, the *RC > C word order, if found in Arabic, can be taken to be a plausible clue of an earlier stage of it*, while its lack points to a later stage.
At this point it seems convenient to call again into the picture the so-called 'transposed NP structure' found in Najdi Arabic and reported by Ingham (1994), actually an instance of Inverted Predication structure (see Section 7.5) :

(814) Najdi Arabic (= 703)

mbaccir a-l-wasim
early *a*-C-winter.rain

'early winter rain'

As already pointed out in Section 7.5, the Def *a* (whose phonological representation is *aw'i*) as well as the C *l-* detectable in (814) qualify it as an ADJPRESEM. Of particular relevance here is that the RC having undergone Predicate Inversion in (814) is a C-final RC too: the subordinate clause *(Op) mbaccir* precedes *l-*.
It follows that the Inverted Predicate / modifier of Najdi Arabic transposed NP structure is an archaic construction because of the word order RC > C (see immediately above).
Combining the two just made typological observations, there emerges a diachronic parallelism between evolution of the NONRERC and that of Def. In particular, the old Def still having a demonstrative value might have plausibly coexisted with the old Word Order RC > C:

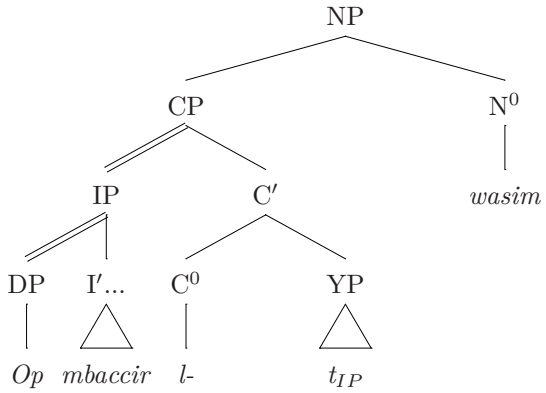
(815) Def: Demonstrative value: YES > NO
NONRERC: Word order RC + Complementizer: YES > NO

By virtue of (815), in the Najdi Arabic transposed NP structure (814) the Def *'i* of the phonological representation *aw'i* underlying the phonetic representation *a* (at least in part) preserves a demonstrative nature, since it displays the word order RC > C in its Inverted Predicate / modifier.
On the other side, the Def found in the derivational step in (812) has lost its demonstrative nature because its modifier displays a word order other than RC > C (only Op precedes C).
It will be argued in the two next Sections that the older transposed NP structure (814) still attested in Najdi Arabic neutralizes the POS Def - NP (*aw'i* - *bāb*) representationally, the more recent Reinstated Word Order derivationally.

8.6.1 Representation: Abstract Incorporation in situ

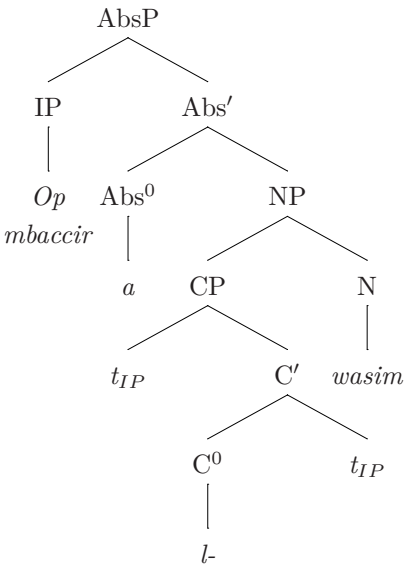
To put the transposed NP structure (814) in minimalist perspective, a step back is needed: the derivational history of this instance of Inverted Predication is still unclear.
(814) being basically a ADJPRESEM passes through the derivational steps of this latter: it minimally but crucially differs from the ADJPRESEM discussed so far in that its Spec, CP hosts not just the Subject of the RC (the relative pronoun DP / Op) but the entire RC, which yields the 'C-final effect'.
It seems therefore safe to assume that the transposed NP structure (814) is base-generated as illustrated in (816), where accordingly functional material (Abs *a*, Conj *w*, Def *'i*) is not yet found:

(816)



Under standard (pre-)minimalist assumptions (see (236) at the end of Section 4.5.3), the RC minus CP is interpreted as IP (which raises to Spec, CP via phrasal movement).
In this derivational step the POS DP - N (*Op - wasim*) occurs along the same lines of Section 8.3. Notice that the IP intervening between DP (Op) and the node NP that symmetrically c-commands DP and N is not relevant, because IP like CP is only a segment (cp. the double line in (816)).
Neutralization of this POS exhibits the same dynamics discussed in Section 8.3. Lack of adjacency of DP - N makes their representational neutralization not viable (*mbaccir* intervenes between them). XP-raising of DP is preferred over X-raising of N (cp. the Economy Principle in (801)). Finally DP (Op), being an anaphoric element, cannot raises alone but as part of its governing category by virtue of the Principle A (see end of Section 8.3).
It should be noted in this connection that *fronting of the RC over C makes the governing category of DP (Op) (816) not CP but IP since in CP itself it is I that governs Op and agrees with it*. Accordingly, what raises to Spec, AbsP is IP (*Op mbaccir*) *while the C is left in situ*:

(817)



But the C *l-* in this syntactic position gives rise to a new POS. Consider NP in (817): it symmetrically c-commands CP and N. Furthermore, CP is monosyllabic and Dependent-less (subsequent to raising of IP to Spec, AbsP). This amounts to saying that CP and N create the POS CP - N (*l - wasim*), which can be in principle neutralized representationally because the

two poles are adjacent: only a trace intervenes between them and such an element does not inhibit Abstract Incorporation, as per Baker (1988, p. 202), probably because it is not audible in Spell-out.

Yet, there is still a requirement to be met in order to achieve Abstract Incorporation: CP and N must be coindexed.

This is actually the case, because *l-* is a C of demonstrative nature (as it often happens crosslinguistically: see Section 5.2.4) and Arab Grammarians' judgments converge with modern typology in assigning the demonstrative the syntactic category 'noun' (see Sections 5.2.1 and 3.2.2.1). The POS CP - N (*l - wasim*) is thus neutralized, yielding the word-like unit *lwasim* or the alike, so that what Syntax 'sees' in (817) after that Abstract Incorporation takes place is a NP identical to the NP *bāb* of the derivational step discussed in Section 8.3.

After all, the NP adorned with the C *l-* (*lwasim*) carries the same phonological information of its *l*-less counterpart *bāb* because it is underlyingly a monosyllabic stem *CVC* (cp. Section 8.4) and affixation of *l- does not modify the syllabic structure (lCVC still counts as one syllable).*

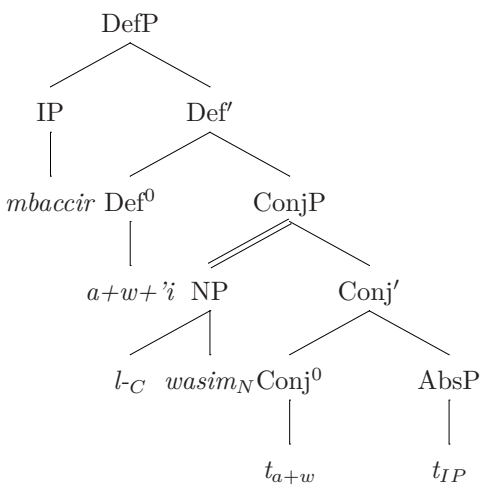
In addition to monosyllabicity, the NP *lwasim* of (817) is Dependent-less because its RC Dependent has almost totally moved outside NP under the form of IP, and the remaining portion of it (C) is no longer a Dependent because it has fused with N into a single word .

The NP *lwasim* thus characterized is right-adjacent to the Head Abs, whose terminal in the phonological representation is *a*, to the effect that they yield the POS Abs - NP (*a - lwasim*). This latter is derivationally neutralized along the same lines of the POS Abs - NP *a - bāb* dealt with in Section 8.4 etc...

More generally, setting the equivalence between the NP *lwasim* of (817) and the NP *bāb* of the derivational step discussed in Section 8.3 implies that derivation of the transposed NP structure *in nuce* (817) proceeds along the lines discussed in the previous Sections, ending up in the Inverted Predication Structure (818), the counterpart of the Inverted Predication Structure (812).

It is worth observing in this connection that the contrast between recursive IP-raising in (818) and recursive CP-raising in (812) *is only apparent, in the sense that in both the cases the raised phrase is the result of one and the same requirement imposed by the Principle A*, namely the need to provide Op with its governing category (see end of Section 8.3).

(818)



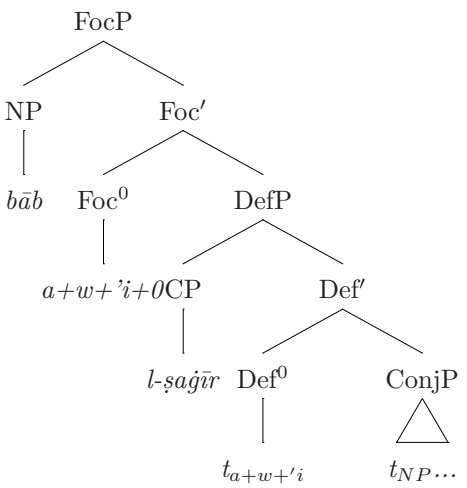
As already observed with regard to (812), under a dynamic antisymmetric approach the key-property of the Arabic Inverted Predication structure occurring in primary predication is the POS Def - NP, in this case manifesting itself as *aw'i - lwasim*.
Insofar as (818) is concerned, neutralization of the POS at issue is representational, on the ground that the functional item *'i* base-generated in Def (to which Abs *a* and Conj *w* head-move as a consequence of Domain Extending Head Movement) still preserves a demonstrative nature (see (815)).
On this basis, Def carries the syntactic feature N and hence turns out to be a Head coindexed for the syntactic feature N with the Head-within-Spec immediately below it, namely the NP *lwasim*.
Once that the POS is neutralized, the syntactic structure in (818) is shipped off to A-P, in order to be spelled out, after that the phonological operations typical of this module have taken place. They are the rules of *hamzah*-Deletion, Glide Contraction, Long Vowel Shortening motivated in Ch. 7, converting the phonological representation *aw'i* into the phonetic representation *a* etc... The output of these processes is precisely the Najdi Arabic transposed NP structure *mbaccir a-l-wasim* (814)[9].
This Section therefore captures in dynamic antisymmetric terms the phrase structure of the Najdi Arabic transposed NP structure and especially the *C-final word order of its RC modifier*, justified on independent grounds (syntactic diagnostics) throughout the previous Chapters.

8.6.2 Derivation: From Spec, ConjP to Spec, FocP

The other side of the token of the principled correlation established in (815) is that an RC modifier not affected by clausal fronting co-occurs with a Def *'i* devoid of a demonstrative value and hence of the syntactic feature N.
Returning to the Inverted Predication Structure depicted in (812), one may capitalize on this observation to find that the Def *'i* of the complex Head *aw'i* can be described as such, because in the RC modifier on its left no clause is fronted.
Within the framework of Dynamic Antisymmetry, this means that the POS Def - NP *aw'i - bāb* cannot be broken resorting to the representational Abstract Incorporation, because the Def *(aw)'i* is not coindexed with the NP *bāb* for the syntactic feature N, and hence movement of the NP *bāb* is forced to apply.
As already discussed in the previous Sections, this happens because XP-raising (of the NP *bāb*) is preferred over X-raising (of the composite Def *aw'i*) by virtue of the Economy Principle.
The NP *bāb* thus seeks for the first Spec able to act as its landing site. The only viable option is Spec, DefP, but it is already fulfilled by the RC modifier *(Op) l-ṣaġīr*.
Consequently, the functional projection FocP is merged into the syntactic structure, immediately above DefP (cp. (791)) in order to provide a landing site for the NP *bāb* to be raised. At the same time the composite Def *aw'i* raises to the covert Head Foc in order to block raising of the RC modifier *(Op) l-ṣaġīr* to Spec, FocP and to enable the NP *bāb* to raise across *(Op) l-ṣaġīr* towards the same position. This is depicted in (819):

[9]Notice that adopting the Left Branch Condition instead of the Principle A would have not yield the desired word order in the case of the transposed NP structure: the Left Branch Condition would have forced raising of CP, thus not leaving C in situ (the C-final effect typical of the (NONRE)RC of the transposed NP structure)..

(819)



The POS Def - NP of (812) is thus neutralized and its neutralization yields the Reinstated Word Order.

Derivation of the Arabic NP is almost complete: the only phenomenon to be still accounted for within the framework of Dynamic Antisymmetry is the (putative) prenominal article[10]. The next Section tackles this issue.

8.7 The POS D - NP: Abstract Incorporation

Section 6.4.3 has offered phonological, semantic and diachronic evidence that the 'prenominal article' *l-* actually is a medial functional item, of Pre-semitic origin, inserted between an Amorite demonstrative base *d̲-* working as a (quasi-)article and the noun.

L-insertion performs the task of distinguishing the demonstrative / definiteness construction at issue (secondary predication) from a copular sentence (primary predication), otherwise identical with it. Generally in Arabic the demonstrative base *d̲-* has a covert realization for phonological reasons, namely a well-documented constraint on the co-occurrence of the two interdentals *d̲ - l*, which typically results in dropping of *d̲-* (cp. (457)).

The (allegedly) *monomorphemic* 'prenominal article' *l-* is thus better understood as a *bimorphemic* entity made of the Amorite demonstrative *d̲-*, the real *locus* of definiteness, sitting in Spec, DP and of the Pre-semitic disambiguating element *l-*, fulfilling the Head D:

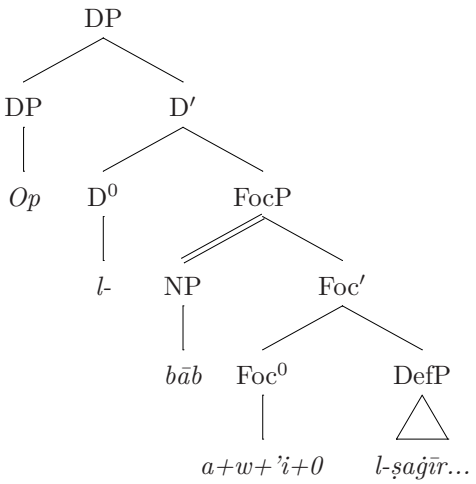
(820) (= 541)
... $[_{\mathrm{DP}}$ Op^{DH} $[_{\mathrm{D}'}$ l- [NP...

The hybrid nature (= Pre-semitic and Amorite) of this construction points to its relatively recent character and in fact similar *bimorphemic* constructions are documented for Semitic only from Phoenician (variety of Byblos) from XIII BC onwards (see (528) in Section 6.4.3). Before the bimorphemic constructions, the demonstrative (or the determiner, if any) occurs alone,

[10]It should be added for the sake of completeness that in (819) prior to raising of NP to Spec, FocP etc... the very merger of the covert Foc to the left of DefP 'virtually' creates a new POS, whose poles are the *covert* Foc and the *covert* DP Op sitting in Spec, CP. It seems however safe to maintain that actually this POS does not exist because both of its poles are covert. If in fact the phenomenon of blind POS discussed for Arabic in Ch. 6 is given an ontological status, at least *one pole of the POS must be overt*, which is clearly not the case for the configuration Foc - DP.

preceding the noun it refers to in the Amorite parastrate (see (520 - 522) in Section 6.4.3), or *follows* it in the Pre-semitic parastrate, although this state of affairs is not documented for Arabic, but only for an archaic Semitic language like Akkadian and for Somali, a Hamitic language (see (502), (518) in Ch. 6 as well as Section 6.4.3).
It is therefore important to bear mind throughout the reading of this Section that the terms ADJPRESEM and NONADJAMOR are used for the sake of simplicity, and abstract away from the fact that the putative prenominal article actually makes these RC types 'impure' types relative to their parastrate.
When Syntax merges the bimorphemic DP *(d) + l* (820) to the left of FocP (cp. (791)) in the relevant stage of language, the following syntactic scenario obtains:

(821)



Adjacency between the D *l-* and the NP sitting in Spec, FocP creates a new special POS. Consider the node DP in (821): it symmetrically c-commands the D *l-* and the NP *bāb* (the node FocP is only a segment and hence does not c-command NP: cp. the double line in (821)). It should be added that NP is (underlyingly) monosyllabic and Dependent-less (see Section 8.4) and this is why the special POS D - NP (*l - bāb*) arises.
It seems safe to maintain that the POS at issue is neutralized representationally, by means of Abstract Incorporation, both because the two poles D and NP are adjacent and because they are coindexed for the syntactic feature N: *l-* is a C of demonstrative nature and demonstratives belong to the syntactic category N (see arguments of Sections 8.6.2).
The POS D - NP, ie *l - bāb* is thus neutralized, yielding the word-like unit *lbāb*. To put if differently:

(822) In Arabic, Word Formation between *l-* and the NP (*l*-prefixation) is a consequence of their representational neutralization *in situ*

Such an account derives the prefixal nature of *l-* without assuming any process of N-to-D raising, as currently assumed in the (pre-)minimalist literature to derive the word-like behavior of *l-* and the noun in Arabic (see e.g. Fassi-Fehri 1993 as well as Hoyt 2006 for more references). Thus, the dynamic antisymmetric account proposed here, while maintaining the idea that *Incorporation* is at work in the complex *l* - noun à la Fassi-Fehri (1993) etc..., implements it in a different way, using *Abstract* Incorporation instead of Incorporation *via Head-movement*.
Invoking Word Formation between D and *NP* (Abstract Incorporation *in situ*) instead of *N-*

to-D raising (Incorporation via Head-movement) has the advantage of solving the problems encountered by this latter: among them, violation of Mirror Principle (which would yield N-D instead of D-N) and the pretheoretical fact that numerals can intervene between *l-* and the noun, as exemplified in (823).

(823) Levantine Arabic (Feghali, 1928, p. 190)

el-'arba‘ eḥṣne
the-four horses

'the four horses'

Indeed, the numeral intevening between the noun and *l-* within the (pre-)minimalist framework is another instance of violation of Mirror Principle. Consider the order in which N and Num are merged into the syntactic structure: N is a substantive category, base-generated as such lower than Num (see (236) at the end of Section 4.5.3).
Assuming that N-to-D raising applies entails that N before head-moving and left attaching to D, head-moves and left-attaches to Num, yielding the word order *l* > N > Num, but clearly this latter is not found in (823)[11].
Under the dynamic antisymmetric account, no such problem arises, because it relies upon the antisymmetric / cartographic scenario à la Cinque (2003) / Shlonsky (2004), where N-raising in abandoned in favor of NP-raising.
Returning to the representational neutralization of the POS D - NP (*l* - *bāb*), it ends the syntactic derivation, so that (821) is sent to A-P, where the rules of *hamzah*-Deletion, Glide Contraction, Long Vowel Shortening motivated in Ch. 7 convert the phonological representation *aw'i* into the phonetic representation *a* etc...
The same holds for all kind of Reinstated Word Order displaying the phonetic representation *i* instead of *a*. Its derivation runs much the same way of its counterpart displaying the phonetic representation *a*, and the former only differs from the latter in that its phonological representation has a covert ABS: *0w'i* instead of *aw'i*, as discussed in the previous Chapter.
This outcomes in spell-out of the Bedouin ADJPRESEM, an instance of Reinstated Word Order along the lines of Najdi Arabic *al-bēt a-l-ṭuwīl* 'the tall house' (787) etc...
The same derivation applies to the Bedouin NONADJAMOR, (although its copulae are less evident on phonological level, because only the Def *ya-* is audible). The final step of its derivation, the equivalent of (821), is depicted in (783).

