The Arabic noun phrase

A minimalist approach

Published by
LOT
Trans 10
3512 JK Utrecht
the Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
http://wwwlot.let.uu.nl/

ISBN: 90-76864-38-1
NUR 632

This dissertation was typeset in LaTeX by the author.

The Arabic noun phrase
A minimalist approach

Een wetenschappelijke proeve op het gebied van de Letteren

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen
op gezag van de Rector Magnificus,
Prof. Dr. C. W. P. M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen
op maandag 16 juni 2003,
des namiddags om 1.30 uur precies

door

Joost Merijn Kremers

geboren op 22 april 1974 te Nijmegen

Promotores:	Prof. Dr. C. H. M. Versteegh
	Prof. Dr. E. J. Reuland (Universiteit Utrecht)
Co-promotor:	Dr. A. P. Foolen
Manuscriptcommissie:	mw. Prof. Dr. A. M. C. van Kemenade (voorzitter)
	Prof. Dr. M. B. H. Everaert
	Dr. W. E. Ditters

Contents

Appendices

Acknowledgements

A number of people have contributed to the realisation of this thesis, and I would like to take this opportunity to thank them.

It has been said before that a thesis cannot be written if one does not have good supervision. The supervision that I received has always been extremely helpful, and the comments from my supervisors have taught me a lot, especially about what a scientific text should and should not be.

Iris Mulders and Vic Schepens have often listened to — or read — long rantings of mine about what I was trying to do and the dead alleys I was running into. Their comments and probing questions have often revealed flaws in my reasonings and lead to solutions to problems I was facing. Many other people have helped me in the same way, although less extensively. I would like to thank them all.

I have also appreciated it very much that Eric Reuland has been willing to co-supervise this thesis, and to take time out of his busy schedule on a regular basis to meet with me and with the other supervisors.

Arabic is not my native language, which means that the help of native speakers has been essential. I would like to thank Rabha Heinen, Achmed Khabbazeh and Mohammed el-Sharkawy in particular for helping me with judgements and comments. I have also used data from the corpus of Arabic newspaper and literary texts collected at the University of Nijmegen. I would like to thank Jan Hoogland for allowing me to use this corpus and for helping me to find in it what I needed to know.

The practical side of writing the documents that make up the dissertation has been made much easier and more satisfying by the use of LaTeX. I must thank Iris Mulders for introducing me to this wonderful system, and thereby saving me endless frustration.

Lastly, a word of thanks must also go to Wolfgang Mederle for providing me with a good English translation of a particularly intricate piece of German scientific writing, and to the various posters in the alt.os.linux.slackware newsgroup, who have helped me with several questions regarding English style and usage.

Introduction

The subject of this thesis is the Arabic noun phrase. In the past decade, Chomsky (1995, 1998, 1999) has initiated a new development in syntactic research, commonly known as *minimalism*. The central idea of this development is that linguistic theory should make use of as few primitive notions as possible. For this reason, standard notions such as X-bar theory and government are abandoned and replaced with more basic notions.

Chomsky does not develop a fully worked-out syntactic theory. Rather, he sets out the basic principles along which such a theory should be developed. In this thesis, I will develop a syntactic theory for the Arabic noun phrase that is based on the minimalist principles proposed by Chomsky. I will give an account for several of the most common phenomena that are known from Arabic noun phrases, such as the genitive construction, word formation, placement of adjectives and other modifiers, adjectival agreement and to a lesser extent the formation of deverbal nouns and participles.

Because Chomsky's *minimalist program* is not intended to be a fully-fledged syntactic theory, but rather a basis for one, there are important parts of it that have not been worked out. One such part is linearisation: the derivation of a linear ordering from a hierarchical tree structure. Since Kayne (1994), the idea that Universal Grammar (UG) specifies a universal Specifier-Head-Complement ordering has become more or less the standard view in generative syntax. However, Kayne's theory is not 'minimalist' in the sense of Chomsky (1995), and the claim that UG specifies a fixed order, although intuitively attractive, is not always supported by the facts.

Furthermore, Chomsky (1995) claims that core syntactic structures are hierarchical and do not have a linear ordering. The reason for this claim is that linear ordering is not relevant for the semantics of a phrase. The semantics are computed on the basis of the hierarchical relations in a structure. The linear ordering is only relevant to the computation of the phonological representation of a phrase. This basically means that UG does *not* specify a fixed order. Instead, a linear ordering is computed during the derivation of the phonological representation.

For these reasons, I develop a linearisation procedure in chapter 2 which is minimalist in the sense that it does not make use of any additional stipulated mechanisms. In chapter 3 I discuss the basic structure of the Arabic noun phrase, and I show how the

linearisation procedure that I develop accounts for the linear ordering of the examples under discussion.

In chapter 4, I look at the adjective phrase in Arabic. There are several interesting phenomena going on in the Arabic adjective phrase, which show that adjective phrases have a clause-like structure. In chapter 5, I look at nouns that are derived from verbs, to see how we can account for them. At the end of both chapters, I will return to the linearisation procedure, to show how it handles the structures presented in the chapters.

Sources of the examples

The Arabic examples in this thesis were collected from a variety of sources. The more straightforward examples come from grammar books. Other examples were taken from two descriptive grammars of Arabic: *Syntax of Modern Arabic Prose* (Cantarino 1975), which I abbreviate as 'Cant.', and *Syntax der Arabischen Schriftsprache der Gegenwart* (El-Ayoubi, Fischer & Langer 2001), which I abbreviate as 'SASG'. Some further examples were taken from the corpus of Modern Arabic texts compiled at the University of Nijmegen and from various books and newspaper articles.

A note on transcription and glosses

Before I begin, let me discuss the system that I use to transcribe the Arabic examples, and how the glosses are composed. Table 1 lists the transcription that I use and the pronunciation of each chacacter in IPA. Note that the pronunciation is only approximate, especially in the case of the vowels. For a full description of the pronunciation of Arabic, see Mitchell (1990).

Strictly speaking, Arabic words never begin with a vowel. An initial vowel is always preceded by a glottal stop, the so-called *hamzah*. Such an initial glottal stop can be 'connective' (*hamzat al-waṣl*) or 'disjunctive' (*hamzat al-qaṭ^{c}*). A disjunctive glottal stop is always retained. A connective glottal stop is dropped, together with its following vowel, if the preceding word ends in a vowel.

I transcribe the disjunctive glottal stop with an apostrophe «'». I do not transcribe the connective glottal stop, I only transcribe the following vowel (which is a certain indication of its existence). When the connective glottal stop is dropped, I replace the vowel with a hyphen. Thus: *ibn-ī* 'my son', but *li -bn-ī* 'to my son'.

The letter ẓ is officially pronounced [ð^{ʕ}], but in not too formal contexts, it is often [z^{ʕ}], hence the convention of transcribing it with ẓ.

The feminine ending is indicated in Arabic script with a special letter, the so-called *tā' marbūṭa*. Its pronunciation is /t/, but when the word appears at the end of a clause or just before a pause, it is dropped, together with the following case ending. In the gloss, I always transcribe this letter with -t-, but when I quote an Arabic word in the text, I often omit it. I do the same with short vowels at the end of the word, which are most often case endings: in the glosses, they are written, in the text they are often omitted.

transcription	IPA	transcription	IPA
’	ʔ	f	f
b	b	q	q
t	t	k	k
$\underline{t}$	θ	l	l
ǧ	ʒ	m	m
ḥ	ħ	n	n
$\underline{h}$	χ	h	h
d	d	w	w
$\underline{d}$	ð	y	j
r	r		
z	z		
s	s	a	a/ɑ
š	ʃ	ā	a:/ɑ:
ṣ	$s^ʕ$	u	ʊ
ḍ	$d^ʕ$	ū	u:
ṭ	$t^ʕ$	i	ɪ
$\underline{z}$	$z^ʕ$	ī	i:
ʿ	ʕ	aw	au/ɑu
ġ	ʁ	ay	ai/ɑi

Table 1: Transcription and pronunciation

Consonants in Arabic can be lengthened, which is usually indicated by doubling them in the transcription. Thus: *kataba* /kataba/ ‘he wrote’, and *kattaba* /kat:aba/ ‘he made s.o. write’. I use this method, too.

Arabic has a number of particles that consist of one consonant plus a short vowel. Although they are clitics and are always attached to the following word in Arabic script, I detach them in the transcription to avoid confusion between these particles and affixes. Thus *wa huwa* ‘and he’, rather than *wa-huwa*.

The definite article in Arabic is *al-*, cliticised onto the noun (and adjective). The /l/ of the article assimilates to the first consonant of the noun if this consonant is alveolair. In the transcription, however, I do not indicate this assimilation, so I always write *al-*.

In the gloss, I give a morph-to-morph translation of the Arabic example. In principle, every affix is separated with a hyphen, although I sometimes leave out the hyphen if the particular morpheme is not relevant to the example at hand. Each morpheme is glossed separately with an English translation or with an abbreviation indicating its function. For instance, I gloss the definite article *al-* with *the-*, but the case endings are glossed with NOM, GEN and ACC.

As in the transcription, affixes in the gloss are attached to the main word with a hyphen. Sometimes, however, I omit some morphemes from the gloss, usually when they are not important for the point at hand. So I write *kitāb-u-n* ‘book-NOM-INDEF’, but on occasion, I use *kitāb-un* ‘book-NOM’. I frequently do this with the feminine

ending: *sayyārat-un* 'car-NOM' rather than *sayyār-at-un* 'car-F-NOM'.[1]

Sometimes a morpheme cannot be properly separated from the main word. When this happens, and when the function of the morpheme is important, I write the morpheme after the main word, and connect the two with a dot. E.g., *al-miyāh-u* 'the-water.PL-NOM'.[2] I also use this method when a single Arabic word must be glossed with more than one English word: *yaktubu* 'he.writes', or *yaḫruǧu* 'he.goes.out'.[3]

[1]Note that when a feminine word is cited in the text, the *-t* of the feminine ending is dropped, along with the case ending, so I write *sayyāra* 'car'.

[2]Mass nouns in Arabic can sometimes have a plural form, such as *mā'* 'water', pl. *miyāh* in this example. This indicates large quantities of the substance.

[3]Generally, the glosses of verbs are *ad hoc*. For example, the tense on the verb in the gloss reflects the proper translation, rather than the Arabic tense, and aspect and mood are not indicated.

1

The minimalist framework

The framework that I will use in my analysis is that developed in Chomsky's recent work (1995, 1998, 1999). This framework is generally known as the *minimalist program*. As this term suggests, it is not a fully worked-out syntactic theory, but rather a program for developing such a theory. The central idea of minimalism is that a syntactic theory should make use of as few principal notions as possible. In the *Government and Binding* (G&B) theories that were developed on the basis of works such as Chomsky (1981, 1986a), many notions were proposed, such as the *X-bar schema, government, s-structure, d-structure, LF, PF, (relativized) minimality, ECP*, etc. The goal of the minimalist program is to try and reduce these notions to a small set of more basic principles.