8.8 Residual issues

The careful reader will have noticed that spell-out of the (NONRE)RC adorned with the prenominal *l-*, irrespective of whether it is ADJPRESEM or NONADJAMOR, phonologically realizes this item as *al-*, to the effect that its phonetic representation manifests an *a*-vowel absent in its phonological representation.
The *a*-vowel (as well as its variants: *ə, i* etc...) preceding the *prenominal l-* is likely to be *epenthetic* and to be kept distinct from the *morphemic a* preceding the *medial l-*, whose phonological representation is *aw'i*.

[11] This derivation glosses over the just pointed out problem that the initial position of *l-* itself violates the Mirror Principle, which makes its plausibility even more dubious.

The former, in fact, has a diagnostic property sharply opposing it to the latter, namely its *invariability*. On the one hand, the morphemic *a* preceding the medial *l-* is realized as such in the Bedouin Branch but as *u* in the Hijazi one and, in a likewise fashion, in CA it surfaces as *a* when non-Subject, but as *u* when Subject.
On the other hand, the epenthetic *a* preceding the prenominal *l- keeps unaltered* across both dialectal variation and the CA 'case-system': e.g., as far as it is known, in both the Hijazi branch and CA, the phonological string **u-l- N-(u - l - RC)* never occurs (cp. (824.b) below). Moreover in CA the *u* intervening between the noun and the medial *l-* introducing the RC modifier to mark the Subject, if co-occurring with a prenominal *l-* in sentence initial position has as its counterpart the *invariable* vowel *a*, as exemplified in (824.a):

(824) Classical Arabic (= 42)

a. (')a-l-rajul-u l-karīm(-u)
the-man-NOM the-noble(-NOM)

b. *u-l-rajul-u l-karīm(-u)
idem

'the noble man'

This can be taken to be evidence enough to claim that the prenominal article *l-* is preceded by a vowel *a* or the alike *only in its phonetic representation*, subsequent to a rule of Epenthesis applying at A-P, so that no phonological material precedes the prenominal *l-* in its phonological representation.
This is desirable, given that the phonological representation associated with a non-epenthetic (morphemic) *a* is *aw'i* and this latter in its turn implies that the constituents linked by them are the result of a derivational history involving Canonical Predication, Inverted Predication, reinstatement of the word order Subject > Predicate, which is not the case for the (covert) Op and the NP. Op, in fact, does not fulfill Spec, DP as a consequence of the transformations examined in the previous Sections, being rather base generated as such.
Setting aside the just discussed derivational step, on the whole the derivation of the Arabic modification structure described throughout this Chapter *recursively applies to any other RC modifier added to the first one*.
In this respect, the antisymmetric / cartographic account of the Arabic NP à la Cinque (2003), Shlonsky (2004), Ihsane (2003) has it that the universal hierarchy of the DP functional projections is characterized by the following word order of modifiers:

(825) Dem > Num > Adj > N

and that the Arabic NP displays a mirror-image of the universal word order in (825), at least insofar as the MCV are concerned, since in these latter the following word order is documented:

(826) Moroccan Arabic (= 64)

lə-wlad l-zwiin-iin l-xams-a hadu
the-boys the-handsome-PL the-five-*at* these

'these five handsome boys'

In order to bridge the gap between the universal word order in (825) and the Arabic word order in (826), Cinque (2003), Shlonsky (2004), Ihsane (2003) contend that there applies in the Arabic NP XP-raising of N (and of the material it acquires along its derivation). This theoretical move

is able to derive the desired word order, but has a relevant drawback: no principled explanation is given of why XP-raising of N is preferred over its Spec-to-Spec movement.
Another problem of this analysis is that it rests upon a traditional interpretation of the adnominal markers *l-, -n, -t* and probably of *u, a, i*, if it is extended to CA, as features intrinsic to the Head N: this working hypothesis is not confirmed by diachronic, semantic and dialectological considerations and is to be rejected, as discussed at length in Section 3.2.2.1.
Thus the present work, in line with the antisymmetric / cartographic assumptions adopted so far, retains Cinque's etc... idea that the universal word order (825) yields the Arabic word order (826) through XP-raising but proposes a different implementation of it, whereby the adnominal markers *l-, -n, -t (u, a, i) are not features intrinsic to the Head N* but, as evidenced throughout the previous Chapters, functional items involved in the structural makeup of a (NONRE)RC, the syntactic structure underlying all the modifiers combined with *l-, -n, -t*.
It is put forward here that this shift of perspective straightforwardly accounts for application of XP-raising of N instead of its Spec-to-Spec movement.
To begin with, consider a syntactic representation of (826) based on the analysis worked out in this Chapter - in particular, the derivational step in which the first modification structure *wlad l-zwiin-iin* has been formed. This is made of the NP *wlad* modified by a conjoined RC introduced by the C *l-*, where the 'adjectival' RC *zwiin* is coordinated with the Numeral Classifier Construction - RC *-iin*. The NP *wlad* sits in Spec, ConjP (WP), its modifier in Spec, AbsP (GenP), as per Section 8.3:

(827) ...[$_{\text{ConjP}}$ wlad [$_{\text{Conj}'}$ Conj [$_{\text{AbsP}}$ [$_{\text{CP}}$ Op [$_{\text{C}'}$ l- [$_{\text{XP}}$ zwiin-iin]]] [$_{\text{Abs}'}$ Abs...

At this point of the derivation, Syntax merges into the structure the Numeral *l-xams-a*, to be analyzed along the same lines of *zwiin-iin*:

(828) ...[$_{2-\text{AbsP}}$ [$_{\text{CP}}$ Op [$_{\text{C}'}$ l- [$_{\text{XP}}$ xams-ah]]] [$_{2-\text{Abs}'}$ 2-Abs [$_{1-\text{ConjP}}$ wlad [$_{1-\text{Conj}'}$ 1-Conj [$_{1-\text{AbsP}}$ [$_{\text{CP}}$ Op [$_{\text{C}'}$ l- [$_{\text{XP}}$ zwiin-iin]]] [$_{1-\text{Abs}'}$ 1-Abs...

The very merger of the second Head Abs (2-Abs) is responsible for a new POS: 2-Abs and the NP *wlad* create the blind POS 2-Abs - NP, to be neutralized derivationally because of lack of coindexing between the two poles. This is the derivational step dealt with in Section 8.4 and it is therefore expected that the NP *wlad* sitting in Spec, 1-ConjP move to Spec, 2-ConjP via Domain Extending Head Movement. This movement would result in the word order N > Num > Adj that however is at odds with (826). (829) illustrates this state of affairs ($\square$ indicates the raised constituent):

(829) ...[$_{2-\text{ConjP}}$ $\boxed{\text{wlad}_i}$ [$_{2-\text{Conj}'}$ 2-Conj [$_{2-\text{AbsP}}$ [$_{\text{CP}}$ Op [$_{\text{C}'}$ l- [$_{\text{XP}}$ xams-ah]]] [$_{2-\text{Abs}'}$ 2-Abs [$_{1-\text{ConjP}}$ $\boxed{\text{t}_i}$ [$_{1-\text{Conj}'}$ 1-Conj [$_{1-\text{AbsP}}$ [$_{\text{CP}}$ Op [$_{\text{C}'}$ l- [$_{\text{XP}}$ zwiin-iin]]] [$_{1-\text{Abs}'}$ 1-Abs...

In order to obtain the correct word order N > Adj > Num, what is needed is raising of the entire 1-ConjP (the entire modification structure *wlad l-zwiin-iin*), not just of its Spec (the NP *wlad*) to Spec, 2-ConjP. This is illustrated in (830), where $\square$ indicates the raised constituent:

(830) ...[$_{2-\text{ConjP}}$ $\boxed{\text{[}_{1-\text{ConjP}}\text{ wlad [}_{1-\text{Conj}'}\text{ 1-Conj [}_{1-\text{AbsP}}\text{ [}_{\text{CP}}\text{ Op [}_{\text{C}'}\text{ l- [}_{\text{XP}}\text{ zwiin-iin]]] [}_{1-\text{Abs}'}\text{ 1-Abs...]]]}_i}$ [$_{2-\text{Conj}'}$ 2-Conj [$_{2-\text{AbsP}}$ [$_{\text{CP}}$ Op [$_{\text{C}'}$ l- [$_{\text{XP}}$ xams-ah]]] [$_{2-\text{Abs}'}$ 2-Abs $\boxed{\text{t}_i}$...

But why is Spec-to-Spec movement of the NP *wlad* alone in (829) barred? Capitalizing on den Dikken (2006, pp. 39-43), it is argued here that the rationale behind preferring XP-raising over

Spec-to-Spec movement is semantic, namely ambiguity of scope. Consider again (829), in the following simplified representation:

(831) ...[wlad] ...[Op l-xams-a]... [Op l-zwiin-iin]...
...X ...Y... Z...

In (831) the lower Op combined with *l-zwiin-iin* by definition is a coreferential element and by virtue of Relativized Minimality (see Section 4.5.1) it refers back to the higher Op combined with *l-xamsa*, immediately preceding it. In particular the higher Op provides the lower Op with the referential meaning this latter needs by means of a 'detour': the NP *wlad* asymmetrically c-commands the higher Op, so that the former transfers onto the latter its referential features making interpretation of the higher Op possible. In its turn, the higher Op transfers the referential features got by the NP *wlad* onto the lower Op. To put it shortly, the lower Op gets its referentiality through the chain: $wlad_i$ - Op_i - Op_i.
This mechanism, however, raises an interpretive problem of scopal ambiguity, because the higher Op scopes over the lower Op, to the effect that the semantic information carried by this latter can be ambiguously interpreted as dependent on the semantic information carried by the higher Op. An informal paraphrase of the meaning of (831), ie of the syntactic structure (829) resulting from Spec-to-Spec movement, is *not* that 'the boys are intrinsically beautiful and at the same time in number of five', which is the actual meaning conveyed in (826).
Instead, the fact that the Op combined with *l-xams-a* takes scope over the Op combined with *l-zwiin-iin* is responsible for a meaning of (831) (= 829) informally paraphrasable as: 'they are beautiful as a consequence of the fact that they are five', which is *not* the intended meaning of (826).
If, on the other hand, XP-raising of N along with its modifier (1-ConjP) applies, there obtains (830), whose simplified representation is (832):

(832) ...[wlad [Op l-zwiin-iin]]... [Op l-xams-a]...
...X... Z...

In (832) the Op combined with *l-xams-a* does not scope over the Op combined with *l-zwiin-iin* because the former is lower than and hence does not asymmetrically c-command the latter. Consequently, the undesired reading 'they are beautiful as a consequence of the fact that they are five' does not arise in (832). In a likewise fashion, the Op combined with *l-zwiin-iin* does *not* scope over the Op combined with *l-xams-a* because the former has not the ability of asymmetrically c-commanding the latter (which would yield the undesired reading 'they are five as a consequence of the fact that they are beautiful'): this is a consequence of the fact that the Op combined with *l-zwiin-iin* is *embedded* into the NP *wlad*.
As such it is this latter that at the same time asymmetrically c-commands *both* the Op combined with *l-zwiin-iin* and the Op combined with *l-xams-a*, which yields the desired reading 'the boys are intrinsically beautiful and at the same time in number of five' - structurally, in fact, both the Op are on the same footing.
What precedes is highly reminiscent of the scenario of semantic ambiguity described by den Dikken (2006) (in his turn developing Cinque 1999), where the adverbials take ambiguous scope over the verbs. Such a scenario leads to motivate the XP-raising posited in the Arabic NP as follows:

(833) Spec-to-Spec movement is barred and XP-raising applies instead for reasons of convergence with C-I: the need to avoid ambiguity in the scope of the Op found in the Arabic RC modifier.

8.9 Relativized Minimality in the pronominal Construct State

The dynamic antisymmetric account worked out in the previous Sections does not derive the pronominal CS because in this construction the modifier of the noun is *not* a (NONRE)RC, at least when it denotes *inalienable* possession (see Ch. 6 and especially Section 6.6.4).
In the best case, such an approach should be able to explain a phenomenon observed by Aikhenvald (2000) for the Classifier-less Possessive Construction of which the pronominal CS is the Arabic instantiation.
This is the correlation between Word Formation and inalienable reading of the Possessee: for example, as already pointed out in Section 6.6.2, in Tariana the (equivalent of) the pronominal CS is a word-like unit if its Possessee is inalienable, while it is better seen as mere juxtaposition of the Possessee and the Possessor if the Possessee is alienable.
Aikhenvald (2000) exemplifies this scenario through the contrast existing in Tariana between (589) and (590). As for Arabic, the same situation is found in the opposition between long vs. short vowel in the pronominal CS. In Arabic, in fact, the long vowel is a clue of a word-like unit and the vowel of the putative case ending in the pronominal CS is long iff and only if the Possessee is (unambiguously) inalienable (Section 6.6.2):

(834) Old / Classical Arabic (= 584, 593)

a. ’ab-$\boxed{\text{ū}}$-ka
father-COP-you

b. tijārat-$\boxed{\text{u}}$-hum
commerce-COP-hum

‘your father’

In the (pre-)minimalist literature, mere juxtaposition of two given syntactic constituents is designated as PF-MERGER or MERGER UNDER ADJACENCY[12].
In its syntactic implementation, this term refers to two constituents whose juxtaposition in Spell-out corresponds to their fulfilling in Syntax two *adjacent but distinct positions*, in the sense that adjacency of these constituents is not the result of Head movement (Incorporation), nor is their adjacency ‘converted’ into a single word by Abstract Incorporation.
A familiar example from English is juxtaposition between the tense / agreement marker *es* and the verbal stem in emphatic forms such as *he does not write*. Their word order in fact does not arise from Head movement, nor Abstract Incorporation makes them a single word because they are not coindexed etc...They surface as such simply because they are base-generated as such into the syntactic structure (see Bobaljik 1994 as well as Section 8.10.2 for details):

(835) ...$[_{I'}$ -es ...$[_{V'}$ write...

[12]As well as LOWERING or AFFIX HOPPING in pre- or non-antisymmetric frameworks (e.g. Chomsky 1957), where juxtaposition is regarded as involving some kind of downwards or rightwards movement. This is not possible in approaches more or less in line with the Kaynean Antisymmetry (see Section 4.5.3), which therefore interpret juxtaposition in Spell-out as juxtaposition in Syntax. See the ongoing discussion.

Focusing on the internal consistency of the theoretical framework, Baker (2002) explicitly states that Merger under adjacency has a theoretical status within an approach combining (Abstract) Incorporation with cartography and antisymmetry, and supports his claim bringing crosslinguistic and typological evidence. This ensures that Merger under adjacency is a plausible working hypothesis for the present work.

The picture that emerges from the discussion so far is that the correlation observed by Aikhenvald (2000) can be schematized as follows:

(836) Word Formation ⇔ Inalienable Possessee
Merger under adjacency ⇔ Alienable Possessee

Aikhenvald (2000, pp. 132ff.) gives a phonosymbolic explanation of the correlation in (836). The gist of her reasoning is that metaphorically the inalienable Possessee noun is thought of as closer to its Possessor than the alienable one: thus, the more distant the Possessor from its Possessee in the real world, the more distant the word denoting the former from the word denoting the latter.

Nevertheless, phonosymbolism, a *semantic* factor, appears to be too a limited trigger of Word Formation: it can be hardly taken to be at work outside the specific case of the pronominal CS referring to an inalienable Possessee. Yet, Word Formation in natural languages is a widespread phenomenon, and therefore Aikhenvald (2000)'s account misses a crucial generalization.

On the other hand, the development of Dynamic antisymmetric put forward in the present work, namely the *representational* neutralization of a POS, has the advantage of offering an alternative rationale for Word Formation, which does not run into the problems of an account à la Aikhenvald (2000) because the former is the exact opposite of the latter: the *representational* neutralization of a POS triggering Word Formation is *semantics-independent* and *general*.

It seems therefore desirable to provide another explanation of the correlation in (836), resorting to Dynamic Antisymmetry.

Taking Ouhalla (2010)'s minimalist account of the pronominal CS (Section 6.6.4) as starting point for the discussion, this construction, when referring to an *inalienable* Possessee would have a syntactic structure along the following lines:

(837) $[_{N'}$ N $[_{DP}$ Pronoun

The Shamaliya Arabic equivalent of 'his sister', ie *xt-hu* (576) exemplifies the structure in (837)

(838) $[_{N'}$ xt $[_{DP}$ hu

Nevertheless, the fact that a Conj intervenes between the Possessee and the pronominal Possessor in CA / OA (cp. (834)) - manifesting itself either as a long or as a short vowel - allows to assume the same state of affairs for Shamaliya Arabic. In this light, OA / CA and Shamaliya Arabic (or more generally the MCV) only differ in the phonological realization of Conj: overt in the former, covert in the latter (ones).

(839) $[_{ConjP}$ 'ab $[_{Conj'}$ ū (= a+w) $[_{AbsP}$ ka...
$[_{ConjP}$ xt $[_{Conj'}$ 0 $[_{AbsP}$ hu...

But the dynamic antisymmetric account of the configuration in (839) worked out throughout this Chapter implies that it is a derivational step preceded by the derivational step in (840):

(840) $[_{AbsP}$ ka $[_{Abs'}$ a $[_{NP}$ t_{ka} $[_{N}$ 'ab...
$[_{AbsP}$ hu $[_{Abs'}$ 0 $[_{NP}$ t_{hu} $[_{N}$ xt...

in its turn derived from the following base-generation structure:

(841) $[_{NP}$ ka $[_{N}$ 'ab]]
$[_{NP}$ hu $[_{N}$ xt]]

It ensues that Dynamic Antisymmetry makes it possible to retain Ouhalla (2010)'s idea that the Possessee and the Possessor are adjacent in the pronominal CS, simply by projecting back this scenario to a more primitive stage of derivation.

Within the framework of Dynamic Antisymmetry, (841) transforms into (840) because of the POS DP - N (on the details of which, see Section 8.3): either *ka - 'ab* in OA / CA or *hu - xt* in the MCV.

Notice that the pronominal DP *ka, hu* etc... differs from the anaphoric DP *Op, d̲ī* found in the same POS (see Section 8.3) in a fundamental respect: it is a pole *adjacent* with the other pole of the POS (N).

If one considers that both DP and N are coindexed for the same syntactic feature N, it is expected that neutralization of this POS be representational (Abstract Incorporation). Yet, this expectation is not met, since the word-like unit **ka-'ab* in OA / CA or *hu-xt* (cp. English *your father, his sister*) is ungrammatical in Arabic.

One may therefore wonder why the representational neutralization of the POS DP - N, albeit possible in principle, in this case does not apply. The answer provided here capitalizes on the alternation C$\bar{V}$-CVC vs. CVCC characterizing the Arabic *stem* dealt with in Section 8.4 (see in particular (800)).

Consider OA / CA clitic pronoun first: in some cells of its paradigm it has the phonological representation C$\bar{V}$ (*kā, kī, hū, hā, nā*: Fleisch 1961, vol. I, p. 5 ff.) and such a form, if combined with an inalienable N, results in the bisyllabic structure C$\bar{V}$-CVC, because all the inalienable Ns under discussion are sequences CVC not only in the phonological representation, but also in the phonetical representation: *'ab, 'ax, ḥam* 'father, brother, father-in-law' (Wright, 1896, vol. I, p. 249).

For example, a more accurate characterization of (841), as far as OA / CA is concerned, is *kā-'ab*, ie:

(842) $[_{NP}$ kā $[_{N}$ 'ab]]
$[_{NP}$ C$\bar{V}$ $[_{N}$ CVC]]

But the bisyllabic structure C$\bar{V}$-CVC in the MCV (along with its variant CVCC) is typically the phonetic representation of a *a single stem*, not of a stem combined with some other constituent (cp. *kātib-* of (800)).

It follows that the phonetic representation C$\bar{V}$-CVC stemming from representational neutralization of the POS DP - N (DP a personal pronoun) threatens the correct pairing of morphs and features taking place at Spell-out, because this bysillabic structure can be ambiguously paired both with a pronominal CS and with a simple N-stem:

(843) C$\bar{V}$ + CVC vs. C$\bar{V}$CVC
Pronoun + N-stem vs. N-stem ... Pronoun?