1.1 Principles and parameters

The minimalist program is a development within the framework of Principles and Parameters (P&P). The central notion of the P&P framework is that the human language faculty can be described with a set of principles together with a set of parameters. The principles express fundamental properties of the language faculty and the parameters express the cross-linguistic variability of language.

One of the principles that governs syntactic research is the idea that grammatical structures are hierarchical: even though a phrase such as *John saw a cat* contains four words that are linearly ordered, the phrase contains a hierarchical structure. It is not just a string of four words, but a structure in which some elements form a unit, which is then combined with other units. The structure of this particular phrase is roughly [John [saw [a cat]]]. That is, the elements *a* and *cat* are combined to form *a cat*, which is then combined with *saw* to form *saw a cat*, etc.

Another important principle in syntactic theory is the *structure dependence principle*, which states that syntactic operations are structure-dependent. This means that syntax operates on (hierarchical) structures. For example, inversion is an operation

that applies to a subject noun phrase and an auxiliary in an interrogative clause:

(1) a. John will go to France
b. will John go to France?

The examples in (1) could lead one to conclude incorrectly that inversion can be described as "Inversion reverses the first two words of a clause". This obviously leads to the wrong results in (2):

(2) a. our uncle will go to France
a′. *uncle our will go to France?
b. John read the book
b′. *read John the book?

These examples show that inversion can only be defined in structural terms: it is not the first two words of the clause that swap places, it is the subject noun phrase, which can consist of more than one word, and the auxiliary verb, that swap places. Therefore, in order to describe the phenomenon of inversion, one needs to make use of a structural description of the phrase.[1]

These principles are principles of the human language faculty, which means that all languages should adhere to them. At the same time, however, we know that languages show great variabitility. Phenomena that appear in one language do not necessarily appear in the same way in another language:

(3) a. (+he) left for work[2]
b. (él) se ha ido al trabajo
(he) SE has gone to.the work
'he has gone to work'

In English, the subject of a clause must be overtly present, as demonstrated in (3a). In other languages, however, a subject pronoun can be dropped. In the Spanish example in (3b), *él* can be present but does not have to be. This variability is captured with parameters. In order to account for the fact that languages can differ in their ability to drop subject pronouns, the so-called *pro-drop* parameter is posited. The setting of this parameter is different in English and Spanish, accounting for the observed variability.

Another example of variability is shown in (4):

(4) a. what did you see ~~what~~?[3]
b. suft^{i} 'ēh?
you.saw what
'what did you see?'

[1]It should be noted that inversion is not an operation of and by itself. Rather, it is the visible effect of several other operations.

[2]I use the (+...) notation to indicate obligatory inclusion. In my opinion, the contrast between (*...) and (+...) for obligatory exclusion and obligatory inclusion respectively, is less confusing than the more common notation of (*...) and *(...).

[3]Note that I indicate movement by striking out the source position of the moved element.

The question word *what* in the English example in (4a) has moved to the first position of the clause. This is the normal position for so-called *wh*-words in English to appear in. Many other languages show the same phenomenon: *wh*-words are moved to the front of the clause. There are languages that do not do this, however. One such example is Colloquial Cairene Arabic, which is shown in (4b). The *wh*-word *'ēh* 'what' does not move to the front of the clause, but stays *in situ*. This variability is described with the so-called *wh*-parameter. The setting of this parameter determines whether or not *wh*-words move to clause-initial position.[4]

1.2 The grammar model

In a Chomskian approach, a linguistic structure has two structural representations: LF and PF. LF is the so-called *logical form*, which is a representation from which the semantic value of the phrase is computed. The LF representation is used by the syntactic system in its interaction with the semantic system: the syntactic system creates a syntactic structure and forms an LF of it, which is then passed on to the semantic system so that it can compute the semantics of the phrase.

The other representation that is created by the syntactic system is PF, the *phonological form*. This representation is sent to the phonological system, where it will be 'spelled out'. Spelling out a structure basically means mapping the hierarchical structure onto a linear structure that contains only the features needed for further phonological processing of the phrase.[5] Both representations are derived through repeated application of the two basic operations of the syntactic system: Merge and Agree, which will be discussed in the following sections. Chomsky argues that the derivation of a clause starts with a so-called numeration, which is an unordered set of all the lexical elements that will eventually appear in the phrase. This numeration is formed by selecting the required lexical elements from the lexicon. One by one, the lexical elements are then taken from the numeration and placed in the tree structure that is being built. At a certain point, the derivation splits up into an LF-derivation and a PF-derivation. One can schematize the grammatical model as in (5):

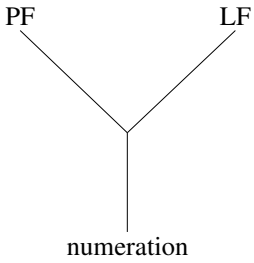
[4]It should be noted that there are more factors that influence the positioning of *wh*-words. E.g. echo questions in English usually have the *wh*-word *in situ*:

(i) A: I bought a dog yesterday!
B: You bought *what*?

In neutral cases, where other factors do not play a role, the *wh*-parameter is very robust.

[5]That is, spelling out does *not* mean the actual pronouncing of a phrase. "Further phonological processing" can be pronunciation but also perception.

(5)

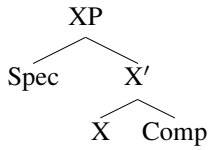


A derivation that reaches LF without violating any principles is said to *converge* at LF. If a derivation does not converge, it is said to *crash*. A derivation that crashes is not a well-formed linguistic structure, i.e. it is ungrammatical.

1.3 Minimalist tree structures and *Merge*

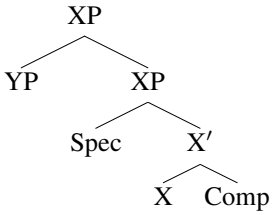
As mentioned before, syntactic structures are thought of as hierarchical structures. They can be represented in tree notation. Originally, trees were very free, e.g. a node in a tree could have any number of branches, but at a certain point the idea developed that trees were restricted in various ways. For example, since Kayne (1984) it is commonly assumed that a tree can only be binary-branching, i.e. a node can only have two subnodes, not more. In G&B theory, the so-called X-bar schema was developed, which stated that a phrase must be represented as in (6):

(6)



The X-bar schema places several restrictions on a tree structure: each phrase must have one and only one head (X in (6)); a tree must be binary-braching; a head combines with a complement to form X′,[6] which in turn can combine with a specifier to form the fully-projected phrase XP. With the X-bar schema, we have two levels of projection: X′ and XP, the maximal projection. An XP can be merged with another maximal projection, say YP, in a configuration known as adjunction:

(7)



[6]Note that X′ is pronounced 'X-bar' here, not 'X-prime'.

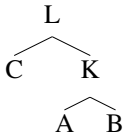
Adjunction is optional, and it does not create a new level of projection. Instead, the same level XP is maintained, as indicated by the fact that there are two XP nodes in (7).

In the minimalist program, however, Chomsky (1995) abandons the X-bar schema and replaces it with the notion *Merge*. Merge is one of the two basic operations of the syntactic system. It is the operation that forms trees, which it does by taking two elements A and B and putting them together to form a (more) complex structure K(A,B). One of the elements A and B passes its label on to K. I will call this node the projecting node.[7] That is, once A and B have been merged, we have the structure in (8):

(8)

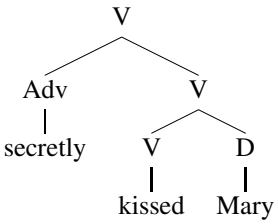
This structure can then be used in a further application of Merge. For example, if we merge K with an element C, we get (9):

(9)



This is all that Merge does. One consequence of this is that there is no principled way to distinguish different levels of projection, as was the case with the X-bar schema. There are ways to obtain these distinctions, but in principle, a head X can be merged with any number of elements, resulting in multiple specifier constructions. Furthermore, the syntactic structure does not distinguish specifiers from adjuncts. To see why this is the case, consider the following structure:

(10)



In (10), the head V is first merged with a D, which expresses the object. The resulting structure has the label of the projecting node, V. Next, it is merged with an adverb, a merger that is traditionally termed adjunction. The resulting structure is again a V.

The structures that are formed by Merge are not X-bar structures in the traditional way. In a more traditional analysis, the merger of *kissed*$_{\mathrm{V}}$ and *Mary*$_{\mathrm{D}}$ would result in the formation of a VP: a projection of V that contains all the internal arguments of

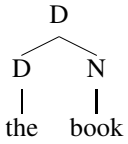
[7] Chomsky calls this node *the head*, but that term has always been used for a slightly different notion: traditionally, a head is a terminal element. I will continue to use *head* in this sense.

the V. The further merger of *secretly*$_{\text{Adv}}$ would result in an adjunction structure such as [$_{\text{VP}}$ secretly [$_{\text{VP}}$ kissed Mary]], where *secretly* is adjoined to the VP, forming a doubly-layered VP, like the doubly-layered XP in (7).

In a minimalist approach, however, each node receives the label of its projecting subnode, which in (10) is V at every level.[8] The structure in (10) is a V, as much as [$_{\text{V}}$ kissed [$_{\text{D}}$ Mary]] is a V. Both phrases have the same combinatorial properties. They can both be taken as a complement by v, for example. Similarly, a bare verb *kissed* is a V. There is of course a difference between the V *kissed* and the V *kissed Mary*: the former still has an unassigned theta role, whereas the second does not.[9]

Chomsky calls this *bare phrase structure*. In a bare phrase structure approach, there are no projection levels anymore. One consequence of this is that a maximal projection is simply the highest projection of a specific category. This means that a node can be a head and a maximal projection *at the same time*. Take the following example:

(11)



In (11), the node [$_{\text{N}}$ book] is a head (i.e. a terminal element) but at the same time it is a maximal projection, because it is the highest projection of category N.

1.4 *Agree* and the feature system

The second basic operation of the syntactic system is *Agree*. Agree establishes a relation between two elements on the basis of feature-match. Lexical and functional elements are thought of as bundles of features. A word such as *man* has grammatical features such as [+N], [+SG], [+MASC], etc. and semantic features such as [+ANIMATE], [+HUMAN], [+MALE], etc. Functional elements only have grammatical features, e.g. the tense projection T has a feature [±TENSE].[10] Agree can establish a relation between two elements if they share certain grammatical features.

For example, many languages show subject agreement on the verb. This means that there must be some operation that establishes this relation between the subject and the verb. In the original implementation, Chomsky (1995) made a distinction between interpretable and uninterpretable features: φ-features (that is, features for person, number and gender) on a noun or pronoun are interpretable because they have

[8] The label VP may still be used at times, but then it is purely a mnemonic device for the highest projection that is still of category V.

[9] In (10), I have indicated the labels of each node with a *categorial* label. But it should be noted that Chomsky prefers an even more minimalist system in which nodes no longer have a categorial label. Instead, the label of a node consists of a representation of the structure of that node. Because the matter is not relevant to the present study, I will not go into it here, and simply label nodes with their category.