The same holds for the MCV, where CVCC sequences like *hu-xt* occur, given that CVCC is another typical phonetic representation of a simple stem (cp. *katb-* of (800)) and then another potential source of ambiguity at Spell-out:

(844) CV + CC vs. CVCC
Pronoun + N-stem vs. N-stem ... Pronoun?

Plainly, the scenario schematized in (843 - 844) potentially makes the derivation crash and therefore the representational neutralization of the POS DP - N responsible for it is abandoned in favor of the derivational one:

(845) Representational neutralization (Word Formation) of a POS yielding at A-P a form potentially identical with an already existing form is blocked, because C-I ambiguously associates with it two different meanings. Hence, neutralization of the POS at issue is procrastinated to a later stage of derivation.

(845) can be referred to as PROCRASTINATED NEUTRALIZATION of a POS.
As a consequence, (842) transforms into (840) and then into (839) along the lines described in the Sections 8.3 and 8.4. Derivational neutralization of the POS DP - N applies via movement of the DP *-ka / -hu* to Spec, AbsP, but the very merger of this functional projection creates the POS Abs - NP (*a - 'ab / 0 - xt*), whose neutralization is derivational because of their lack of coindexing and involves movement of the NP *'ab / xt* to Spec, ConjP.
This is the derivational stage illustrated in (839) where the Conj, irrespective of whether it is covert or realized as *(a)w*, gets a demonstrative feature by virtue of its *rapport associatif* with some Amorite Conj (the MCV equivalent of *ṯumma*, in case of the null Conj, *dī* in case of *(a)w*: see Ch. 6).
Given that (crosslinguistically and in Arabic) pronouns are themselves demonstrative elements, and that Conj and DP create a POS (*aw - 'ka / 0 - hu*), coindexing between the two poles for the demonstrative feature enables Syntax to break the POS at issue representationally.
This outcomes in a word-like unit (Word Formation), which is sent at A-P for Spell-out. Since here the complex Head Conj *aw* is followed by a sequence CV(:) and they are a word, the glide *aw* surfaces as a long *ū* by virtue of the Glide Contraction Rule dealt with in Section 7.6. In greater detail the sequence CV(:) is a phonological environment blocking application of Long Vowel Shortening.
With all this in place, it is now possible to turn to the instance of pronominal CS where *alienability* of its Possessee correlates with a *short* vowel. It has been seen in Section 6.6.2 that the alienable meaning of the Possessee in both Numeral Classifier Languages and Arabic is expressed by means of a (NONRE)RC whose predicate is a Possessive Classifier combined with a non-pronominal Possessor, the so-called Possessive Classifier Construction.
Focusing on its sound-side in OA / CA, it exhibits a short Conj (*u, a, i*) because this latter is followed by the complementizer *l-* and by a noun, which in OA / CA is always CV-initial. More accurately, the phonological representation *(a)w, (a)w'i* turns into a short *u, a, i* because the phonological environment following Conj, ie *lCV* (CCV) forces application of Long Vowel Shortening after Glide Constraction (Section 7.6).
On these grounds, the vowel of the pronominal CS denoting an alienable Possessee is short because this construction too, by virtue of its meaning, is underlyingly a (NONRE)RC / Possessive Classifier Construction: actually, it is the pronominal counterpart of the non-pronominal CS denoting an alienable Possessee.
Accordingly, the short vowel of the pronominal CS denoting an alienable Possessee is modelled after the short vowel of its non-pronominal counterpart.

In a likewise fashion, the Possessive Classifier has a covert realization: no phonological environment in this kind of pronominal CS is responsible for such phenomenon, only analogy with its non-pronominal counterpart. It is worth observing in this connection that there is no need of inserting an overt complementizer *l-* in the pronominal CS denoting an alienable Possessee since it cannot be ambiguously interpreted as a possessive sentence: *clitic pronouns never occur in isolation as Predicates*.
To sum up:

(846) a. $[_{\text{ConjP}}$ ḥarr $[_{\text{Conj}'} \Downarrow$ u (=a+w) $[_{\text{CP}}$ Op $[_{\text{C}'}$ l- $[_{\text{DP}} \Downarrow 0^{DHU}$ [šams...
b. $[_{\text{ConjP}}$ 'tijārat $[_{\text{Conj}'}$ u (=a+w) $[_{\text{CP}}$ Op $[_{\text{C}'}$ pro $[_{\text{DP}} 0^{DHU}$ [ka...

Within the framework of Dynamic Antisymmetry, derivation of (846.b) parallels that of (846.a) and, more generally, of the Arabic (NONRE)RC modifier discussed in the previous Sections.
Hence, lack of lengthening of Conj is due to the fact that in (846.b) the two poles of the POS Conj - DP undergoing Abstract Incorporation and thus creating a word-like unit are *aw - Op*, not *aw - ka*, because it is the DP Op to be adjacent to Conj, not the DP *-ka*, fulfilling a more distant position in the syntactic structure.
As such, the pronominal Possessor of this kind of pronominal CS is sent to Spell-out without undergoing any process of Abstract Incorporation etc..., to the effect that when found on the sound-side immediately after Conj, their juxtaposition simply corresponds to their order of base-generation in Syntax: the word order Conj > pronominal Possessor in this case is an instance of Merger under adjacency, as much as English *(he do)es write* (835) mentioned above.
In this respect the pronominal CS denoting an alienable Possessee sharply differs from its counterpart denoting an inalienable Possessee schematized in (839), where the POS Conj - DP involves the DP *ka* since no constituent intervenes between it and Conj. This contrast is illustrated in (847).

(847) a. Inalien.: $[_{\text{ConjP}}$ 'ab $[_{\text{Conj}'}$ ū (=a+w) $[_{\text{AbsP}}$ [ka]... ⇒
b. Alien.: $[_{\text{ConjP}}$ 'tijārat $[_{\text{Conj}'}$ u (=a+w) $[_{\text{AbsP}}$ [Op ... ka]... ⇒
Word Form.
Merger under Adjac.

It follows that the alternation long vs. short vowel opposing (847.a) to (847.b) is the morphophonological reflex of a syntactic scenario of Relativized Minimality (Locality), à la Rizzi (1990) (see Section 4.5.1).
In (847.b) Op intervening between the Conj *(a)w* and the DP *-ka* blocks Abstract Incorporation of the DP *-ka* into the the Conj *(a)w* (and hence their Word Formation / Vowel Lengthening) because it is Op that undergoes Abstract Incorporation into the Conj *(a)w* by virtue of its closeness to this latter. (848) illustrates this scenario ($\square$ indicates Abstract Incorporation):

(848) a. Inalien.: [aw...ka] = [X...Z]
b. Alien.: [aw...Op]...ka = [X...Y]...Z

Dynamic Antisymmetry thus explain the correlation (851) observed by Aikhenvald (2000) as follows:

(849) The alternation long / short vowel in the Arabic pronominal CS is the morpho-phonological correlate (terminals) of the morpho-syntactic interplay between representational neutralization of a POS and Relativized Minimality (nonterminals)

In more traditional terms, this is tantamount to saying that:

(850) In the pronominal CS denoting an inalienable Possessee, the Possessee and the Possessor undergo Word Formation and hence Vowel Lenghtening because these constituents are adjacent. In the pronominal CS denoting an alienable Possessee, they do not undergo Word Formation and Vowel Lenghtening because they are not adjacent. Since this latter construction is underlyingly a (NONRE)RC, in fact, a (covert) relative pronoun intervenes between them thus blocking their Word Formation.

(851), actually a simplified representation of (848), schematizes (850) ($\boxed{\ }$ indicates Word Formation):

(851) Word Form. $\Leftrightarrow$ Inalien. Poss.ee = $\boxed{\text{Possessee + Possessor}}$
Merger under Adjac. $\Leftrightarrow$ Alien. Poss.ee = $\boxed{\text{Possessee + Rel. Pronoun}}$ + Possessor

The explanation of the correlation (851) observed by Aikhenvald (2000) is therefore recast in structural and syntactic terms, by simply combining the phenomenon of representational neutralization of a POS with Rizzi's Relativized Minimality.

8.10 Theoretical implications

This conclusive Section highlights some theoretical implications brought out by the 'case study' in the Arabic NP discussed in the previous Sections.

8.10.1 An inversion of perspective

The transformational account of the Arabic NP developed in this work capitalizes on the proposal made in Chapters 5 and 6 that the representational neutralization of a POS (Abstract Incorporation) blocks its derivational neutralization (movement).
But, as first noticed in Baker (1988), the core of Abstract Incorporation (what *blocks* movement) is coindexing, that is *pairing of two identical features*. Adopting the Checking Theory (on which, see end of Section 4.5.2), the same configuration yields the totally opposite result, since in this approach *pairing of two identical features* is responsible for *triggering* movement[13].
If on the right track, the dynamic antisymmetric account of movement developed in the present work therefore *inverts the perspective currently adopted in MP*:

(852) Pairing of features does not trigger movement, blocks it.

This shift of perspective might be plausibly have the advantage of avoiding a paradoxical situation found in the Checking Theory.
A closer examination of this latter, in fact, reveals that pairing of features both *triggers* movement and *blocks* it, given that such a theory may resort to Abstract Incorporation in order to explain some syntactic phenomena (see e.g. Kayne 1994, p. 157, fn. 24 and p. 140, fn. 10). This problem does not arise in Dynamic Antisymmetry, where pairing of features characterizes

[13]Notice that the uninterpretable / interpretable opposition of these otherwise identical features, advocated in many versions of Checking Theory (see end of Section 4.5.2) does not undermine this point, given that such an opposition is not necessary in another influential version of Checking Theory, that worked out by Rizzi (see e.g. Rizzi 1997 and references therein).

Abstract Incorporation only.
In a similar vein, the just outlined dynamic antisymmetric account *inverts* the relationship cause effect posited in Cinque's cartographic / antisymmetric account based on the Checking Theory.
In the former, it is XP-raising (remnant movement of NP) that triggers X-raising (AbsP / GenP to ConjP / WP): this scenario is summarized in (449).
In the latter it is the contrary, because Gen 'raises to a head W, thereby activating Spec,WP, which attracts the remnant NP' (Cinque, 2003, p. 68).
In this respect, the dynamic antisymmetric account seems to preferable because the Checking Theory does not explain *why* Gen (Abs) should raise to W (Conj), a phenomenon derived by Dynamic Antisymmetry invoking Domain Extending Head Movement.

8.10.2 Parameter-setting as POS neutralization

This Section explores the consequences of assuming a twofold strategy of POS neutralization, as proposed in the present work, for the so-called theory of Principles and Parameters, an influential model meant to account for language acquisition and linguistic variation (Chomsky 1995 and much subsequent work), and strictly interlocked with the Checking Theory.
In Moro (2000)'s original proposal, which dispenses with the Checking Theory, the derivational neutralization of a POS explains presence of movement: the undesirable adjacency of two poles is eliminated leaving just one pole in situ by *separating* them.
The extension of Moro (2000)'s idea put forward in this work is that the representational neutralization of a POS explains lack of movement: the undesirable adjacency of two poles is eliminated leaving just one pole in situ by *fusing* them.
In more 'traditional' versions of MP, adopting the Checking Theory, presence vs. lack of movement is taken care of by the just mentioned Principles and Parameters model. In its essence this model takes as its departure point the Checking Theory and holds that a given interpretable feature (iT) found on a substantive item (e.g. V) raises to be paired with a higher uninterpretable feature (uT) found on a functional item (e.g. I, the locus of tense and agreement). Such a pairing, in fact, would allow deletion of the uT (see end of Section 4.5.2).
In this light, presence of movement is the ability for a given substantive item (e.g. V) to delete the uT carried by a given functional item (e.g. I or, in more traditional terms, the *tense / agreement markers)*.
Focusing on V and I, descriptively natural languages display different syntactic word orders and hence different syntactic positions of V relative to I, even in one and the same language. Consider the following example, from Classical Arabic[14]:

(853) a. Perfective: katab-ta = 2SG + write.PF 'you wrote'
b. Imperfective: ta-ktub- = 2SG + write.IMPF 'you write, are writing, will write'

If one assumes in line with standard minimalist assumptions (cp. Section 4.5.3) that:

- items of Lexicon are merged into Syntax in a bottom-up fashion

[14]On the interpretation of the morphemic material attached to the verbal stem in (853) as agreement markers, see p. 3.2.2.3. The syntactic account of the perfective - imperfective contrast given here is partly based on Benmamoun (2000).

- substantive item are merged into Syntax before functional items

it follows that V is merged before than I and fulfills a lower position than I prior to displacement phenomena. In other words, no movement has applied in the following configuration:

(854) I...V

On these grounds, *ta-ktub-* (853.b) reflects lack of movement (V is lower than I), *katab-tu* (853.a) reflects presence of movement (V higher than I). Within the Principles and Parameters model the different behavior of these two verbal tenses implies that in CA the I *ta* carries an uT as well as that the CA perfective stem *katab* (853.a) carries the iT able to delete the uT of *ta*, so that it moves to the left of *ta*.
It also implies that the CA imperfective stem *ktub-* (853.b) does *not* carry the iT able to delete the uT of the I *ta*, so that *ktub-* does *not* move to the left of *ta*.
For the sake of illustration, the uT / iT at under discussion can be taken to be tense (T)[15].
The Principles and Parameters model explains this state of affairs contending that *presence vs. lack of the relevant iT on V, the ultimate source of presence vs. lack of movement of V* is a matter of how Lexicon stores its functional (= grammatical) features in a given language.
It is an empirical hypothesis in MP that Lexicon in all natural languages stores a fixed (universal) set of functional features, among them T. Such an obligatory presence of T in the Lexicon is dubbed a PRINCIPLE.
However, Lexicon varies in this or that natural language as to whether it encodes a given functional feature, e.g. T, into a given substantive item or not. This YES / NO option is referred to as PARAMETER. For example, Lexicon may decide to encode T into the subtantive item V or not: for the sake of simplicity, it could be said that such a decision making is arbitrary (and is usually referred to as PARAMETER SETTING). In particular, the parameter concerning presence vs. lack of T on V is designated in literature as V STRENGTH-PARAMETER
Accordingly, in the CA imperfective stem *ktub-* (853.b), the V Strength Parameter has the value NO, in the CA perfective *katab* (853.a) onto YES.
But once that the Lexicon sets *katab* onto the YES value of the V Strength Parameter, providing it with iT, the Lexicon enables *katab* to delete the uT carried by the I *ta* and hence to move to this latter.
On the other hand, when the Lexicon sets *ktub-* onto the NO value of the V Strength Parameter, not providing it with iT, the Lexicon disables *ktub-* to delete the uT carried by I *ta* and hence blocks movement of *ktub-* . This is schematized in (855):

(855) a. V Strenght Par.: YES ⇒ Mov.: YES ⇒ ta katab > katab ta t_{katab}
b. NO ⇒ YES ⇒ ta ktub- > blocked

From a morpho-phonological standpoint, however, the NO value of the V Strength Parameter is more complex than at first sight it can appear. While, in fact, the movement-less *taktub-* is a word, its English (quasi-)equivalent *(do)es write* is not.
Recall from Section 4.2 that the diagnostic property of a word is its uninterruptability: while Arabic *taktub-* is a word because it cannot be broken up, English *(do)es write* is not a word

[15]Recently, Chomsky proposes to identify the iT / uT responsible for movement with EPP (see e.g. Chomsky 2008), but this does not affect the main point made here, namely that in any case a pair uT / iT is at the core of the Principles and Parameters model.

because it can: cp. *(do)es not write*. This is evidenced also by the presence of the dummy support *do*, whose insertion is needed precisely because the affixal *-es* is not fused with *write*. Thus, *taktub-* and *does write* differ in that the former has undergone Word Formation, while the latter is an instance of Merger under adjacency (cp. (835) in Section 8.9).

It seems convenient at this point to bear into comparison the Principles and Parameters model with Dynamic Antisymmetry.

Functionally, the YES-parameter setting of the former corresponds to the derivational neutralization of the POS of the latter in its role of triggering movement and, in a likewise fashion, the NO-parameter setting of the former corresponds to the representational neutralization of the POS of the latter in its role of blocking movement.

A more accurate statement of this parallelism, based on the just made morpho-phonological considerations, has it that the NO-parameter setting blocks movement outcoming either in Word Formation (*taktub-*) or Merger under adjacency (*(do)es write*), and that its functional equivalents in Dynamic Antisymmetry are the representational neutralization of the POS and its non-manifestation, respectively.

Such a parallelism, summarized in Table 8.1, strongly suggests that a version of Dynamic Antisymmetry combining Moro (2000)'s proposal with the findings of the present work, might plausibly replace the Principles and Parameters model. This is a research topic that falls far beyond the scope of this dissertation, and only a piece of evidence will be brought here in favor of this claim.

Consider again the data set forth above: *katabtu, taktub-, (do)es write*. The standard assump-

Table 8.1: From Principles and Parameters to Dynamic Antisymmetry

Phenomenon	Principles and Parameters		Dynamic Antisymmetry	
Movement	Parameter	YES	Neutralization	Derivational
Word Formation	Parameter	NO	Neutralization	Representational
Merger under Adjac.	Parameter	NO	Neutralization	None

tion in literature is that the terminals they are made of, the tense / agreement markers and the verbal stem, have as their syntactic correlate the Heads I, V respectively:

(856) a. ...[$_{I'}$ katab+tu [$_{VP}$ [$_{V'}$ t_{katab}...

b. ...[$_{I'}$ ta [$_{VP}$ [$_{V'}$ ktub-...

c. ...[$_{I'}$ es [$_{VP}$ [$_{V'}$ write...

Nevertheless, Cinque (2006) recently argues that under an antisymmetric / cartographic approach the verbal stem subject to movement is a phrase VP instead of a Head V. Developing this line of reasoning, (856) is rewritten as (857):

(857) a. ...[$_{IP}$ katab [$_{I'}$ ta [$_{VP}$ t_{katab}...

b. ...[$_{I'}$ ta [$_{VP}$ ktub-...

c. ...[$_{I'}$ es [$_{VP}$ write...

In (857) the verbal stem à la Cinque (2006) is Dependent-less, given that no other constituent sits in VP. It is monosyllabic as well: this appears evident at least for the English verbal stem *write* and the CA imperfective stem *ktub-*, and this scenario comes as no surprise once that

Benveniste (1935)'s and Bohas (1997)'s research on the Indo-european and Semitic root is taken into account (see Section 8.4 and (803)).
As for the CA perfective stem *katab*-, diachronic evidence, especially from Akkadian, shows that it is a monosyllabic sequence CVCC (*katb*-) with the vowel *a (i, u)* intervening between the second and the third consonant (V2) in order to avoid the cluster CC-CV arisen from suffixation of the tense / agreement marker CV (*ta* etc...) to the verbal stem (CV)CC (Garbini & Durand, 1994, p. 112)[16]:

(858) katb > katb + ta > *CC + CV > V2-insertion > kat a bta

Only later does V2 grammaticalize in OA / CA into an apophonic vowel (Garbini & Durand 1994, p. 115[17], Guerssel & Lowenstamm 1996) but Guerssel & Lowenstamm (1996, p. 131) demonstrate that even in these varieties of Arabic the V2 *a* under certain cirucmstances cannot be analyzed as apophonic and preserves its epenthetic nature *in synchrony, that is in its phonological representation*[18].
The apparently bisyllabic perfective stem CVCVC, e.g. *katab*-, being underlyingly CVCC (*katb*-), is thus realigned with the monosyllabic stems *ktub*- and *write*.
The Dependent-less and monosyllabic stem is therefore a VP acting as a Head, so that when I is merged immediately above it into the syntactic structure, the Head I and the Head-like VP create the special POS I - VP (*ta - katab, ta - ktub, es - write*) and the different word orders that their terminals display when spelled-out is plausibly the result of some strategy of POS neutralization.
Starting with the clearest cases, the CA forms *katab-ta, ta-ktub* instantiate the derivational and representational neutralization of the POS I - VP (*ta - katab, ta - ktub*) respectively: from the standpoint of (854) displacement is evident in *katabta*, Word Formation in *taktub*-.
Turning to the most complicated case, English *(do)es write*, the pretheoretical fact that the negation *not* can intervene between *(do)es* and *write* points to its lack of word-like status, whereas the word order I > VP shows that no movement has occurred in it: hence no neutralization of the POS I - VP (*es - write*) has taken place - neither representational (Word Formation) nor derivational (movement).
As already discussed in Section 8.9, this behavior qualifies English *(do)es write* as an instance of Merger under adjacency. It has been seen throughout the present work that a potential special POS X - XP is not neutralized (in more traditional terms, it is an instance of Merger under adjacency) under the following circumstances:

- XP is not Dependent-less and / or monosyllabic: no POS (cp. (479 - 482) in Section 6.3.2)
- XP is preceded by another phrase YP closer to X: Relativized Minimality (848)
- X and XP if representationally neutralized yield a form identical with an already existing form (845)

[16] '[coniugazione] bisillabica con vocale epentetica nella seconda sillaba quando questa risulti chiusa per i triconsonantici [...]: acc[adico] *kab(i)t* «[essere] pesante, importante», *mar(u)ṣ* «[essere] malato».'