[10] Elements can also have phonological features. Lexical elements always do, functional elements can, but do not have to.

a role to play in the semantics of the noun or pronoun. (E.g., a pronoun with the features [3M,SG] refers to a different element than a pronoun with the features [3F,PL]). The same features on the finite verb, however, are uninterpretable, because they have no meaning there.

The idea was that uninterpretable features, because they have no semantic value, must be erased during the derivation. Erasing features can be done by establishing an agree relation with an element that has the same features. Therefore, a finite verb, having a set of uninterpretable φ-features, will try to agree with a noun phrase (the subject) which also has a set of φ-features. If the agree relation is successfully established, the uninterpretable features on the verb are erased. If for some reason, the uninterpretable features cannot be erased, the derivation crashes.

Originally, (c.f. Chomsky 1995), features were binary: an element would either have a certain feature or it would not. In *Derivation by Phase* (1999), however, Chomsky seems to move towards a valued feature system, although the system is not fully worked out.

The valued feature system seems to be proposed to overcome a basic problem with the notion of (LF-)interpretability of features. LF-interpretability of features is a semantic notion: a feature is interpretable if it has semantic content. The problem is that the syntactic system, in order for it to delete all uninterpretable features, must be able to determine which features are interpretable and which are not. That, however, requires a form of look-ahead: the syntactic system must "look ahead" in the derivation to the semantic system in order to determine whether a feature is interpretable or not. Ideally, however, the syntactic system does not need to access the semantic system in order to do its work. The two systems are separate modules, and only communicate through the LF representation that the syntactic system forms. The syntactic system does not have access to the semantic system to determine whether a certain feature is uninterpretable, which means that it cannot determine whether it is to erase a specific feature or not.

To solve this problem, Chomsky proposes a form of valued feature system. In such a system, a feature is an item such as DEFINITENESS, CASE, TENSE etc. Every feature can take different values. For example, the feature DEFINITENESS can take the values DEFINITE, INDEFINITE. CASE can take the features NOMINATIVE, GENITIVE, ACCUSATIVE, etc.[11] The idea is that a lexical item can enter the derivation with some features still *unvalued*; i.e., certain features have not been given a value yet. In order for the derivation to converge, the unvalued features must be valued during the derivation.

Chomsky proposes that a head can enter into an agree-relation if it is *active*, that is, if it has an unvalued feature. An active head will try and value its unvalued features by *probing* for a *goal*. That is, it will try and find an element that has the same features, but in which these features are already valued. If such a match is found, the active

[11] I will write both the features and their values in small caps. In order to distinguish between a feature and its value(s), I will write the former with a large initial capital, while the latter are all small caps. A valued feature will be written as [CASE: NOM]. I will not do this with features that have only binary values. Instead, I will write [+TENSE] and [–TENSE]. These notations can be considered short-hand for [TENSE: FINITE] and [TENSE: INFINITIVAL].

features of the head will be valued with the values of the matching features on the goal. Chomsky says that features thus valued can be deleted. This last addition, that a feature valued in the derivation can be deleted, essentially means that the features that are unvalued at the onset of the derivation are the uninterpretable features.

In this way, the problem of look-ahead disappears. The deviration no longer has to determine which features are uninterpretable. Instead, it can simply check if a feature is valued or not, and if it is not, it must try and value it by finding a matching element in the c-command domain of the head.

Note that Chomsky argues that a head can only establish an agree relation with a matching goal if the goal is active itself, i.e., if it has unvalued features. These unvalued features will be valued by the head in the same agree relation, if the head has the same features. This means that an agree relation is generally a two-way relation. The typical example is again the agreement between subject and verb. Not only does the subject value the φ-features of the verb, the verb also assigns (nominative) case to the subject. The idea is that the subject enters the derivation with a valued set of φ-features and with an unvalued CASE feature, whereas the verb enters the derivation with a set of unvalued φ-features and with a valued CASE feature.[12]

If the goal is not active, the agree relation fails. Note, however, that the probe cannot try and find another goal lower in the tree. This is called the *defective intervention effect*: α cannot agree with β if there is an intervening γ that matches with α. As a result, the features of the probe remain unvalued and the derivation crashes.

The operation Agree can optionally be followed by the operation Merge. If this happens, the goal moves and merges with the probe. This usually takes place in subject-verb agreement. Subject-verb agreement is established by the head T, which carries the TENSE feature. The subject is generated as an argument of the verb.[13] T, being active because of the set of unvalued φ-features, will probe for a goal and finds the subject. This establishes the agree relation and values the features on either side. After this, the subject merges with T, ending up in spec,TP.

1.5 Phases

It has long been known that there are domains that are impervious to certain syntactic operations. For example, although movement is a very common operation, there are certain domains that do not allow the movement of any element out of them. Consider the *wh*-movement we already saw above:

(12) what did you see ~~what~~?

[12] In fact, Pesetsky & Torrego (2001) argue that the CASE feature on the verb is really the TENSE feature. This means that nominative case is really a reflex of TENSE on the subject. This account makes sense if one considers the fact that (structural) case is not interpretable semantically on either the verb or the noun. Since the minimalist program argues that there can be no syntactic features that are not semantically interpretable, CASE is problematic. TENSE is interpretable on the verb, which means that if CASE were to reduce to tense, the problem would disappear.

[13] More specifically, in spec,vP, unless the verb is unaccusative, in which case the subject is generated in the VP.

In (12), the object, the interrogative element *what*, has moved from the canonical object position immediately after the verb to clause-initial position. This means that movement of the object to clause-initial position is grammatical under certain conditions. What is often not possible is to move an element that is *part* of the object:

(13) *whose did you see [~~whose~~ dog]?

The intended reading of (13) is "Whose dog did you see?", but the phrase is ungrammatical. It is not possible to move an element out of the noun phrase *whose dog*.

Another example is the impossibility to move a *wh*-word out of an adjunct clause:

(14) *he is the person to whom they left before speaking ~~to whom~~

In (14), the relative marker *to whom* is extracted from the complement position of the verb *speaking*.[14] This movement is not possible, which leads to the conclusion that an adjunct clause, such as *before speaking to whom*, is impervious to extraction, or opaque, as it is sometimes called.

Phenomena such as these make it clear that there are certain barriers in linguistic structures that operations cannot cross. In the current theory, these barriers are implemented as phases. A phase is a piece of structure that is closed off for further operations.

In the original definiton in Chomsky (1998), phases are defined propositionally: a phase is the closest syntactic counterpart to a proposition. This basically comes down to saying that CP and vP are phases, but TP and VP are not. A VP cannot be the syntactic counterpart of a proposition, because it lacks a subject. It is vP that introduces this subject. TP is likewise not sufficient, because it is part of a clause, but essential elements of the clause, i.e. the force markers, topic and focus markers, are positioned in the CP domain.[15]

It is now also generally assumed that DP is a phase as well, even though it does not appear to contain a full proposition. If, however, we take seriously the idea that phases can be fronted and extraposed, we must conclude that DPs are indeed phases, because they can undergo these operations. Following the general consensus, I will assume that DPs count as phases as well.

With the inclusion of DP, the notion "phase" becomes very close to what Grimshaw (1991) calls "maximal extended projection": the highest projection of a lexical category. The lexical categories are V, N and A, and the maximal extended projections are the categories to which they project. In the case of V, this is CP, for N this is DP and

[14]The intended meaning of (14) is "he is the person that they did not speak to before they left". The relative clause in (14) is built on the structure in (i):

(i) they left before speaking to him

(i) is grammatical. What (14) attempts to do is to extract the complement of the verb *speaking*, which turns out to be impossible.

[15]Chomsky further notes that there are considerations on what he calls the "sound side" that support this idea of phases: vPs and CPs can be fronted, extraposed etc. and can serve as response fragments. TPs and VPs cannot.

for A it is DegP.[16] The two notions are not exactly the same, however, because vP is considered a phase, but it is not a maximal extended projection.

Whatever the exact definition is, a phase is a piece of syntactic structure that in some way operates as a whole: it cannot be split up, synctactic operations that affect the phase can only affect the phase as a whole, not parts of it.[17]

Chomsky (1999) develops this idea. He argues that there is a rationale behind the fact that phases are independent pieces of syntactic structure: the derivation of a syntactic structure takes place phase by phase. Specifically, this means that the syntactic system builds a phase and sends it to PF to be spelled out. Once it has been spelled out, it can be merged into another syntactic structure, but because it has already been spelled out, it has been stripped of its syntactic information. As a result, the syntactic system cannot "look" into it anymore, and it must necessarily treat it as a whole.

As an example, suppose the syntactic system is deriving a phrase such as in (15):

(15) John said to Mary that he was going to leave her

The subclause *that he was going to leave her* is a CP, and hence a phase. As such, Chomsky argues, it is built separately from the matrix clause, and it is spelled out separately. Then, when the matrix clause is being built, it is the result of the spell-out of the embedded CP that is merged into the matrix structure.

[16]Whether P should be seen as a lexical category is a difficult question. It generally seems to be an intermediate category. For example, it is a not a closed class like C or D, but on the other hand it is also not as open as N or V: new prepositions are sometimes introduced, but not with the same ease or frequency with which new nouns and verbs are introduced.

[17]This is in fact not entirely true: it is generally accepted that the so-called edge of a phase *is* still accessible to syntactic operations from the outside. The edge of a phase is the outer layer of specifiers plus the highest head. That is, the edge of the CP phase is formed by the head C and all elements in spec,CP. The edge of the phase functions as a sort of "escape hatch" through which elements can escape.

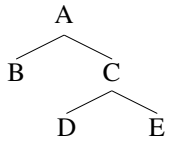
2

Linearisation

2.1 Tree structures

In syntactic theory, syntactic structures are represented as trees. Trees traditionally define two different types of relations: they define hierarchy relations, in the sense that one node can dominate another, and they define ordering relations, in the sense that one node can precede another. For example, take the tree in (1):

(1)



Dominance relations exist between A and all other nodes, and between C and D, and between C and E. Precedence relations are defined between nodes B and C, and between D and E. As one can see, two specific nodes are either in a dominance relation, or in a precedence relation. They cannot be in both relations at the same time.

A third relation between nodes in a tree has been developed by Reinhart (1976), (further discussed by Chomsky 1981 and Aoun & Sportiche 1983) which is the relation of *c-command*. Formally, it is defined as follows:[1]

(2) α c-commands β iff α does not dominate β and every γ that dominates α dominates β
(Chomsky 1986a, p. 8)

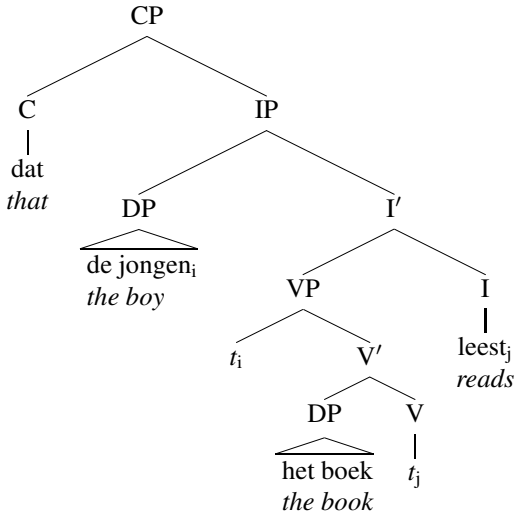
What this definition comes down to is that every node c-commands its sister node and the nodes dominated by its sister. So in (1), B c-commands C, D and E, C c-commands B, and D and E c-command each other. A does not c-command any node.