[17] 'Appare chiaro in tale schema quanto l'analogia, combinata con la morfologizzazione progressiva delle vocali d'appoggio...'

[18] 'Thus, *ḍaraab* and *kataab* with apparently identical vocalism differ, we claim, as follows: *kataab* involves genuine *a* and manifests the $a \rightarrow u$ apophonic class, whereas *ḍaraab* (underlyingly *ḍarxb*) involves x, manifesting, as such, the $x \rightarrow i$ apophonic class.[...] x, the true underlying vocalization of surface *ḍaraab*, is the null element $\oslash$.'

These findings predict that Merger under Adjacency / lack of neutralization of the POS I - VP (*es - write*) in English falls into one of the just mentioned cases.
The first is ruled out, because in the POS under discussion XP is Dependent-less and monosyllabic. So is the second, because no YP intervenes between X and YP in the CA equivalents of the English POS I - VP (see (857)).
A clue arguing in favor of the third instance of lack of neutralization is that in this case neutralization of the POS X - XP *applies at a later derivational step*: as a matter of fact, *-es has fused into a single word with do.* In other words, the English POS I - VP *(do)es - write* is to be identified with the Procrastinated Neutralization of a POS (845) and this interpretation is likely to be corroborated by the following remark: the string *(do)es + verb* minus the dummy support *do* is potentially identical with the string PL *s + verb*, as exemplified in (859):

(859) s = I of V: the boy s write (intended as: 'the boy does write')
s = PL of N: the boy s write (intended as: 'the boys write')

As a first approximation, (857) is rewritten as (860), where ⇔ indicates interference between two forms:

(860) a. ...[$_{\text{IP}}$ katab [$_{\text{I}'}$ ta [$_{\text{VP}}$ t$_{katab}$...
b. ...[$_{\text{I}'}$ ta$_i$ [$_{\text{VP}}$ ktub-$_i$... > [$_{\text{I}'}$ ta$_i$ + ktub-$_i$ [$_{\text{VP}}$ t$_{ktub-}$...
c. ...[$_{\text{I}'}$ es$_i$ [$_{\text{VP}}$ write$_i$...⇔ PL s write >

Among the details of the just outlined analysis in need of refinement left for future research is the exact nature of the coindexing between I and VP responsible for their representational neutralization in (860b), (860c).
Assuming that (860) is basically correct, the picture that emerges from the discussion so far is that *katabta, taktub, (do)es write* are to be identified, respectively, with derivational, representational, and no neutralization of a POS (later 'rescued' by the procrastinated neutralization of a POS), to the effect that their analysis lends empirical support in favor or the hypothesis diagrammed in Table 8.1, namely:

(861) Parameter-setting (Principles and Parameters) is better understood as POS Neutralization (Dynamic Antisymmetry)

Notice in particular that replacing the former with the latter has the straightforward advantage of *deriving the twofold morpho-phonological realization that crosslinguistically seems to be associated with the (alleged) NO-parameter setting, left unaccounted for in the Principles and Parameters model: Word Formation and Merger under Adjacency.*
In the best case, this latter (or, in dynamic antisymmetric terms, lack of POS neutralization) boils down to the three-fold typology developed throughout the present dissertation (no clitic-like status, Relativized Minimality, Procrastinated Neutralization of a POS).

8.10.3 A note on den Dikken's Linker

The special POS X - XP has another interesting property: when XP is an anaphoric expression, it 'pied-pipes' (moves along with) all its governing category, in order to satisfy the Principle A. In so doing, the Principle A in the dynamic antisymmetric account proposed here *performs*

the same function of pro-licensing in den Dikken (2006)'s theory of Relators and Linkers (Section 5.2.2 and especially pp. 206, 242), a device that den Dikken makes use of to derive Predicate Inversion:

(862) Derivational neutralization of the special POS X - XP, XP an Op, if combined with the Principle A replaces pro-licensing in deriving Predicate Inversion

This happens because the covert pro, according to den Dikken (2006), needs to be checked against its overt counterpart, the head (his Linker) of a projection (his FP) higher than that headed by pro (say, XP): hence pro raises to Spec, FP via phrasal movement of the entire XP, in order to be checked against the overt Linker via the Spec-Head relation.
Adopting den Dikken's approach entails that any modifier undergoing Predicate Inversion contains a pro and Ch. 5 plausibly shows that this is actually the case for secondary predication, because the analysis of crosslinguistic data carried out in that Chapter has led to set up the typology in (437 - 439), reproduced below as (863 - 865):

(863) ConjP, the syntactic manifestation of non-restrictiveness. Its modifier is a reduced RC, that is a CP headed by pro. A subtype of ConjP is the Numeral Classifier Construction.

(864) DegP, the syntactic manifestation of comparison. Its modifier is a scalar DP headed by pro

(865) DegP, the syntactic manifestation of comparison. Its modifier is a reduced RC (CP headed by pro)

This typology, left vague in den Dikken (2006), is characterized by the presence of pro in the modifier undergoing Predicate Inversion, the modifier at issue being a CP / DP, which crosslinguistically its headed by a functional item C / D of pronominal origin. As such, this typology seems to corroborate den Dikken (2006)'s analysis.
Implicit in an account along these lines, however, are two assumptions:

- the Linker is always pronominal (in order to get coindexed and then to license pro)
- the C / D is always pro

The case study in the Arabic NP made in this Chapter shows that the first assumption is falsified by the behavior of the Arabic (NONRE)RC, where the (alleged) Linker *ya- / 'i* preserves its pronominal / demonstrative nature only in an isolated and old-fashioned instance of Predicate Inversion (the Najdi Arabic transposed NP structure (814)), whereas in the more widespread instance of Predicate Inversion referred to here as Reinstated Word Order the (alleged) Linker *ya- / 'i* is *not* pronominal / demonstrative (see Section 8.6.2).
Another problem raised by this assumption is that descriptively a copula may but must not be of pronominal / demonstrative nature, a problem already pointed out in Section 5.3.5 (cp. Table 5.5).
The second assumption too is falsified by the following counterexample: the Arabic C *l-* heading the (NONRE)RC modifier undergoing Predicate Inversion is *overt*.
It follows that the Arabic data lend empirical support in favor of the proposal in (862). Remarkably, (862) is also confirmed by crosslinguistic evidence: the typology in (863), in fact, involves a CP / DP modifier with an Op sitting in Spec, CP / DP: a relative pronoun in the

case of the NONRERC (CP), a demonstrative in the case of DP.
Adopting (862) has the advantage of solving a non-trivial problem raised by an account relying upon pro-licensing, namely violation of antisymmetry (cp. Table 5.5): pro-licensing, in fact, is not able to bridge the gap between the Canonical Predication structure (possibly subject to Predicate Inversion) and its base-generation structure.
On the other side, Op as part and parcel of the special POS X - XP is involved in deriving these derivational steps. More generally the special POS X - XP is *a recursive trigger of movement deriving all the syntax of secondary predication, from its base-generation to the Reinstated Word Order*, as it will be seen in the next Section.

8.10.4 The syntax of predication as a 'chain reaction' of POS

A bird's eye view to the derivation of the Arabic NP reveals that all its derivational steps, described in literature in terms of pro-licensing or feature checking, actually stem from one and single source, namely the special POS X - XP (regardless of the ordering between X and XP):

(866) Base Generation → POS: DP - N →
Merger of AbsP (= GenP) → POS: Abs - NP →
Merger of ConjP (= WP) → POS: Conj - DP →
Merger of DefP (= FP) → POS: Def - DP →
Merger of FocP (= ZP)

Abandoning a semantically-based account of movement in favor of Dynamic Antisymmetry therefore opens the possibility that

(867) Syntax of predication, and perhaps syntactic derivation more generally, is a 'chain reaction' (recursive manifestation) of POS

8.10.5 Deriving optional movement in secondary predication

According to den Dikken (2006), pro-licensing has the ability to derive not only manifestation of Predicate Inversion - the Inverted Predication structure - but also its non-manifestation, that is the Canonical Predication structure: on his view, this latter is no more than an instance of Abstract Incorporation between pro and a higher Relator.
Considerations of Relativized Minimality (Minimal Link Condition) make it necessary to regard Op rather than pro as the constituent incorporating into the adjacent (putative) Relator, to the effect that once again Op replaces pro from an interpretive standpoint, as shown at the end of Section 5.2.1.
On the other hand, den Dikken (2006) (as well as den Dikken & Singhapreecha 2004) make no attempt to extend the pro-licensing based account to the remaining instance of secondary predication, namely the Reinstated Word Order (RWO), sticking to the Checking Theory instead. But this theoretical move fails to explain why RWO occurs in some languages and lacks in some other. This is the problem referred to as 'optionality of RWO' in Section 5.3.5 and Table 5.5.
The dynamic antisymmetric phenomenon of special POS X - XP (see previous Section), and in particular the contrast between its derivational and representational neutralization, straightforwadly accounts for this phenomenon.

Consider the Indo-european manifestation of the optionality of RWO, which opposes the English construction *that idiot of a doctor* (on which, see Appendix 5.3.6) to the following French construction:

(868) une pizza^{NEW} de chaude^{OLD}
a pizza of hot
'a hot pizza'

In literature (868) is regarded as a marked modification structure instantiating RWO because of its informational profile, which distinguishes it from its unmarked counterpart *une pizza chaude*, where *pizza* represents old information, *chaude* new one (see den Dikken 2006 and references cited there)[19].

As such, French *une pizza de chaude* is opposed to English *that idiot of a doctor* in that the former has undergone reinstatement of the word order typical of Canonical Predication (RWO: YES), the latter has not (RWO: NO):

(869) a. une pizza de chaude = RWO: YES
b. that idiot of a doctor = RWO: NO

To account for this optionality in manifesting the RWO, it seems convenient to concentrate on another opposition observable in the pair *une pizza de chaude* (869.a) - *that idiot of a doctor* (869.b) .

This opposition concerns the modifier of (869) in an early derivational stage, prior to RWO and merger of DP, namely (381), replicated below as (870):

(870) a. [$_{XP}$ chaude [$_{X'}$ de [$_{NP}$ t$_{chaude}$ pizza]]] > RWO: YES
b. [$_{XP}$ that idiot [$_{X'}$ of [$_{NP}$ t$_{thatidiotof}$ doctor]]] > RWO: NO

where *chaude* is probably an RC (CP) and *that idiot* a DP on the basis of the typology in (863). A simplified representation of (870) is the following:

(871) a. chaude de pizza > RWO: YES
b. that idiot of doctor > RWO: NO

It can be easily seen from (871) that the opposition between French *pizza de chaude* (871.a) and English *that idiot doctor* (871.b) has as its distinctive feature the lack vs. presence of a nominal constituent: the demonstrative *that* unambiguously qualifies *idiot* as a N in (871.b).

This fact is taken here not be accidental. The DP *that idiot*, in fact, carries the syntactic feature N because it contains the N *idiot*, whereas CP cannot do so, because it contains the adjective *chaude*.

In its turn, DP, by virtue of its fulfilling Spec, XP transfers the syntactic feature N onto the Head X (*of*) via the Spec - Head relation, while no such a feature can be transferred from Spec, CP to the X *de*.

As a consequence, the X *of* is coindexed for the syntactic feature N with the NP immediately right adjacent to it, whereas the X *de* is not. This is schematized in (872), where $\square$ indicates the contrast between lack (872.a) vs. presence (872.b) of coindexing:

[19]Thanks to Benjamin Fagard for helpful discussion about (868). Benjamin Fagard points out that this construction sounds to him very marked, and that adding a demonstrative to it, in order to syntactically calquing its English counterpart *that idiot of a doctor* (say, *cette pizza de chaude*) would make it quite unnatural. In this connection he interestingly remarks that a corpus linguistics-oriented research plausibly confirms unnaturalness of a construction like *cette pizza de chaude* because it does not seem to occur in a web-based corpus.

(872) a. $[_{XP}$ chaude $[_{X'}$ de $[_{NP}$ t pizza]]] > RWO: YES
b. $[_{XP}$ that idiot $[_{X'}$ $\boxed{\text{of}_i}$ $[_{NP}$ t $\boxed{\text{doctor}_i}$]]] > RWO: NO

But the NP *pizza, doctor* in (870 - 872) is parsed as monosyllabic (see Section 8.4) and is Dependent-less too (subsequent to remnant movement of *chaude / that idiot* from Spec, NP to Spec, XP), so that the special POS X - NP (*de - pizza*, *of - doctor*) arises.

While in the English manifestations of this POS the two poles are coindexed for the syntactic feature N and then a good candidate for representational neutralization, in its French counterpart they are not and then the POS at issue must be broken derivationally. As a result, the English NP *doctor* remains in situ, (finally) yielding the Inverted Predication structure *that idiot of a doctor* (Appendix 5.3.6), whereas the French NP *pizza* raises across its modifier, yielding the Reinstated Word Order *une pizza de chaude* (868).

The details of this derivation can be illustrated as follows. In French (872.a) the derivational neutralization manifests itself as raising of the *phrase pizza*, across *chaude*, with subsequent raising of the Head X (*de*) to the covert Head Y, along the lines of (449) and as illustrated in (873), where $\square$ indicates remnant or phrasal movement:

(873) $[_{XP}$ chaude $[_{X'}$ de $[_{NP}$ t pizza]]] $\rightarrow$
$[_{YP}$ $\boxed{[\text{t pizza}]_x}$ $[_{Y}$ de + t $[_{XP}$ chaude $[_{X'}$ t_{de} $[_{NP}$ $\boxed{t_x}$]]]]]

As for English (872.b), the antisymmetric account à la Kayne (1998) developed in Appendix 5.3.6 strongly suggests that it is followed by the derivational steps (429 - 435), repeated in the following as (874 - 878)

(874) = 429
... $[_{YP}$ $\boxed{[_{XP}$ that idiot $[_{X'}$ of $[_{NP}$ ...doctor...]]]} $[_{Y'}$ Y t...

(875) = 431
... $[_{ZP}$ $[_{Z'}$ Z $[_{YP}$ $\boxed{[_{XP}$ that idiot $[_{X'}$ of $[_{NP}$ ...doctor...]]]} $[_{Y'}$ Y ...

(876) = 432
... $[_{ZP}$ $\boxed{[_{NP}$ ...doctor...]} $[_{Z'}$ Z $[_{YP}$ $\boxed{[_{XP}$ that idiot $[_{X'}$ of $t_{NP}]]}$ $[_{Y'}$ Y ...

(877) = 434
... $[_{DP}$ $[_{D'}$ a $[_{ZP}$ $\boxed{[_{NP}$ ...doctor...]} $[_{Z'}$ Z $[_{YP}$ $\boxed{[_{XP}$ that idiot $[_{X'}$ of $t_{NP}]]}$ $[_{Y'}$ Y ...

(878) = 435
... $[_{DP}$ $\boxed{[_{XP}$ that idiot $[_{X'}$ of $t_{NP}]]}$ $[_{D'}$ a $[_{ZP}$ $\boxed{[_{NP}$ ...doctor...]} $[_{Z'}$ Z $[_{YP}$ t_{XP} $[_{Y'}$ Y ...

The just outlined derivation implies that the pole NP (*doctor*) has the ability to move, ie to be derivationally neutralized, and that hence coindexing is not enough for triggering the expected representational neutralization of the POS X - NP (*of - doctor*). In other words, (872.b) is an instance of Procrastinated Neutralization of a POS (845), so that representational neutralization of the POS at issue might be plausibly blocked by some interference on the level of form, whatever its exact nature.

This makes it possible to conclude that:

(879) Apparent optionality in instantiating or not the Reinstated Word Order actually boils down to a contrast between derivational and (procrastinated neutralization of a POS, itself a subtype of) representational neutralization

8.11 Summary

This Chapter has provided a dynamic antisymmetric account of the Arabic NP based on the findings of the previous chapters, motivated on independent semantic and phonological grounds. These findings, if couched within an antisymmetric framework, bring out a hierarchy of functional projections highly reminiscent of that posited in previous antisymmetric / cartographic accounts of the Arabic NP (e.g. Cinque 2003, Shlonsky 2004, Ihsane 2003), at least insofar as the left and the right peripheries are concerned:

(880) DP > FocP > DefP ... WP > GenP > NP

Thus, the Agglutination Hypothesis discussed in the previous Chapter corroborates the hypothesis that the Arabic NP has the following functional layer:

(881) DP > ZP > FP ... ConjP > AbsP > NP

But investigation of the semantic and syntactic properties of (880 - 881) results in identifying the (apparently) language-specific hierarchy in (881) with the universal hierarchy of functional projections in (880). Nonetheless, the dynamic antisymmetric account proposed in this Chapter has departed from the standard antisymmetric / cartographic approach in that the former totally dispenses with the Checking Theory of movement adopted in the latter. Instead, neutralization of a POS is invoked as the real trigger of movement.
In consequence of this, the entire derivational history of the Arabic NP turns out to be a chain reaction of POS:

(882) Base Generation → POS: DP - N →
Merger of AbsP (= GenP) → POS: Abs - NP →
Merger of ConjP (= WP) → POS: Conj - DP →
Merger of DefP (= FP) → POS: Def - DP →
Merger of FocP (= ZP)

Such an approach is able to replace the Checking Theory in all of its facets: it explains not only why and how movement takes place (derivational neutralization or movement), but also how and why it does not (representational neutralization or Abstract Incorporation). In other words, derivational neutralization of a POS invoked by Dynamic Antisymmetry might plausibly take over the function of the parameter having the value YES in the Principles and Parameters model, representational neutralization of a POS that of the parameter having the value NO. A brief case study of Arabic and English verb is carried out in order to corroborate this hypothesis.
Returning to the Arabic NP, the same contrast between derivational and representational neutralization of a POS is able to account for the dialectal opposition between the Hijazi and the Bedouin branch in terms of lack vs. presence of Predicate Inversion (and hence, indirectly, for the opposition Subject vs. non-Subject arising from it in the CA 'case-system').
The ultimate source of this opposition is likely to be the ability for the Amorite RC type to representationally neutralize its POS 0_{Conj} - $\underline{d\bar{\imath}}$ in the Hijazi branch vs. its inability to do so in the Bedouin branch, so that in this latter derivational neutralization of the same POS, ie Predicate Inversion (ending up with the Reinstated Word Order) is invoked instead.
This happens because in the Hijazi branch, the pole $\underline{d\bar{\imath}}$ of the Amorite POS 0_{Conj} - $\underline{d\bar{\imath}}$ (Conj -

DP) is of demonstrative origin and therefore gets coindexed for the Demonstrative feature with Conj, marked for the same feature (see Ch. 6). On the other hand, in the Bedouin branch, the pole *d̲ī* involved in the same POS has lost its demonstrative function, as clearly shown by the phenomenon of desemantized (otiose) *d̲-* in it, and therefore it cannot get coindexed with Conj, nor is it able to be neutralized representationally.

Turning to the Pre-semitic RC type, it neutralizes its POS *w - Op* (Conj - DP) by analogy with its Amorite counterpart, to the effect that in the Hijazi branch *w - Op* is representationally neutralized as much as 0_{Conj} - *d̲ī* is, and by the same token in the Bedouin branch derivational neutralization (Predicate Inversion and Reinstated word Order) of *w - Op* applies, since this latter mimicks the POS 0_{Conj} - *d̲ī* characterized by the desemantized demonstrative:

(883) a. Hij. ...$[_{\text{Conj}'}$ $\boxed{0_i}$ $[_{\text{CP}}$ $\boxed{\text{d̲ī}_i}$... ⇔ ...$[_{\text{Conj}'}$ $\boxed{\text{w}_i}$ $[_{\text{CP}}$ $\boxed{\text{Op}_i}$... > Abstract Inc.
b. Bed. desemantized (otiose) *d̲* ⇔ ...$[_{\text{Conj}'}$ $\boxed{\text{w}}$ $[_{\text{CP}}$ Op... > Pred. Inv.

The same opposition between derivation and representation in neutralizing a POS derives the difference between Predicate Inversion and Reinstated Word Order in a later stage of derivation, namely when the copula of Predicate Inversion (Def) is merged above ConjP (a contrast left unexplained in den Dikken 2006's theory of Relator and Linkers).

The very merger of Def, in fact, is responsible for the emergence of the POS Def - NP, neutralized representationally if Def and NP are coindexed for the syntactic feature N (which keeps the Inverted Predication structure unaltered), or derivationally if they are not (which yields the Reinstated Word Order).

The Inverted Predication structure survives as a relic in Najdi Arabic in the so-called transposed NP structure, the Reinstated Word Order in the modification structure displaying an intervening *a, i* between the Noun and its (RC) modifier, which is the standard modification structure in Najdi Arabic and the modification structure markd for ACC, OBL in CA. The structural (and semantic-less) link between the transposed NP structure and the Reinstated Word order is set in descriptive work by Ingham (1994).

The English construction *that DP of DP* and the French construction *une pizza de chaude* have been accounted for in the same way (representational vs. derivational neutralization of a POS structurally highly reminiscent of Def - NP).

Finally a dynamic antisymmetric account, if combined with Relativized Minimality, is also able to translate in formal terms the contrast long vs. short 'case-ending' (copula) in the pronominal CS. Such an account capitalizes on a more traditional explanation along the following lines. In the pronominal CS denoting an inalienable Possessee, the Possessee and the Possessor undergo Word Formation and hence Vowel Lenghtening because these constituents are adjacent. In the pronominal CS denoting an alienable Possessee, they do not undergo Word Formation and Vowel Lenghtening because they are not adjacent. Since this latter construction is underlyingly a (NONRE)RC, in fact, a (covert) relative pronoun intervenes between them thus blocking their Word Formation.

Chapter 9

Conclusions

9.1 Main findings and results

9.1.1 Arabic studies

Two main generalizations are involved in the description of the findings below:

- Arabic belongs to the agglutinative type
- The phonological constraint *d;l* (ordering irrelevant) operates in it

(I) The Amorite RC marker *llaḏī* is a Conjunction signalling a NONRERC, to be likened to English *and* in *Enrico, and he is the smartest of us all, got the answer in seven seconds*. This kind of item typically acts as a copula of Canonical Predication

(II) This Conjunction signalling a NONRERC develops out of a demonstrative pronoun acting as a relative pronoun via reanalysis (ie the pronoun *ḏī* fuses with a null Conjunction signalling a NONRERC that immediately precedes it)

(III) *llaḏī* can be defined as a copula of Canonical Predication and as different from a demonstrative pronoun (hence, as a Conjunction signalling a NONRERC) on the basis of two diagnostic properties: *la*-prefixation and resistance to case-marking (cp. the separating pronoun *huwa*: it does not have a demonstrative meaning and it acts as a copula)

(IV) *llaḏī* later evolves into a complementizer because of dropping of *ḏī*: this results in the phonological resemblance of the remaining string *elli* or the alike with the pre-Semitic complementizer *l*- and hence the former is reinterpreted along the lines of the latter via (phonological) analogy

(V) The putative case-ending NOM *u / w* is to be identified with the Conjunction *u / w* 'and', widespread all across Semitic. Its ability to act as a Conjunction signalling a NONRERC in Arabic is paralleled by the Somali RC Conjunction *oo* 'and', able to perform the same function. This state of affairs qualifies the putative case-ending NOM *u / w* as a Pre-semitic RC marker

(VI) The parallelism between Arabic and Somali NONRERC is confirmed by the following diagnostic property: its ability to get both a NONRERC and an adverbial reading

(VII) More accurately, the phonological representation of the putative case ending *u / w* is *aw*, an agglutinative complex made of a fossilized ABS marking the Predicate / Modifier and of the Conjunction *u / w*. A similar instance of ABS, in fact, is also found in the Somali NONRERC

(VIII) As for the putative case-endings *a, i* (ACC, OBL), they are still Conjunctions signalling a NONRERC, because they are commutable with *u* in OA, but they are copulae of Predicate Inversion rather than of Canonical Predication because they are also found in contexts of Predicate Inversion (e.g. Najdi Arabic *mbaccir a-l wasim* = RC + N = Inverted Predication structure vs. CA *ṣāliḥ-u l 'axlāq* = N + N = CS, *not* Inverted Predication Structure)

(IX) Hence, the construction N + *a / i* + RC, e.g. *bāb-*$\boxed{a}$ *l-ṣaġīr* is an instance of Reinstated Word Order, to be likened to French *une pizza* $\boxed{de}$ *chaude*, as opposed to the Inverted Predication structure *mbaccir* $\boxed{a}$*-l wasim / un imbécile* $\boxed{de}$ *garçon* and to the Canonical Predication structure *bāb-*$\boxed{u}$ *l-ṣaġīr / une pizza* $\boxed{0}$ *chaude*

(X) Capitalizing on Arab Grammarians' analysis of the same morphemic material in the same phonological environment, the phonological representations of the putative case-endings *a, i* turn out to be *aw'i, u'i* and they are agglutinative complexes made of the relic ABS marking the Predicate / Modifier and realized either as *a* or zero (contrast ACC *aw'i*, with OBL *u'i*), the Semitic Conjunction *u / w* signalling the NONRERC and a morpheme *'i*.

(XI) In other words, the phonological representation of ACC, OBL *a, i* reveals that they are no more than NOM *u* plus a morpheme *'i* (on which, see immediately below)

(XII) The morpheme *'i* is documented in some MCV as a variant of the demonstrative base *ya-*, which in Najdi Arabic acts as a copula of Predicate Inversion (Predicate + *(iy)ya* + Subject) and also occurs in the Reinstated Word Order (Noun / Subject + *ya-lli* + RC / Predicate). Hence, the real *locus* of Predicate Inversion in the putative case endings ACC, OBL, *a, i* is the underlying item *'i*

(XIII) From a morpho-phonological standpoint, the alternation *'i / ya-* observed in the copula of Predicate Inversion is consistent with the Pre-semitic and Amorite RC types in which they are found respectively: cp. the alternation *l- / la* observed in the Semitistic literature (Testen, 1998).

(XIV) The CA putative case-endings (*bāb-*$\boxed{a}$ *l-ṣaġīr*) develop out of copulae (of Canonical Predication and Predicate Inversion, e.g. *(al-)bēt* $\boxed{a}$*-l-ṭuwīl*) as follows.
CA case-endings *u(:), a(:), i(:)* are better understood as markers of an opposition Subject *u(:)* vs. non-Subject *a(:), i(:)*. This semantic opposition is originally a dialectal opposition between the Hijazi *u(:)* and the Bedouin *a(:), i(:)*, where these items are copulae intervening between a noun and its RC modifier. Development of Hijazi *u:* into a Subject role and of Bedouin *a(:), i(:)* into a non-Subject role is a consequence of their informational profile. While the Hijazi *u:* does not occur in a modification structure having undergone Predicate Inversion, the Bedouin *a(:), i(:)* do. But a modification structure thus characterized is less suitable to function as a Subject because of the newness of information encoded on its noun: crosslinguistically the Subject represents old information.

Thus, when the Hijazi and the Bedouin branch fuse into a single language (CA), the modification structure containing *u:* is preferred for the Subject role over that containing *a(:), i(:)*, which accordingly polarize into the non-Subject role

(XV) Probably, the semantic role of non-Subject further specializes into ACC, OBL via polarization of the agglutinative complex *aw'i, u'i*, where the overt vs. covert realization of the relic ABS (also attested in Hamitic) is morphologized and assigned a new contrastive value: non-Subject Object vs. non-Subject other than Object. Cp. also (XXV)

(XVI) Ability of *l-* to introduce the (NONRE)RC signalled by the Conjunction copula *u / w* in OA qualifies it as a complementizer introducing a (NONRE)RC ('medial' or 'adjectival' *l-*)

(XVII) The pretheoretical fact that Arabic 'adjectives' in OA and MCV (the most conservative manifestations of Arabic) are introduced by the same functional items (*l-, elli*) introducing the (NONRE)RCs forces to reinterpret Arabic 'adjectives' as (NONRE)RCs (substitution test)

(XVIII) As for the Arabic *prenominal l-* ('article'), it arises as a disambiguating element intervening between the demonstrative *d̲-* and the noun

(XIX) Its NP-initial, Indo-european-like position is a deceiving effect of a phonological phenomenon of dropping of the demonstrative *d̲-* because of the non-euphonic cluster *d̲-l*

(XX) Its (originally) medial position (between the demonstrative and the noun) as well as its ability to act as a complementizer introducing a (NONRE)RC qualify it as a Determiner

(XXI) The *tanwīn -n* is an Animate Numeral Classifier turned into a General Classifier

(XXII) As Numeral Classifier, *-n* tends to be combined with a covert quantifier 'one' in the SG and with the quantifier *(')a* 'few' in the (SOUND) PL, a behavior widely attested crosslinguistically: the numeral 'one' is deleted because of its being the most unmarked member of the series of numerals (Greenberg, 1990a)

(XXIII) Hence, the demarcative function of the (apparently) word-final *-n* is a deceiving effect of the unmarked nature of 'one', whereas the (apparently) monomorphemic nature of *-na* is a deceiving effect of a rule of *hamzah*-Deletion turning the agglutinative complex Classifier plus quantifier *n + 'a* into a string *-na*

(XXIV) The Arabic noun marked for what is traditionally referred to as SG *-n* and SOUND PL *-na*, being a Numeral Classifier Construction, is underlyingly a NONRERC, in the wake of Greenberg (1990a)

(XXV) Expression of PL number further specializes into DU (= PL lesser than 3) and SOUND PL (= PL more than 3) via polarization of the agglutinative complex *aw'i, u'i* preceding *-na* (ie the complex Classifier plus quantifier) and surfacing as *ā, ī*, respectively. Thus, the overt vs. covert realization of the relic ABS is morphologized and assigned a new contrastive value: (SOUND) PL vs DU. Cp. also (XV)

(XXVI) The DU-ending *ā-na* (attested in OA) further differentiates from its SOUND PL counterpart *ī-na* because of a Low Vowel Raising rule, turning *ā-na* into *ā-ni* (and, sometimes, *ī-na* into *ī-ni*). This comes as no surprise, since this rule is widely attested in the Bedouin branch (Abboud, 1979) the copulae *a, i* belong to (see (XIV))

(XXVII) In its turn, Low Vowel Raising sets the condition for application of the *i*-Deletion rule, affecting a word-final *i*: *ā-ni, ī-ni* turn into *ā-n, ī-n*. Remarkably, some instances of *ā-n* in the MCV still function as (SOUND) PL: the so-called PSEUDO-DU

(XXVIII) As for BROKEN PL, a comparison between its manifestation in the Bedouin branch, where it conveys a paucal meaning and ends in *-in*, and the SOUND PL plausibly shows that difference between them boils down to a matter of phonological variation: (PAUCAL) BROKEN PL is no more than a (SOUND PL ending in *ī-n* (see (XIV)) affected by a rule of Long Vowel Shortening observable in OA and CA (*lam yaṣīr* > *lam yaṣir*).

(XXIX) Long Vowel Shortening applies for two reasons. Either as a consequence of the deletion of the quantifier *'a* via *hamzah*-Deletion, Low Vowel Raising (XXIII), (XXVI) and *i*-Deletion, which convert *ī + n + 'a* into *īn*, the target of this rule (cp. *nāsin*), or because of fronting of the quantifier at issue, a behavior typical of quantifiers combined with a General Classifier (Greenberg, 1990a): e.g. *'a*-*ḥruf-in*. In this latter case, in fact, fronting of the quantifier 'leaves in situ', so to speak, an open syllable, the target of Long Vowel Shortening

(XXX) The very shortening of the vowel - actually a copula - preceding the (General) Classifier *-n* resyllabifies the entire word. While the string: stem-final C + *ī + n + 'a* is bysillabic (*Ci:* - *na*), the string stem-final C + *in* is monosyllabic (*Cin*). In all likelihood, resyllabification (or more precisely, Long Vowel Shortening) thus triggers all the phenomena traditionally ascribed to BROKEN PL: C2-lengthening, gemination, affixation etc.... Partial evidence for this claim has been brought here through the study of the phenomenon of C2-lengthening and a parallel has been set between Arabic BROKEN PL and Germanic apophonic PL (*nam(a)l-in* > *nimāl-in* > *nimāl* ≈ *foot-i* > *feet-i* > *feet*)

(XXXI) As General Classifier, *-n* tends to take over the function of all the other Classifiers, a behavior widely attested crosslinguistically

(XXXII) The bound *-t* is an Inanimate Numeral Classifier. As such its evolution into a feminine marker and superimposing of *-n* on it (cp. Arab Grammarians' *muqābalah* and the Triptotic *'asadun* supplanting the Diptotic *'usamat-*) is a consequence of the tendency of the General Classifier *-n* to replace the Inanimate Classifier *-t*

(XXXIII) In this light, the thematic (ie morphemic) *a* preceding the bound *-t* is to be identified with the copular *a* preceding *-n*, the (NONRE)RC modifier etc...

(XXXIV) MCV evidence that the non-pronominal FS is in reality a Possessive Classifier Construction, where a noun denoting property, wealth (e.g. *mta', ḏū, ḥagg*) etc... grammaticalizes into a Possessive Classifier mediating the relationship between an *alienable* Possessee and its Possessor.

(XXXV) The ability for the non-pronominal FS to be coordinated with the Numeral Classifier Construction, itself an (underlying) NONRERC, qualifies the non-pronominal FS denoting an *alienable* Possessee in the same way (coordination test)

(XXXVI) The non-pronominal CS displaying a medial *l-* is no more than a non-pronominal FS (Possessive Classifier Construction) in which the Possessive Classifier *ḏū* is deleted because of the non-euphonic cluster *l-ḏ*. The Possessive Classifier *ḏū* in fact is immediately preceded by the complementizer *l-* introducing the (NONRE)RC

(XXXVII) The *l*-less non-pronominal CS traditionally referred to as *tamyīz* is no more than a non-pronominal FS (Possessive Classifier Construction) in which the General Classifier *-n* takes over the function of the Possessive Classifier *ḏū*, a behavior attested crosslinguistically

(XXXVIII) The *l*-less non-pronominal CS traditionally referred to as indefinite CS is no more than a non-pronominal FS (Possessive Classifier Construction) in which the General Classifier *-n* takes over the function of the Possessive Classifier *ḏū* and is subsequently deleted because of its semantic weakening, a behavior attested crosslinguistically

(XXXIX) In the pronominal CS denoting an *inalienable* Possessee, the Possessee and the Possessor undergo Word Formation and hence Vowel Lenghtening because these constituents are adjacent (e.g. *'abū-ka*). In the pronominal CS denoting a (potentially) *alienable* Possessee, they do not undergo Word Formation and Vowel Lenghtening because they are not adjacent (e.g. *tijāratu-ka*). This latter construction, in fact, because of its denoting a (potentially) *alienable* Possessee is underlyingly a (NONRE)RC (cp. (XXXV)) and then a (covert) relative pronoun intervenes between the Possessee and the Possessor thus blocking their Word Formation

(XL) On the whole, all the modification structures found in Arabic (in a broad sense, which includes *-n, -t*) are NONRERC

(XLI) The only notable exception is the *pronominal* CS denoting an *inalienable* Possessee, where the Pronominal Possessor is juxtaposed *tout court* to its Possessee, without being embedded into a NONRERC

(XLII) On the whole, manifestation of Classifier is a pervasive phenomenon in Arabic (Numeral Classifier Constructions and Possessive Classifier Constructions), a state of affairs which to a good extent characterizes it as a Numeral Classifier Language (a position already hinted at in Greenberg 1990a and Aikhenvald 2000)

(XLIII) In this light, the threefold paradigm of number marking posited by Arab Grammarians, having as its input form a COLL able to function both as SG and PL, is straightforwardly accounted for, since this is the conceptualization of number marking typically found in Numeral Classifier Languages (see Section 9.2.3 below)

(XLIV) All Arabic adnominal markers, Classifiers included (after a process of grammaticalization), act as copula (see Section 9.2.1 below).

9.1.2 Minimalist Program

Consider Moro (2000)'s Dynamic Antisymmetry first: Given that Syntax is 'atemporal', two adjacent items in it cannot be located on the time axis and hence are unpronounceable. This configuration is designated as Point of Symmetry. To allow their pronunciation, their undesirable adjacency is eliminated leaving just one item in situ by *separating* them. This is why movement applies. In other words, movement is *derivational* neutralization of a POS.
Developments of Moro's theory put forward in this work are as follows:

(I) If two adjacent items are coindexed, their undesirable adjacency (inhibiting their pronunciation) is eliminated leaving just one item in situ by *fusing* them. This is why Word Formation applies. In other words, Word Formation is *representational* neutralization of a POS.