[1] I give the basic definition here. There are several extended versions of the definition, but they all have this definition at their core.

With tree structures that define ordering relations between sister nodes, linearisation is trivial. For example, if we look at the variation between VO languages and OV languages, the traditional way of G&B theories of the eighties to account for this variation was to assume that in OV languages such as Dutch and Japanese, the complement of V was on the left branch of the V′ node, rather than on the right branch, as it was for English:

(3) a. . . . dat de jongen het boek leest
. . . that the boy the book reads
'. . . that the boy reads the book'

b.



In this Dutch example, both I and V are head-final, resulting in the VP coming before I and the object coming before V.

So we see that tree structures define two basic relations, dominance and precedence, and a derived relation, c-command. The linear ordering of the terminal elements in the tree is determined by the dominance and the precedence relations of the elements in the tree. At a certain point, however, the question was raised whether it would be possible to do away with either of the two basic notions, and derive a linear ordering solely with one of the two. Kayne (1994) argues that this is indeed possible. He claims that one can derive linear ordering by using the relation c-command, which is defined in terms of dominance, but which does not make use of the notion precedence as a primitive.

2.2 Kayne's antisymmetry

Kayne starts out by noting that the linear ordering relation of a given set of terminal elements has three properties.[2] It is transitive, i.e., if X $>$ Y and Y $>$ Z then X $>$ Z; it

[2]It should be noted that the linear ordering that Kayne is talking about here is the linear ordering of a set of end nodes, or terminals, which is not the same as the precedence relation that is defined between two

is total, i.e., it covers all members of a given set of elements; and it is antisymmetric, i.e., if X > Y then it is not possible that Y > X.[3]

Next, Kayne examines the relations *dominance* and *c-command*. He notes that dominance is transitive and antisymmetric, but not total. Dominance is only total in a limited domain: the domain of all nodes dominating a specific node. The dominance relation, therefore, is *locally total*, Kayne claims.

When examining the c-command relation, Kayne notes that it is only transitive: if a node X c-commands a node Y, and Y c-commands Z, then X will also c-command Z. The relation is not antisymmetric, because two nodes can c-command each other (e.g., in the case of two sister nodes.) But one can define the notion of asymmetric c-command:

(4) X asymmetrically c-commands Y iff X c-commands Y and Y does not c-command X
(Kayne 1994, p. 4)

The new relation of asymmetric c-command is transitive and antisymmetric, but it is not total, because in a given tree there can be two nodes neither of which (asymmetrically) c-commands the other. However, the relation is *locally* total if trees are restricted to binary branching nodes. As Kayne states, "(...) in a binary branching tree, if Y asymmetrically c-commands X and Z (distinct from Y) also asymmetrically c-commands X, then it must be the case that either Y asymmetrically c-commands Z or Z asymmetrically c-commands Y." (Kayne 1994, p. 4–5). That is, just like the dominance relation, which is total in the set of all nodes dominating a specific node X, the asymmetric c-command relation is total in the set of all nodes asymmetrically c-commanding a specific node X.[4]

With these steps, Kayne has reached a situation in which he has obtained three relations which share the same properties: there is a relation of linear precedence, which is defined over a set of terminal elements, and there are the relations of dominance and asymmetric c-command, which are defined over non-terminal nodes in a tree. Linear precedence is transitive, total and antisymmetric, and dominance and asymmetric c-command are transitive, *locally* total and antisymmetric. Kayne then states: "The intuition that I would like to pursue is that there should be a very close match between the linear ordering relation on the set of terminals and some comparable relation on non-terminals." (Kayne 1994, p. 5). Of the two relations under consideration, "... it is natural to take asymmetric c-command to be the one that is closely matched to the linear ordering of the set of terminals" (*ibid.*).[5]

sister nodes.

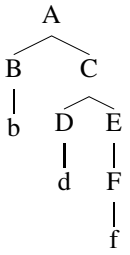
[3] Strictly speaking, the property that Kayne describes here is "asymmetry". A relation is antisymmetric if for all domains $\forall x \forall y((xRy \wedge yRx) \rightarrow x = y)$, which is a less strict property: a relation that is asymmetric will also be antisymmetric, but a relation that is antisymmetric is not necessarily asymmetric.

[4] This is not the case in trees that allow more than two branches: two sister nodes of a ternary branch that both asymmetrically c-command a node X in the third branch c-command each other. In other words: there is no asymmetric c-command relation defined over them.

[5] Note that there is actually no real reason why the dominance relation should be made to be similar to the linear ordering relation. It will become clear later on why Kayne takes this step.

Let us look at a sample tree structure, to see how the asymmetric c-command relation operates:

(5)



As Kayne states, asymmetric c-command is defined over the non-terminal nodes,[6] which are the capital nodes in (5). So B asymmetrically c-commands D, E and F, and D asymmetrically c-commands F. B and C c-command each other, and therefore neither B nor C *asymmetrically* c-commands the other. The same is true for D and E.