(II) An Head-Head construction, either resulting from head-movement or from Abstract Incorporation, creates a POS that is not neutralized derivationally (via movement) but *representationally (via word-formation).* Representational neutralization of the POS makes the Head - Head construction legitimate in Syntax

(III) An Head - Head construction / POS resulting from Abstract Incorporation blocks movement of the poles it is made of because it neutralizes its POS representationally via Word Formation. Hence movement, which performs the same task derivationally, is not needed

(IV) An Head - Head construction / POS resulting from Head-movement blocks further movement of the poles it is made of because it neutralizes its POS representationally via Word Formation. Hence further head-movement, which performs the same task derivationally, is not needed

(V) X cannot raise as much as a XP does in order to neutralize the special POS X - XP, because of Economy Principle. After that XP has been selected for movement, X is too because of Last Resort. XP cannot raise to its landing site because of an intervening YP (Minimal Link Condition) and only movement of X makes XP and YP equidistant to the landing site, thus allowing XP to raise to it (Domain Extending Head Movement)

(VI) In the derivational neutralization of a POS, the copula is a Head X merged into the syntactic structure to provide a landing site (Spec, XP) with the raised pole of a POS and hence to make the derivation converge at A-P. If the landing site is already fulfilled by another constituent, the raised pole of the POS is forced to move to the next available Spec via Domain Extending Movement. But if the copula X and its already occupied landing site Spec, XP are the edge of a phase (CP / DP), no next available Spec can be made available in order to raise the pole and to derivationally neutralize the POS. Accordingly, Merge applies, which injects new syntactic structure in the edge itself, ie immediately above Compl, CP / DP

(VII) Representational neutralization (Word Formation) of a POS yielding at A-P a form potentially identical with an already existing form is blocked, because C-I ambiguously associates it with two different meanings. Hence, neutralization of the POS at issue is procrastinated to a later stage of derivation. This phenomenon can be designated as Procrastinated Neutralization of a POS

(VIII) Parameter-setting (Principles and Parameters) is better understood as POS Neutralization (Dynamic Antisymmetry). In the best case, derivational neutralization of a POS (Dynamic Antisymmetry) might be plausibly replace the parameter having the value YES (Principles and Parameters), representational neutralization of a POS the parameter having the value NO

(IX) Derivational neutralization of the special POS X - XP, XP an Op, if combined with the Principle A replaces pro-licensing in deriving Predicate Inversion

(X) Syntax of predication, and perhaps syntactic derivation more generally, is a 'chain reaction' (recursive manifestation) of POS

(XI) Apparent optionality in instantiating or not the Reinstated Word Order actually boils down to a contrast between derivational and (procrastinated neutralization of a POS, itself a subtype of) representational neutralization

(XII) All the modification structures found in Arabic (in a broad sense, which considers *-n, -t* as modifiers) are NONRERCs (ie RCs signalled by Conj) (cp. (XL)):

(884) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ Conj $[_{\text{CP}}$ Op $[_{\text{C}'}$ C...

(XIII) The only notable exception is the pronominal CS denoting an inalienable Possessee, where the Pronominal Possessor is juxtaposed *tout court* to its Possessee, without being embedded into a NONRERC (cp. (XLI)):

(885) ... $[_{\text{ConjP}}$ NP $[_{\text{Conj}'}$ Conj $[_{\text{DP}}$ Pronoun...

(XIV) The alternation long / short vowel in the Arabic pronominal CS is the morpho-phonological correlate (terminals) of the morpho-syntactic interplay between representational neutralization of a POS and Relativized Minimality (nonterminals) (cp. (XXXIX))

(886) a. Inalienable Poss.ee: $\boxed{\text{Conj...DP}} = \boxed{\text{X...Z}}$
b. Alienable Poss.ee: $\boxed{\text{Conj...Op}}\text{...DP} = \boxed{\text{X...Y}}\text{...Z}$

(XV) In the Arabic modification structure Spec-to-Spec movement is barred and XP-raising applies instead for reasons of convergence with C-I: the need to avoid ambiguity in the scope of the Op found in the Arabic (NONRE)RC modifier. This results in the ordering of N and of its modifier which is the mirror-image of English

9.1.3 Linguistic theory

(I) What is traditionally called 'word' is (recursive) 'packaging' of two items that cannot be located on the time axis and hence are unpronounceable. Their 'packaging' enables them to be located on the time axis and to be pronounced. Technically, (recursive) 'packaging' is representational neutralization of a POS

(II) Adopting a more familiar notion, it is as if the item A and the item B are two summands (say, 2 and 5). Before addition is performed, summands are possessed of the property of *commutativity* - in simpler words, summands are *unordered*: one can arrange them both in

a sum left-to-right, and right-to-left, with no change in the final result (2 + 5 = 7, 5 + 2 = 7). On the other hand, the word C resulting from their Word Formation can be likened to the sum of summands. After addition is performed, the sum of summands by definition is not possessed of commutativity (it is just one number: say, 7) but preserves the numerical value of its summands (2 + 5) and in fact their equality can be asserted by an equation (e.g. 2 + 5 = 7 etc...). By the same token, A + B = C insofar as the phonological content is concerned. It ensues that in both the sum of summands and in the word, information is preserved but lack of ordering is no longer found and, furthermore, both the sum of summands and word are recursive entities. Focusing on the word, eliminating lack of ordering in it allows its spell-out

(III) the POS Head - Head construction featuring two overt poles and neutralized via Word Formation is what is traditionally called agglutinative type

(IV) the POS Head - Head construction featuring a covert and an overt pole and neutralized via Word Formation is what is traditionally called fusive type

(V) More temptatively, the POS Head - Head construction featuring two overt poles, typical of agglutination, if for some unknown reason is not neutralized via Word Formation is what is traditionally called isolating type

(VI) The syntactic structure of the Arabic possessive sentence is not: Possessor + Preposition-Pronoun + Possessee^{Indefinite}. Rather, it is to be interpreted as: Possessor + Preposition-Copula + Possessee^{Indefinite}.

9.2 Synopsis

9.2.1 Copulae in Arabic: a taxonomy

Adnominal marker		Predication Structure			Cp.
Funct. item	Cat. status	Copula?	Subject	Predicate	
l-	D	YES	Dem. (Op)	NP	afto <u>to</u> vivlio
	C	YES	Rel. PRN (Op)	IP	the man <u>that</u> arrived yesterday
(lla)d̲ī and var.	DP	NO	-	-	the man <u>who</u> arrived yesterday
	Conj	YES	NP	(NONRE)RC	Enrico, <u>and</u> he is the smartest...
	C	YES	Rel. PRN (Op)	IP	the man <u>that</u> arrived yesterday
u = aw	Conj	YES	NP	(NONRE)RC	Cali <u>oo</u> Maryam la hadlayá
a = (aw)'i	Def (Dem.-like)	YES	(NONRE)RC	NP	un imbécile <u>de</u> garçon
			NP	(NONRE)RC	une pizza <u>de</u> chaude
i = (w)'i	Def (Dem.-like)	YES	NP	(NONRE)RC	une pizza <u>de</u> chaude
ya-	Def (Dem.-like)	YES	NP	(NONRE)RC	une pizza <u>de</u> chaude
-n	CLF (DP-like)	NO	-	-	ome <u>war</u> po
	C	YES	Rel. PRN (Op)	IP	the man <u>that</u> arrived yesterday
-t	CLF (DP-like)	NO	-	-	ome <u>war</u> po

9.2.2 Copulae in Arabic: the 'Russian-doll' effect

(887) [0_{DEM} ⇐ l ⇒ [bāb ⇐ u ⇒ [0_{PRN} ⇐ l ⇒ ṣaġīr]]]

a. [*Subject:* 0_{DEM} ⇐ l ⇒ [$_{\text{Noun Phrase}}$ *Predicate:* bāb-u 0_{PRN} l-ṣaġīr]]

b. [*Subject:* bāb ⇐ u ⇒ [$_{\text{Relative Clause}}$ *Predicate:* 0_{PRN} l-ṣaġīr]]

c. [*Subject:* 0_{PRN} ⇐ l ⇒ [$_{\text{Noun Phrase}}$ *Predicate:* ṣaġīr]]

9.2.3 Classifiers in Arabic

(888) Old / Classical Arabic (= 160)

baqar	/	baqar-at	WĀḤID /	baqar-āt ('a)
cow.COLL		cow-CLF	Q	cow-CLF Q

(889) Old / Classical Arabic (= 161)

šajar	/	šajar-at	WĀḤID	/	('a) šjār-	un
tree.COLL		tree-CLF	Q		Q tree	CLF

(890) Old Arabic (Lane, 1863, vol. I, p. 295)

taba'	/	taba'-	un	WĀḤID	/	tābi'	ūn ('a)
companion.COLL		companion	CLF	Q		companion	CLF Q

9.2.4 The 'chain reaction' of POS

(891) Base Generation → POS: DP - N →
Merger of AbsP (= GenP) → POS: Abs - NP →
Merger of ConjP (= WP) → POS: Conj - DP →
Merger of DefP (= FP) → POS: Def - DP →
Merger of FocP (= ZP)

References

SECONDARY SOURCES

Abboud, P. F. (1979). The Verb in Northern Najdi Arabic. *Bulletin of the School of Oriental and African Studies*, 3, 467-499.

Abboud-Haggar, S. (2003). *Introducción a la dialectología de la lengua árabe*. Granada: Fundación El Legado Andalusí.

Abboud-Haggar, S. (2006). Dialects: Genesis. In K. Versteegh, M. Eid, A. Elgibali, M.Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 613-622). Leiden: Brill.

Abney, S. P. (1987). *The English Noun Phrase in its Sentential Aspect*. PhD thesis, MIT.

Abu Chacra, F. (2007). *Arabic: An Essential Grammar*. London: Routledge.

Acquaviva, P. (2008). *Lexical Plurals: A Morphosemantic Approach*. Oxford: Oxford University Press.

Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

Aikhenvald, A. Y. (2003). *A Grammar of Tariana*. Cambridge: Cambridge University Press.

Al-Qahtani, D. M. (2000). The Nature of Arabic Sonorant Consonants. *Langues et Linguistique*, 6, 1-21.

Al-Sharkawi, M. (2006). Pre-Islamic Arabic. In M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 689-699). Leiden: Brill.

Alexiadou, A., Haegeman, L., & Stavrou, M. (2007). *Noun Phrase in The Generative Perspective*. Berlin: Mouton de Gruyter.

Alhawary, M. T. (2003). Elicitation Techniques and Considerations in Data Collection in Early Arabic Grammatical Tradition. *Journal of Arabic Linguistics Tradition*, 1, 1-24.

Ali, Y. (1934). *The Holy Quran: Translation and Commentary*. Lahore: Ahmadiyya Anjumani-Ishaat-i-Islam.

Anghelescu, N. (2004). *La langue arabe dans une perspective typologique*. Bucharest: Bucharest University Press.

Angoujard, J. P. (1990). *Metrical Structure of Arabic*. Dordrecht: Foris.

Arberry, A. J. (1955). *The Koran Interpreted: A Translation*. London: Allen & Unwin.

Aronoff, M. (1998). Isomorphism and Monotonicity: Or the Disease Model of Morphology. In

D. K. Brentari *et al.* (Eds.), *Morphology and Its Relation to Phonology and Syntax* (pp. 411–418). Stanford: CSLI Publications.

Ayoub, G. (1991). La nominalité du nom ou la question du *tanwin*. *Arabica*, 38, 151–213.

Ayoub, G. (2006). Tanwin. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski

(Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume IV (pp. 442-445). Leiden: Brill.

Baalbaki, R. (2004). *Grammarians and Grammatical Theory in the Medieval Arabic Tradition*. Farnham: Ashgate Publishing.

Baalbaki, R. (2006). Aṣl. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 191-195). Leiden: Brill.

Baalbaki, R. (2008). *The Legacy of the Kitāb: Sibawayhi's Analytical Methods Within the Context of the Arabic Grammatical Theory*. Leiden: Brill.

Baker, M. C. (1988). *Incorporation. A Theory of Grammatical Function Changing*. Chicago: Chicago University Press.

Baker, M. C. (2002). Building and Merging, not Checking: The Nonexistence of (Aux)-S-V-O Languages. *Linguistic Inquiry*, 33, 321-328.

Bardeas, S. (2009). *The Syntax of the Arabic DP*. PhD thesis, University of York.

Barth, J. (1894). *Die Nominalbildung in den semitischen Sprachen*. Leipzig: Hinrichs.

Barth, J. (1904). Beiträge zur Pluralbildung des Semitischen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 58, 431-446.

Barth, J. (1913). *Die Pronominalbildung in den Semitischen Sprachen*. Leipzig: Hinrichs.

Bartoli, M. G. (1925). *Introduzione alla Neolinguistica*. Geneve: Olschki.

Bassiouney, R. (2009). *Arabic Sociolinguistics*. Edinburgh: Edinburgh University Press.

Bauer, H. (1912). Noch einmal die Semitischen Zahlworter. *Zeitschrift der Morgenlandischen*

Gesellschaft, 66, 267-270.

Belletti, A. (1990). *Generalized Verb Movement*. Turin: Rosenberg & Selliers.

Belnap, R. K. (1999). A New Perspective on the History of Arabic Variation in marking Agreement with Plural Heads. *Folia Linguistica*, 33, 169-185.

Belnap, R. K. & Gee, J. (1994). Classical Arabic in Contact: The Transition to Near Categorical Agreement Patterns. In V. Cantarino, M. Eid & K. Walters (Eds.), *Perspectives on Arabic Linguistics VI* (pp. 121-149). Amsterdam: Benjamins.

Belnap, R. K. & Shabaneh, O. (1994). Variable Agreement and Nonhuman Plurals in Classical

and Modern Standard Arabic. In E. Broselow & M. Eid (Eds.), *Perspectives on Arabic Linguistics IV* (pp. 245-62). Amsterdam: Benjamins.

Benmamoun, E. (1999). Arabic Morphology: The Central Role of the Imperfective. *Lingua*, 108, 175-201.

Benmamoun, E. (2000). *The Feature Structure of Functional Categories*. Oxford: Oxford University Press.

Benmamoun, E. (2006). Construct State. In M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 477-482). Leiden: Brill.

Bentolila, F. (1981). *Grammaire fonctionnelle d'un parler berbère: Ait Seghrouchen d'Oum Jeniba (Maroc)*. Paris: SELAF.

Benveniste, E. (1935). *Origines de la formation des noms en indo-européen*. Paris: Maisonneuve.

Benveniste, E. (1954). Tendances récentes en linguistique générale. *Journal de Psychologie*, 48, 130-145.

Benveniste, E. (1960). "Être" et "avoir" dans leurs fonctions linguistiques. *Bulletin de la Société Linguistique de Paris*, 55, 113-134.

Blake, B. J. (2001). *Case*. Cambridge: Cambridge University Press.

Blanc, H. (1970). Dual and Pseudo-dual in the Arabic Dialects. *Language*, 46, 42-57.

Blau, J. (1979a). Redundant Pronominal Suffixes denoting Intrinsic Possession. *Journal of the Ancient Near Eastern Society*, 11, 31-37.

Blau, J. (1979b). The Use of Arabic Stative Verbs in the Suffix-tense to mark the Present. In

J. Blau (Ed.), *Studia Orientalia Memoriae D.H. Baneth Dedicata* (pp. 181-192). Jerusalem: The Magnes Press.

Blau, J. (1981). *The Emergence and Linguistic Background of Judaeo-Arabic*. Jerusalem: Ben-Zvi Institute.

Blau, J. (2002). *A Handbook of Early Middle Arabic*. Jerusalem: Max Schloessinger Memorial Foundation.

Bloomfield, L. (1933). *Language*. New York: Henry Holt and Company.

Bobaljik, J. (1994). What does Adjacency do? *MIT Working Papers in Linguistics*, 22, 1-32.

Boeckx, C. (2006). *Linguistic Minimalism: Origins, Concepts, Methods, and Aims*. Oxford: Oxford University Press.

Boeckx, C. & Stjepanovic, S. (2001). Head-ing toward PF. *Linguistic Inquiry*, 32, 345-355.

Bohas, G. (1997). *Matrices, étymons, racines: Éléments d'une théorie lexicologique du vocabulaire arabe*. Louvain: Peeters.

Bohas, G. (2000). *Matrices et étymons: Développements de la théorie*. Lausanne: Éditions du

Zèbre.

Bohas, G., Guillaume, J.-P., & Kouloughli, D. E. (1990). *The Arabic Linguistic Tradition*. London: Routledge.

Boskovic, Z. & Lasnik, H. (2007). *Minimalist Syntax: The Essential Readings*. Malden, Mass.: Blackwell.

Boukhris, F., Boumal, A., El Moujahid, E. H., & Souifi, H. (2008). *La nouvelle grammaire de l'amazighe*. Rabat: Publications de l'Institut Royal de la Culture Amazighe.

Boussofara-Omar, N. (2006). Diglossia. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 629-636). Leiden: Brill.

Bravmann, M. (1961). Genetic Aspects of the Genitive in the Semitic Languages. *Journal of the American Oriental Society*, 81, 386-394.

Brockelmann, K. (1908). *Grundriss der vergleichenden Grammatik der Semitischen Sprachen*. Hildesheim: Olms.

Brockelmann, K. (1910). *Précis de linguistique sémitique*. Paris: Geuthner.

Brustad, K. (2000). *The Syntax of Spoken Arabic*. Washington, DC: Georgetown University Press.

Brustad, K. (2007). Drink Your Milks: *-aat* as Individuation Marker in Levantine Arabic. In B. Gruendler & M. Cooperson (Eds.), *Classical Arabic Humanities in Their Own Terms: Festschrift for Wolfhart Heinrichs* (pp. 1-19). Leiden: Brill.

Buccellati, G. (1972). Review of Studi sui pronomi determinativi semitici. *Journal of the American Oriental Society*, 92, 296-297.

Buccellati, G. (1996). *A Structural Grammar of Babylonian*. Wiesbaden: Harrassowitz.

Campbell, R. (1996). Specificity Operators in Spec, DP. *Studia Linguistica*, 50, 161-188.

Cantineau, J. (1950). Racines et schèmes. In *Mélange offert à William Marçais par l'Institut d'études islamiques de l'Université de Paris* (pp. 119-124). Paris: Maisonneuve.

Caplice, R. I. & Snell, D. C. (2002). *Introduction to Akkadian*. Rome: Biblical Institute Press.

Carroll, D. W. (2008). *Psychology of Language*. Belmont, CA: Thomson & Wadsworth.

Carter, M. (1972). Twenty Dirhams in the Kitāb of Sibawahi. *Bulletin of School of Oriental And African School*, 35, 485-496.

Carter, M. (2004). *Sibawayhi*. London: I.B. Tauris and Oxford University Press.

Carter, M. (2006). Grammatical Tradition: History. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume II (pp. 182-190). Leiden: Brill.

Chekayri, A. (2005). Weak Verbs in Arabic. In A. Elgibali (Ed.), *Investigating Arabic. Current*

Parameters in Analysis and Learning (pp. 65-84). Leiden: Brill.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1966). *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. New York: Harper & Row.

Chomsky, N. (1970). Remarks on Nominalizations. In R. A. Jacobs & P. S. Rosenbaum (Eds.), *Readings in English Transformational Grammar* (pp. 184-221). Waltham, MA: Blaisdell.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, N. (2008). On Phases. In R. Freidin, C. P. Otero, & M. L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud* (pp. 133-166). Cambridge, Mass.: MIT Press.

Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Chomsky, N., Hauser, M. D., & Tecumseh Fitch, W. (2002). The Faculty of Language: What is It, Who has It, and how did It evolve? *Science*, 298, 1569-1579.

Chomsky, N., Hauser, M. D., & Tecumseh Fitch, W. (2005). The Evolution of the Language Faculty: Clarifications and Implications. *Cognition*, 97, 179-210.

Cinque, G. (1994). On the Evidence for Partial N-movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths Towards Universal Grammar. Studies in Honor of Richard S. Kayne*. Georgetown University Press.

Cinque, G. (1996). The "Antisymmetric" Programme: Theoretical and Typological Implications. *Linguistic Inquiry*, 32, 447-464.

Cinque, G. (1999). *Adverbs and Functional Heads*. Oxford: Oxford University Press.

Cinque, G. (2003). On Greenberg's Universal 20 and the Semitic DP. In L. O. Delsing, C. Falk, G. Josefsson, & H. A. Sigurdsson (Eds.), *Grammatik i fokus* (pp. 65-73). Lund: Wallin & Dalholm.

Cinque, G. (2005). A Note on Verb/Object Order and Head/Relative Clause Order. *University of Venice Working Papers in Linguistics*, 15, 49-104.

Cinque, G. (2006). *Restructuring and Functional Heads*. Oxford: Oxford University Press.

Cinque, G. (2010). *The Syntax of Adjectives*. Cambridge, Mass.: MIT Press.

Cinque, G. & Rizzi, L. (2008). The Cartography of Syntactic Structures. *STiL - Studies in Linguistics*, 2, 42-58.

Collett, M., Collett, T. S., & Srinivasan, M. V. (2006). Insect Navigation: Measuring Travel Distance across Ground and through Air. *Current Biology*, 16, 887-890.

Collett, M. & S., C. T. (2000). How do Insects use Path Integration for Their Navigation? *Biological Cybernetics*, 83, 245-259.

Comrie, B. (1981). *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

Corbett, G. G. (200). *Number*. Cambridge: Cambridge University Press.

Corriente, F. (1971). On the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology. *The Jewish Quarterly Review*, 62, 20-50.

Corriente, F. (1973). Again on the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology. *The Jewish Quarterly Review*, 64, 154-163.

Cowell, M. W. (1964). *A Reference Grammar of Syrian Arabic*. Washington, DC: Georgetown University Press.

Cuny, A. & Feghali, M. (1924). *Du genre grammatical en sémitique*. Paris: Geuthner.

De Dominicis, A. (2003). *Fonologia. Modelli e tecniche di rappresentazione*. Roma: Carocci.

de Lacy O'Leary, E. (1923). *Comparative Grammar of the Semitic Languages*. London: Trübner.

De Sacy, S. (1831). *Grammaire arabe à l'usage des élèves de l'École Spéciale des Langues Orientales Vivantes*. Paris: Imprimerie Royale.

de Vries, M. (2001). Patterns of Relative Clauses. In T. der Wouden & H. Broekhuis (Eds.), *Linguistics in the Netherlands 2001* (pp. 231-243). Benjamins: Amsterdam.

Delfitto, D. & Melloni, C. (2009). Compounds don't come Easy. *Lingue e linguaggio*, 1, 75-104.

den Dikken, M. (2006). *Relators and Linkers : the Syntax of Predication, Predicate Inversion, and Copulas*. Cambridge, Mass.: MIT Press.

den Dikken, M. (2007). Amharic Relatives and Possessives: Definiteness, Agreement, and the Linker. *Linguistic Inquiry*, 38, 302-320.

den Dikken, M. & Singhapreecha, P. (2004). Complex Noun Phrases and Linkers. *Syntax*, 7, 1-54.

Diem, W. (1998). *Fa-waylun li-l-qāsiyati qulūbuhum: Studien zum Arabischen Adjektivischen Satz*. Wiesbaden: Harrassowitz.

Diessel, H. (2005). Pronominal and Adnominal Demonstratives. In M. Haspelmath, M. S. Dryer, D. Gil & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 174-177). Oxford: Oxford University Press.

Drodzik, L. (1998). Functional Variations of the So-called Feminine Marker in Arabic. *Asian and African Studies*, 7, 23-44.

Dryer, M. S. (2005a). Definite Articles. In M. Haspelmath, M. S. Dryer, D. Gil & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 154-157). Oxford: Oxford University Press.

Dryer, M. S. (2005b). Indefinite Articles. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 157-161). Oxford: Oxford University Press.

Dryer, M. S. (2005c). Order of Adposition and Noun Phrase. In M. Haspelmath, M. S. Dryer, D. Gil & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 346-349). Oxford: Oxford University Press.

Edzard, L. (2006). Article, Indefinite. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 188-191). Leiden: Brill.

Eisele, J. (2001). Review of Parallels in Semitic Linguistics: The Development of Arabic *la-* and Related Semitic Particles. *Journal of the American Oriental Society*, 121, 327-329.

Eksell, K. (1984). On the Development of *d-* Particles as Genitive Exponents in Arabic Dialects. *Acta orientalia*, 45, 21-42.

Eksell, K. (2006). Pseudo-dual. In M. Eid, A. Elgibali, M. Woidich, & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 736-738). Leiden: Brill.

Enç, M. (1991). The Semantics of Specificity. *Linguistic Inquiry*, 22, 1-25.

Fassi-Fehri, A. (1993). *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.

Fassi-Fehri, A. (1999). Arabic Modifying Adjectives and DP Structures. *Studia Linguistica*, 53, 105–154.

Fassi-Fehri, A. (2000). Typological Ingredients of Mediterranean Attributive Adjectives. In S. Cristofaro & I. Putzu (Eds.), *Languages in the Mediterranean Area: Typology and Convergence* (pp. 83–104). Milano: Franco Angeli.

Fassi Fehri, A. & Vinet, M.-T. (2004). Distribution of Number and Classifier in Arabic and Chinese and Parametrization. *Linguistic Research*, 9, 9-52.

Fassi Fehri, A. & Vinet, M.-T. (2007). Number and Classifier Distributions in Arabic and Chinese and Their Complementarity. In A. Zaenen, J. Simpson, T. Holloway King, J. Grimshaw, J. Maling, & C. Manning (Eds.), *Architectures, Rules, and Preferences : Variations on Themes of Joan Bresnan* (pp. 163-197). Stanford: CSLI Publications.

Feghali, M. (1928). *Syntaxe des Parlers Arabes Actuels du Liban*. Geuthner: Paris.

Ferguson, C. (1959a). The Arabic Koine. *Language*, 35, 616-630.

Ferguson, C. (1959b). Diglossia. *Word*, 15, 325-340.

Ferguson, C. (1989). Grammatical Agreement in Classical Arabic and the Modern Dialects: A Response to Pidginization Hypothesis. *Al-Arabiyya*, 22, 5-18.

Ferguson, C. (1991). Epilogue: Diglossia Revisited. *Southwest Journal of Linguistics*, 10, 214-234.

Ferrando, I. (2000). Le morphème de liaison /an/ en arabe andalou: notes de dialectologie comparée. *Oriente Moderno*, 19, 25-46.

Ferrando, I. (2006). The Plural of Paucity in Arabic and Its Actual Scope: On Two Claims by Sibawayhi and Al-Farra'. In S. Boudelaa (Ed.), *Perspectives on Arabic Linguistics, XVI* (pp.

39-62). Amsterdam: Benjamins.

Fischer, W. (1982). *Grundriss der Arabischen Philologie*. Wiesbaden: Reichert.

Fischer, W. (2006a). Adjective. In K. Versteegh, M. Eid, A. Elgibali, M.Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 16-21). Leiden: Brill.

Fischer, W. (2006b). Classical Arabic. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 397-405). Leiden: Brill.

Fischer, W. & Jastrow, O. (1980). *Handbuch der Arabischen Dialekte*. Wiesbaden: Harrassowitz.

Fleisch, H. (1961). *Traité de philologie arabe*. Beyrouth: Imprimerie catholique.

Fleischer, H. L. (1854). *Kleinere Schriften*. Leipzig: Hirzel.

Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, Mass.: MIT Press.

Frascarelli, M. & Puglielli, A. (2006). A Comparative Analysis of Restrictive and Appositive Relative Clauses in Cushitic Languages. In L. Brugè, G. Giusti, N. Munaro, W. Schweikert, & G. Turano (Eds.), *Contributions to the XXX Incontro di Grammatica Generativa* (pp. 279–303). Venezia: Cafoscarina.

Freeze, R. (1992). Existentials and Other Locatives. *Language*, 68, 553–595.

Fück, J. (1950). *Arabiya. Untersuchungen zur Arabischen Sprachund Stilgeschichte*. Berlin: Akademie-Verlag.

Garber, P. A. (1987). Foraging Strategies among Living Primates. *Annual Review of Anthropology*, 16, 339-364.

Garber, P. A. & Jelinek, P. E. (2006). Travel Patterns and Spatial Mapping in Nicaraguan Mantled Howler Monkeys (Alouatta Palliata). In A. Estrada, P. A. Garber, M. Pavelka, & L. Leandra (Eds.), *New Perspectives in The Study of Mesoamerican Primates* (pp. 207-310). New York: Springer.

Garbini, G. & Durand, O. (1994). *Introduzione alle lingue semitiche*. Brescia: Paideia.

Gelb, I. (1930). La mimazione e la nunazione nelle lingue semitiche. *Rivista di Studi Orientali*, 12, 217-263.

Gensler, O. (2004). The Arabic *mā* - *min* Construction: Rethinking the Priority of Restricted

Relatives. Talk Presented at *Syntax of the World's Languages (SWL 1)*, Leipzig, August 5-8,

2004.

Giacalone Ramat, A. (2005). Persistence and Renewal in the Relative Pronoun Paradigm: the

Case of Italian. *Folia Linguistica Historica*, 26, 115-38.

Gil, D. (2005). Genitives, Adjectives and Relative Clauses. In M. Haspelmath, M. S. Dryer, D. Gil & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 246-250). Oxford: Oxford University Press.

Giorgi, A. (2010). *About the Speaker: Towards a Syntax of Indexicality*. Oxford: Oxford University Press.

Giusti, G. (2001). The Birth of a Functional Category. From Latin *ille* to the Romance Article and Personal Pronoun. In G. Cinque (Ed.), *Current Studies in Italian Syntax: Essays offered to Lorenzo Renzi* (pp. 99-132). Amsterdam: Elsevier.

Giusti, G. (2002). The Functional Structure of Noun Phrases: A Bare Phrase Structure Approach.In G. Cinque (Ed.), *Functional Structure in DP and IP* (pp. 54-90). Oxford: Oxford University Press.

Goldenberg, G. (1988). Subject and Predicate in Arab Grammatical Tradition. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 138, 39-73.

Grande, F. (2003). La struttura originaria della radice araba nel pensiero di Ibn Ginni. *Annali di Ca' Foscari*, 42, 57-83.

Grande, F. (2011). Deriving Arabic Nunation and Its Properties in the Light of Greenberg's Numeral Classifier Hypothesis. Talk Presented at Jil Jadid Conference, Austin, February 18-19, 2011.

Grandy, D. A. (2010). *Everyday Quantum Reality*. Bloomington, Indiana: Indiana University Press.

Greenberg, J. H. (1950). The Patterning of Root Morphemes in Semitic. *Word*, 6, 162-181.

Greenberg, J. H. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In J. H. Greenberg (Ed.), *Universals of Language* (pp. 73-113). Cambridge, Mass.: MIT Press.

Greenberg, J. H. (1990a). Dynamic Aspects of Word Order in the Numeral Classifier. In K. Denning & S. Kemmer (Eds.), *On Language: Selected Writings of Joseph H. Greenberg* (pp. 227-240). Stanford: SUP.

Greenberg, J. H. (1990b). Numeral Classifiers and Substantival Number. In K. Denning & S. Kemmer (Eds.), *On Language: Selected Writings of Joseph H. Greenberg* (pp. 166-193). Stanford: SUP.

Guella, N. (2010). On the Relative Clause in the Arabic Dialects of the Maghreb. In H. Dhaouadi & A. Ghersetti (Eds.), *Synergies Monde Arabe* (pp. 101-110). Riyadh: GERFLINT.

Guerssel, M. & Lowenstamm, J. (1996). Ablaut in Classical Arabic Measure I Active Verbal Forms. In J. Lecarme, J. Lowenstamm, & U. Shlonsky (Eds.), *Studies in Afro-asiatic Grammar* (pp. 123-134). Amsterdam: Holland Academic Graphics.

Guillaume, J. P. (1992). Le statut de l'adjectif dans la tradition grammaticale arabe. *Histoire Epistémologie Langage*, 14, 59-74.

Guillaume, J. P. (1998). Les discussions des grammairiens arabes à propos du sens des marques d'i'rāb. *Historie Epistémologie Langage*, 20, 43-62.

Guillaume, J. P. (2000). Grammaires arabes. In B. Colombat & E. Laczano (Eds.), *Corpus représentatif des grammaires et des traditions linguistiques, tome II* (pp. 159-184). Paris:

Shels.

Guillaume, J. P. (2006). Grammatical Tradition: Approach. In K. Versteegh, M. Eid, A. Elgibali, M.Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume II (pp. 175-182). Leiden: Brill.

Gully, A. (1995). *Grammar and Semantics in Medieval Arabic: A Study of Ibn Hisham's Mughni-l-labib*. Richmond: Curzon Press.

Haddad, G. & Kenstowicz, M. (1980). A Note on the Parallels between the Definite Article and the Relative Clause Marker in Arabic. *Studies in the Linguistic Sciences*, 10, 141-148.

Haegeman, L. (1991). *Introduction to Government and Binding Theory*. Malden, Mass.: Blackwell.

Hallaq, W. B. (1993). Was Al-Shafi'i the Master Architect of Islamic Jurisprudence? *International Journal of Middle East Studies*, 25, 587-605.

Hallaq, W. B. (2005). *The Origins and Evolution of Islamic Law*. Cambridge: Cambridge University Press.

Harrell, R. S. (1962). *A Short Reference Grammar of Moroccan Arabic*. Washington, DC: Georgetown University Press.

Harrell, R. S., Abu-Talib, M., & Carroll, W. S. (1965). *A Basic Course in Moroccan Arabic*. Washington, DC: Georgetown University Press.

Hashemi, A. (2006a). Ḥāl. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 55-555). Leiden: Brill.

Hashemi, A. (2006b). Gender. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume II (pp. 155-163). Leiden: Brill.

Hasselbach, R. (2007a). Demonstratives in Semitic. *Journal of the American Oriental Society*, 127, 1-27.

Hasselbach, R. (2007b). External Plural Markers in Semitic: A New Assessment. In C. L. Miller (Ed.), *Studies in Semitic and Afroasiatic Linguistics Presented to Gene B. Gragg* (pp. 123-138). Chicago: The Oriental Institute.

Haywood, J. A. (1965). *Arabic Lexicography: Its History, and Its Place in the General History of Lexicography*. Leiden: Brill.

Heine, B. (1997). *Possession. Cognitive Sources, Forces and Grammaticalization*. Cambridge: Cambridge University Press.

Heine, B. & Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction*. Oxford: Oxford University Press.

Heringa, H. (2007). Appositional Constructions: Coordination and Predication. In *Proceedings of the Fifth Semantics in the Netherlands Day* (pp. 67-82).

Hetzron, R. (1967). Agaw Numerals and Incongruence in Semitic. *Journal of Semitic Studies*, 12, 169-197.

Hetzron, R. (1976). Two Principles of Genetic Reconstruction. *Lingua*, 38, 89-108.

Hetzron, R. (1978). On the Relative Order of Adjectives. In H. J. Seiler (Ed.), *Language Universals: Papers from the Conference held at Gummersbach / Cologne, Germany, October 3-8, 1976* (pp. 165-184). Tübingen: Gunter Narr Verlag.

Hetzron, R. (1997). *The Semitic Languages*. London: Routledge.

Higgins, R. (1979). *The Pseudo-Cleft Construction in English*. New York: Garland.

Holes, C. (2004). Tanwin in the Arabic Dialects of Eastern and South-eastern Arabia. *Estudios de dialectología norteafricana y andalusí*, 8, 89-97.

Hopper, P. J. & Closs Traugott, E. (2003). *Grammaticalization*. London: Cambridge University Press.

Hoyt, F. (2006). Noun Phrase. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 428-434). Leiden: Brill.

Ihsane, T. (2003). Demonstratives and Reinforcers in Arabic, Romance and Germanic. In J. Lecarme (Ed.), *Research in Afroasiatic Grammar II* (pp. 263-285). Amsterdam: Benjamins.

Ingham, B. (1982a). *North-east Arabian Dialects*. London: Kegan Paul International.

Ingham, B. (1982b). Notes on the Dialect of the Dhafir of North-eastern Arabia. *Bulletin of the School of Oriental and African Studies*, 45, 245-259.

Ingham, B. (1986). Notes on the Dialect of the Al murra of Eastern and Southern Arabia. *Bulletin of the School of Oriental and African Studies*, 49, 271-291.

Ingham, B. (1994). *Najdi Arabic: Central Arabian*. Amsterdam: Benjamins.

Ingham, B. (1997). *Arabian diversions: Studies on the Dialects of Arabia*. Reading: Ithaca Press.

Ingham, B. (2006). Najdi Arabic. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 326-334). Leiden: Brill.

Iwasaki, S. & Ingkaphirom, P. (2005). *A Reference Grammar of Thai*. Cambridge: Cambridge University Press.

Johnstone, T. M. (1961). Some Characteristics of the Dosiri Dialect of Arabic as Spoken in Kuwait. *Bulletin of the School of Oriental and African Studies*, 24, 249-297.

Jones, A. (2005). *Arabic Through the Qurían*. The Islamic Texts Society: Cambridge.

Julien, M. (2002). *Syntactic Heads and Word Formation*. Oxford: Oxford University Press.

Kahle, E. (1975). *Studien zur Syntax des Adjektivs im vorklassischen Arabisch*. Erlangen: Lüling.

Kaiser, L. (1999). *The Morphosyntax of Clausal Nominalization Constructions*. PhD thesis, University of Yale.

Kayne, R. S. (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.

Kayne, R. S. (1998). Overt vs. Covert Movement. *Syntax*, 1, 128-191.

Kern, S. (2010). Prenominal Relative Clauses: A Typology Study. Ms., Unité Mixte de Recherche Dynamique Du Langage.

Khan, G. (1984). Object Markers and Agreement Pronouns in Semitic Languages. *Bulletin of the School of Oriental and African Studies*, 47, 468-500.

Khan, M. (1996). *The Noble Quran*. Madinah: King Fahd Printing Complex.

Kouloughli, D. E. (1979). Sur le traitement des glides dans la phonologie de l'arabe standard. *Analyses Théories*, 1, 54-93.

Kouloughli, D. E. (1999). Y a-t-il une syntaxe dans la tradition arabe? *Historie Epistémologie Langage*, 21, 45-64.

Kouloughli, D. E. (2001). Sur le statut linguistique du tanwin: contribution à l'étude du système déterminatif de l'arabe. *Arabica*, 48, 20-50.

Kouloughli, D. E. (2007). Sur la valeur du tanwin. Nouvelle contribution a l'étude du système determinatif de líarabe. *Arabica*, 54, 94-131.

Kremers, J. (2003). Adjectival Agreement in the Arabic Noun Phrase. In M. van Koppen, J. Sio, & M. de Vos (Eds.), *Proceedings of ConSOLE XI* (pp. 1-13). Leiden: s.e.

Kurilowicz, J. (1950). La mimation et l'article en arabe. *Archiv Orientalni*, 18, 323-328.

Kurilowicz, J. (1961). *L'apophonie en sémitique*. Warsaw: Polska Akademia Nauk.

Kurilowicz, J. (1972). *Studies in Semitic Grammar and Metrics*. Warsaw: Polska Akademia Nauk.

Lambdin, T. O. (1971). The Junctural Origin of the West Semitic Definite Article. In H. Goedicke (Ed.), *Near Eastern Studies in Honor of W. F. Albright* (pp. 315-333). Baltimore: Johns Hopkins University Press.

Lancioni, G. (1996). *Ordini lineari marcati in arabo. Un'analisi generativa*. Roma: Bardi.

Lane, E. W. (1863). *Arabic-English Lexicon*. London: Williams and Norgate.

Langacker, R. W. (1977). Syntactic Reanalysis. In C. N. Li (Ed.), *Mechanisms of Syntactic Change* (pp. 57-139). Austin: University of Texas Press.

Larcher, P. (1995). Où il est montré qu'en arabe classique la racine n'a pas de sens et qu'il n'y a pas de sens à dériver d'elle. *Arabica*, 42, 291-314.

Larcher, P. (2007). Linguistique arabe: état de la recherche. *Arabica*, 54, 246-261.

Lawler, J. R. (1992). *Rimbaud's Theatre of the Self.* Cambridge, Mass.: Harvard University Press.

Leaman, O. (2009). *Islamic Philosophy*. Cambridge: Cambridge University Press.

Lehmann, W. P. (1993). *Theoretical Bases of Indo-European Linguistics*. London: Routledge.

Lepschy, G. C. (1966). *La linguistica strutturale*. Torino: Einaudi.

Leu, T. (2008). *The Internal Syntax of Determiners*. PhD thesis, New York University.

Levin, A. (1986). The Medieval Arabic Term *kalima* and the Modern Linguistic Term morpheme: Similarities and Differences. In M. Sharon (Ed.), *Studies in Islamic History and Civilization in Honour of Professor David Ayalon* (pp. 423-446). Leiden: Brill.

Levin, A. (2006). Kalima. In M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume II (pp. 545-548). Leiden: Brill.

Lipinski, E. (1997). *Semitic Languages : Outline of a Comparative Grammar*. Leuven: Peeters.

Lévi-Strauss, C. (1962). *La Pensée sauvage*. Paris: Plon.

Lyons, C. (1999). *Definiteness*. London: Cambridge University Press.

Lyons, J. (1968). *Introduction to Theoretical Linguistics*. London: Cambridge University Press.

MacDonald, M. C. A. (2008). Ancient North Arabian. In R. D. Woodard (Ed.), *The Ancient Languages of Syria-Palestine and Arabia* (pp. 179-224). Cambridge: Cambridge University Press.

Mahfoudhi, A. (2007). The Place of the Etymon and the Phonetic Matrix in the Arabic Mental Lexicon. *Journal of Arabic and Islamic Studies*, 7, 74-102.

Marçais, P. (1956). *Le parler arabe de Djidjelli, Nord constantinois, Algérie*. Paris: Maisonneuve.

Marantz, A. (1997). No Escape from Syntax: Don't try Morphological Analysis in the Privacy of Your Own Lexicon. *University of Pennsylvania Working Papers in Linguistics*, 4, 201-225.

Marçais, W. (1930). La diglossie arabe. *L'Enseignement Public*, 97, 401-409.

Marotta, G. (1999). Degenerate feet nella fonologia metrica dell'italiano. In P. Benincà, A. Mioni, & L. Vanelli (Eds.), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia* (pp. 97-116). Roma: Bulzoni.

Mascitelli, D. (2006). *L'arabo in epoca preislamica: formazione di una lingua*. Roma: L'Erma di Bretschneider.

Matras, Y. (2009). Defining the Limits of Grammatical Borrowing. In A. Marcantonio (Ed.), *The Indo-European Language Family: Questions about Its Status* (pp. 1-20). Washington, DC: Institute for the Study of Man Inc.

Matras, Y. & Bakker, P. (2003). The Study of Mixed Languages. In Y. Matras & P. Bakker (Eds.), *The Mixed Language Debate: Theoretical and Empirical Advances* (pp. 1-20). Berlin: Mouton de Gruyter.

Matushansky, O. (2002). Tipping the Scales: The Syntax of Scalarity in the Complement of

seem. Syntax, 5, 219-276.

Matushansky, O. (2006). Head Movement in Linguistic Theory. *Linguistic Inquiry*, 37, 69-109.

McCarthy, J. J. (1979). *Formal Problems in Semitic Phonology and Morphology*. PhD thesis, MIT.

Medjell, G. (2003). Notes on (Non)restrictedness in Relative Clauses in Egyptian Arabic. In J. Aguadé & I. Ferrando (Eds.), *AIDA 5th Conference Proceedings* (pp. 539-544). Cadiz: Servicio de Publicaciones, Universidad de Cadiz.

Mion, G. (2009). L'indétermination nominale dans les dialectes arabes: une vue d'ensemble. In A. Arioli (Ed.), *Miscellanea Arabica 2009* (pp. 215-231). Rome: Nuova Cultura.

Moro, A. (2000). *Dynamic Antisymmetry*. Cambridge, Mass.: MIT Press.

Moro, A. (2004). Linear Compression as a Trigger for Movement. In A. Breitbarth & H. van Riemsdijk (Eds.), *Triggers* (pp. 387-429). Berlin: Mouton de Gruyter.

Moro, A. (2008). *The Boundaries of Babel: the Brain and the Enigma of Impossible Languages*. Cambridge, Mass.: MIT Press.

Moscati, S. (1958). On Semitic Case-endings. *Journal of Near Eastern Studies*, 17, 142-144.

Moscati, S., Spitaler, A., Ullendorf, E., & von Soden, W. (1980). *An Introduction to the Comparative Grammar of the Semitic Languages : Phonology and Morphology*. Wiesbaden: Harrassowitz.

Mughazy, M. (2006). Relative Clause. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume IV (pp. 60-70). Leiden: Brill.

Nöldeke, T. (1897). *Zur Grammatik des classischen Arabisch*. Wien: s.e.

Nyberg, S. H. (1920). Wortbildung mit Prafixen in den Semitischen Sprachen. *Le Monde Oriental*, 14, 177-288.

Ouhalla, J. (2000). Possession in Sentences and Noun Phrases. In J. Lecarme, J. Lowenstamm, & U. Shlonsky (Eds.), *Research in Afroasiatic Grammar* (pp. 221-242). Amsterdam: Benjamins.

Ouhalla, J. (2004). Semitic Relatives. *Linguistic Inquiry*, 35, 288-300.

Ouhalla, J. (2010). Variation and Change in Possessive Noun Phrases: The Evolution of the Analytic Type and Loss of the Synthetic Type. Ms., University College Dublin.

Ouwayda, S. A. (2009). Men and Their Apples: Dividing Plural and Agreement Plural. Screening Paper, University of Southern California.

Owens, J. (1988). *The Foundations of Grammar: an Introduction to Medieval Arabic Grammatical Theory*. Amsterdam: Benjamins.

Owens, J. (1990). *Early Arabic Grammatical Theory: Heterogeneity and Standardization*. Amsterdam: Benjamins.

Owens, J. (2001). Arabic Sociolinguistics. *Arabica*, 48, 419-469.

Owens, J. (2005). The Grammatical Tradition and Arabic Language Teaching: A View from Here. In A. Elgibali (Ed.), *Investigating Arabic. Current Parameters in Analysis and Learning* (pp. 103-116). Leiden: Brill.

Owens, J. (2006). *A Linguistic History of Arabic*. Oxford: Oxford University Press.

Palmer, L. R. (1954). *The Latin Language*. London: Faber and Faber.

Palva, H. (2006). Dialects: Classification. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich &

A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 604-613). Leiden: Brill.

Passino, D. (2009). Phonological Traces of Morphosyntactic Phases. The Phonology/Syntax Interface in the Varieties of Italian. Ms., Università dell'Aquila.

Pat-el, N. (2009). The Development of the Semitic Definite Article. *Journal of Semitic Studies*, 54, 19-49.

Peled, Y. (2006). Ḍamīr. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 555–559). Leiden: Brill.

Pennacchietti, F. (1968). *Studi sui pronomi determinativi semitici*. Napoli: Istituto Universitario Orientale.

Pesetsky, D. & Torrego, E. (2004). Tense, Case, and the Nature of Syntactic Categories. In J. Guèron & J. Lecarme (Eds.), *The Syntax of Time* (pp. 495-537). Cambridge, Mass.: MIT Press.

Peters, J. R. T. M. (1976). *God's created Speech: A Study in the Speculative Theology of the Mu'tazili*. Leiden: Brill.

Petracek, K. (1965). Die Isomorphie im System der Arabischen Pluralbildung. In A. Heintz, M. Karag, T. Milewski, J. Safarewicz, & W. Taszycki (Eds.), *Symbolae Linguisticae in Honorem Georgii Kurylowicz* (pp. 227-229). Warsaw: Polska Akademia Nauk.

Petracek, K. (1981). Le système de l'arabe dans une perspective diachronique. *Arabica*, 28, 162-177.

Pinkerton, J. (1811). The Curious and Remarkable Voyages and Travels of Marco Polo. In J. Pinkerton (Ed.), *A General Collection of the Best and Most Interesting Voyages and Travels in All Parts of the World* (pp. 101–178). London: Longman, Hurst, Rees, and Orme.

Prochazka, T. (1988). *Saudi Arabian Dialects*. London: Kegan Paul International.

Prunet, J.-F., Béland, R., & Idrissi, A. (2000). The Mental Representation of Semitic Words. *Linguistic Inquiry*, 31, 609-647.

Putzu, I. & Ramat, P. (2001). Articles and Quantifiers in the Mediterranean Languages: A Typological-Diachronic Analysis. In W. Bisang (Ed.), *Aspects of Typology and Universals* (pp. 99-132). Berlin: Akademie Verlag.

Rabin, C. (1951). *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rabin, C. (1955). The Beginnings of Classical Arabic. *Studia Islamica*, 4, 19-37.

Rabin, C. (1969). The Structure of the Semitic System of Case-endings. In *Proceedings of the International Conference on Semitic Studies, held in Jerusalem, 19-23 July 1965* (pp. 190-204). Jerusalem: Israel Academy of Sciences and Humanities.

Radford, A. (2004). *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.

Ratcliffe, R. (1998). Defining Morphological Isoglosses: The "Broken" Plural and Semitic Subclassification. *Journal of Near Eastern Studies*, 57, 81-123.

Ratcliffe, R. (2006). Number. In M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 439-447). Leiden: Brill.

Ray, T. S. (2003). Oriya. In G. Cardona & D. Jain (Eds.), *The Indo-Aryan Languages* (pp. 485-522). New York: Routledge.

Rebuschi, G. (2003). Basque Semi-free Relative Clauses and the Structure of DPs. *Lapurdum*, 8, 457-478.

Reckendorf, H. (1895). *Die Syntaktischen Verhältnisse des Arabischen*. Leiden: Brill.

Retsoe, J. (1995). Pronominal State in Colloquial Arabic: A Diachronic Attempt. *Studia Orientalia*, 75, 183-192.

Ritter, E. (1991). Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew. *Syntax and Semantics*, 25, 37-62.

Rizzi, L. (1986). Null Objects in Italian and the Theory of pro. *Linguistic Inquiry*, 17, 501-557.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Rizzi, L. (1997). The Fine Structure of the Left Periphery. In L. Haegeman (Ed.), *Elements of Grammar* (pp. 281-337). Dordrecht: Kluwer.

Rosenhouse, J. (2006). Bedouin Dialects. In K. Versteegh, M. Eid, A. Elgibali, M.Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 259-269). Leiden: Brill.

Rosenthal, F. (1989). *The History of al-Tabari, Volume I - General Introduction and From the Creation to the Flood*. Albany: State University of New York Press.

Ross, J. R. (1967). *Constraints on Variables in Syntax*. PhD thesis, MIT.

Rowlett, P. (2007). *The Syntax of French*. Cambridge: Cambridge University Press.

Rubin, A. (2005). *Studies in Semitic Grammaticalization*. Winona Lake, IN: Eisenbrauns.

Saah, M. (2010). Relative Clauses in Akan. In E. O. Aboh & J. Essegbey (Eds.), *Topics in Kwa Syntax* (pp. 91-108). Springer: New York.

Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.

Schachter, P. & Otanes, F. T. (1983). *Tagalog Reference Grammar*. Berkeley and Los Angeles:

University of California Press.

Scheer, T. (2004). *A Lateral Theory of Phonology: What is CVCV, and why should It be?* Berlin: Mouton de Gruyter.

Scheer, T. (2007). (Direct) Interface without Big Brothers. Handout Presented at the Second Student Conference on Formal Linguistics, Poznan, 21-22 April 2007.

Scheer, T. (2008). Why the Prosodic Hierarchy is a Diacritic and why the Interface must be Direct. In J. M. Hartmann, V. Hegedüs, & H. van Riemsdijk (Eds.), *Sounds of Silence: Empty Elements in Syntax and Phonology* (pp. 145–184). Amsterdam: Elsevier.

Seidensticker, T. (2006). Lexicography: Classical Arabic. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume III (pp. 30-37). Leiden: Brill.

Shibatani, M. & Bin Makhashen, K. A. (2009). Nominalization in Soqotri, a South Arabian Language of Yemen. In L. W. Wetzels (Ed.), *Endangered Languages: Contributions to Morphology and Morpho-syntax* (pp. 9-31). Leiden: Brill.

Shlonsky, U. (1997). *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. Oxford: Oxford University Press.

Shlonsky, U. (2004). The Form of the Semitic Noun Phrase. *Lingua*, 114, 1465-1526.

Simpson, A. (2005). Classifiers and DP Structure in Southeast Asia. In G. Cinque & R. Kayne (Eds.), *The Oxford Handbook of Comparative Syntax* (pp. 806-838). Oxford: Oxford University Press.

Stassen, L. (2005). Zero Copula for Predicate Nominals. In M. Haspelmath, M. S. Dryer, D. Gil & B. Comrie (Eds.), *The World Atlas of Language Structures* (pp. 486-489). Oxford: Oxford University Press.

Suleiman, Y. (1999). *Arabic Grammatical Tradition: A Study in Ta'lil*. Edinburgh: Edinburgh University Press.

Suleiman, Y. (2006). 'Arabiyyah. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume I (pp. 173-178). Leiden: Brill.

Suzuki, I. (1973). Outline of the Topography of Syria. *Bulletin of Universities and Institutes, The University Museum, The University of Tokyo.*

Szemerényi, O. J. L. (1996). *Introduction to Indo-European Linguistics*. Oxford: Oxford University Press.

Tanase Dogaru, M. (2008). Pseudo-partitives and Silent Classifiers in Romanian. In S. Blaho, C. Constantinescu, & E. Schoorlemmer (Eds.), *Proceedings of ConSOLE XV* (pp. 295-320). Leiden: s.e.

Tao, L. (2009). The Importance of Discourse Analysis for Linguistic Theory. In Z. Frajzyngier, A. Hodges, & H. D. Rood (Eds.), *Linguistic Diversity and Language Theories* (pp. 285-318). Amsterdam: Benjamins.

Testen, D. (1998). *Parallels in Semitic linguistics: The Development of Arabic la- and Related Semitic Particles*. Leiden; Boston: Brill.

Thornton, A. M. (2005). *Morfologia*. Roma: Carocci.

Townsend, D. J. & Bever, T. G. (2001). *Sentence Comprehension: The Integration of Habits and Rules*. Cambridge, MA: MIT Press.

Tropper, J. (2004). Gedanken zum Pluralmarker *u* im Semitischen. *Journal of Semitic Studies*, 54, 19-49.

Troseth, E. (2009). Degree Inversion and Negative Intensifier Inversion in the English DP. *The Linguistic Review*, 26, 37-65.

Trudgill, P. (1974). Linguistic Change and Diffusion: Description and Explanation in Sociolinguistic Dialect Geography. *Language in Society*, 3, 215-246.

Trussell, S. (1979). *Kiribati (Gilbertese): Grammar Handbook*. Battleboro, VT: School for International Training.

Ullendorff, E. (1965). The Form of the Definite Article in Arabic and Other Semitic Languages. In G. A. Makdisi (Ed.), *Arabic and Islamic Studies in Honor of Hamilton A. R. Gibb* (pp. 631-637). Leiden: Brill.

Unterbeck, B. (1999). Gender: New Light on an Old Category. In B. Unterbeck & M. Rissanen (Eds.), *Gender in Grammar and Cognition* (pp. xv - xlvi). Berlin: Mouton de Gruyter.

Veccia Vaglieri, L. (1937). *Grammatica teorico-pratica della lingua araba*. Roma: Istituto per l'Oriente.

Versteegh, K. (1977). *Greek Elements in Arabic Linguistic Thinking*. Leiden: Brill.

Versteegh, K. (1987). *Pidginization and Creolization: The Case of Arabic*. Amsterdam: Benjamins.

Versteegh, K. (1995). *The Explanation of Linguistic Causes. Az-Zaǧǧāgī's Theory of Grammar*. Amsterdam: Benjamins.

Versteegh, K. (1997a). *The Arabic Language*. Edinburgh: Edinburgh University Press.

Versteegh, K. (1997b). *Landmarks in Linguistic Thought III: The Arabic Linguistic Tradition*. Edinburgh: Edinburgh University Press.

Versteegh, K. (2004). Pidginization and Creolization Revisited: The Case of Arabic. In M. Haak, R. de Jong, & K. Versteegh (Eds.), *Approaches to Arabic Dialects* (pp. 343-358). Leiden: Brill.

Vicente, A. (2006). Relative Pronoun (Arabic Dialects). In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic Language and Linguistics*, volume IV (pp. 70-72). Leiden: Brill.

Visser, R. (2005). Diglossia and Language Acquisition in Egypt. Bachelor Thesis, University of Amsterdam.

Voigt, R. (1997). Zur Nominal- und Verbalnasalierung im Semitischen. *Wiener Zeitschrift für*

die Kunde des Morgenlandes, 87, 207–230.

Voigt, R. (1998). Der Artikel im Semitischen. *Journal of Semitic Studies*, 43, 221–258.

Voigt, R. (1999). Nominal (and Verbal) Nasalization and the Nominal Plural Morphemes in Semitohamitic. In M. Lamberti & L. Tonelli (Eds.), *Afroasiatica Tergestina: Papers from the 9th Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics* (pp. 11-22). Padua: Unipress.

Vollers, K. (1906). *Volksprache und Schriftsprache im alten Arabien*. Amsterdam: Oriental Press.

von Soden, W. (1955). *Grundriss der akkadischen Grammatik*. Rome: Pontificio Istituto Biblico.

Walter, M. A. (2003). Kalaam, Kalaarbaam: An Arabic Speech Disguise in Hadramaut. In I. Mey, G. Pizer, H.-Y. Su, & S. Szmania (Eds.), *Proceedings of the 10th Annual Symposium about Language and Society - Austin (SALSA)* (pp. 1-10). Austin: s.e.

Watson, J. C. E. (2002). *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.

Woidich, M. & Heinen-Nasr, R. (2004). *Kullū Tamām! An Introduction to Egyptian Colloquial Arabic*. Cairo: The American University in Cairo Press.

Wright, W. (1890). *Lectures on the Comparative Grammar of the Semitic Languages.*Cambridge: Cambridge University Press.

Wright, W. (1896). *A Grammar of the Arabic Language*. Cambridge: Cambridge University Press.

Zabbal, Y. (2002). The Semantics of Number in the Arabic Noun Phrase. Master's thesis, University of Calgary.

Zaborski, A. (1976). The Semitic External Plural in an Afroasiatic Perspective. *Afroasiatic Linguistics*, 3, 1-9.

Zaborski, A. (1995). First Person Pronouns in Arabic in the Light of Arabic and Hamito-semitic Dialectology. *Studia Orientalia*, 75, 289-294.

Zanuttini, R. (1997). *Negation and Clausal Structure: A Comparative Study of Romance Languages*. Oxford: Oxford University Press.

PRIMARY SOURCES

’Abū Bišr ‘Amr ’Ibn ‘Uthmān ’Ibn Qanbar al-Bišrī Sībawayhi. *Al-Kitāb*. Ed. H. Derenbourg, Imprimerie Nationale, Paris, 1881.

’Abū l-Qāsim Maḥmūd Ibn ‘ Umar Al-Zamaxšarī. *al-Kašsāf ‘an ḥaqā’iq ġawāmiḍ al-tanzīl wa-‘uyūn al-’aqāwīl fī wujūh al-Ta’wīl*. Maṭba‘ Nassau Lees, Calcutta, 1856.

Jalāl al-Dīn ’Abū l-Faḍl ‘Abd al-Raḥmān Al-Suyūṭī. *Al-Muzhir fī ‘ulūm al-luġah wa-’anwā‘i-hā*. Maktabat Dār al-Turāṯ, al-Qāhirah, s.d.

Muḥammad Ibn Mukarram Ibn ‘Alī Riḍwān Ibn Manẓūr. *Lisān al-‘Arab*. Dār al-Ma‘ārif, al-Qāhirah, 1981.

Muwaffaq al-Dīn ’Abū l-Barā’ Ya‘īš Ibn Ya‘īš. *Šarḥ al-Mufaṣṣal*. Dār al-Kutub al-‘Ilmiyyah, Bayrūt, 2001.

Rāḍī l-Dīn Muḥammad Ibn al-Ḥasan Al-’Astārābāḏī. *Šarḥ al-Kāfiyah*. Manšūrāt Jāmi‘at Qāryūnis, Binġāzī, 1996.