Kayne now formulates the *Linear Correspondence Axiom* or LCA, which basically states that the set of asymmetric c-command relations (defined on non-terminals) corresponds to the set of precedence relations (which is defined on terminals). Each non-terminal node dominates one or more terminal nodes. Kayne's LCA states that if an ordered pair <X,Y> is part of the (maximal) set of asymmetric c-command relations, the terminal nodes that X and Y dominate should be ordered accordingly. For example, in (5), the ordered pair <B,D> is part of the set of asymmetric c-command relations. The terminal nodes that B and D dominate, b and d, should therefore be ordered in the same way. In other words, the ordered pair <b,d> should be in the set of linear ordering relations defined over (5).

As stated, the set of asymmetric c-command relations in (5) is {<B,D>, <B,E>, <B,F>, <D,F>}. This means that the set of linear ordering relations is {<b,d>, <b,f>, <b,f>, <d,f>}.[7] This set of ordered pairs of terminals gives a linear ordering, which in this case would be <b,d,f>.

In order to allow for specifiers, Kayne has to make a change to the standard definition of c-command:

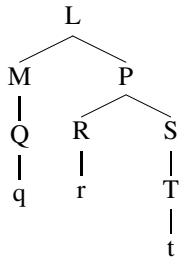
(6) X c-commands Y iff *X and Y are categories* and X excludes Y and every category that dominates X dominates Y
(Kayne 1994, p. 16)

[6] Kayne means by this that the only asymmetric c-command relations that are relevant are those between non-terminals, not those that involve terminal elements.

[7] The pair <b,f> is in the set twice, once because of <B,E> and once because of <B,F>. This is not problematic.

To see why this change must be made, consider the following tree:

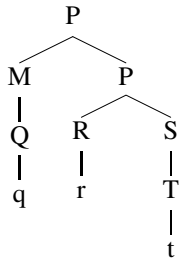
(7)



In (7), M asymmetrically c-commands R, which means that the pair <q,r> is in the set of linear ordering relations. But here P also asymmetrically c-commands Q, which means that the pair <r,q> is also in the set of linear ordering relations. This is contradictory: it is not possible for <r,q> and <q,r> to be in the set of linear ordering relations of a given tree.

If the tree is modified to the following, this problem disappears, given the definition of c-command in (6):

(8)



Here, the top node L is replaced with P. P is now a category consisting of two *segments*. In this tree, M still asymmetrically c-commands R, because every category that dominates M (i.e. P) dominates R. However, P no longer asymmetrically c-commands Q, because the lower P is not a category, but merely a segment of a category. The pair <r,q> is therefore no longer part of the set of linear ordering relations, and the contradiction is eliminated.

As Kayne points out, this discussion leads to the conclusion that specifiers and adjuncts cannot be distinguished. The tree in (7) shows that under the LCA a structure in which a non-head is the sister of another non-head (M and P in (7)) cannot exist because such a configuration cannot be linearised. The only way in which two non-heads can be sisters is in the configuration in (8), where M is adjoined to P. This means that specifiers must be adjoined, or in other words, specifiers are adjuncts.

2.3 Problems with Kayne's theory

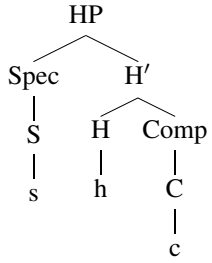
The theory that Kayne proposes has two major problems. One is conceptual in nature, the other is of a more practical kind. The conceptual problem lies in the fact that

although the theory defines a linear ordering of terminal elements, it does not define whether this ordering is precedence or subsequence. The practical problem is encountered when one tries to explain word order variation across languages with a syntactic theory based on this approach.

2.3.1 Precedence and subsequence

As Kayne points out in his chapter 4, the LCA maps asymmetric c-command relations onto linear ordering. More specifically, in a tree such as the one in (9), the LCA forces the specifier and the complement to appear on different sides of the head:

(9)



In (9), the node Spec asymmetrically c-commands H, meaning that the ordered pair <s,h> is in the set of linear ordering relations of (9). Furthermore, H asymmetrically c-commands C, which means that <h,c> is also in the set of linear ordering relations. Because linear order is transitive, we know that <s,c> is also in the set.

In other words, we get a linear order of s-h-c. But as Kayne himself points out, there is nothing in the theory so far that forces us to interpret the ordered pair <s,h> as "s precedes h". We may well interpret it as "s follows h". Which means that we can just as easily derive the order c-h-s. The only thing that follows from his theory is that in any given X-bar structure, the specifier and the complement must be on opposite sides of the head. It does not specify on which side each of them must be.

Obviously Kayne notes this shortcoming. His solution to the problem is far from satisfying, however. He states that given these two possibilities, a choice needs to be made. He argues that the choice for a universal Spec-Head-Comp ordering is much more plausible than for a universal Comp-Head-Spec ordering. The reason for preferring Spec-Head-Comp, he claims, is that Spec-Head orderings are generally much more common than Head-Spec orderings. In some categories, he states, there may be variation between Spec-Head and Head-Spec orderings, but there are categories in which there is a strong tendency toward Spec-Head orderings. He specifically mentions spec,CP and spec,IP.

Kayne's argument rests on the assumption that it is possible to observe an underlying universal order from the ordering relations that occur across languages. This is a problematic assumption, however. The two cases he mentions, spec,CP and spec,IP,[8] are probably the only cases where Spec-Head ordering is strongly predominant. Note

[8] Where we can take spec,CP to be a shorthand for other specifier positions to which fronted elements move, like Top and Focus.

that according to Kayne's theory, specifiers and adjuncts are treated the same way. Therefore, if we consider Spec-Head orderings, we must consider the position of adjuncts also. It is well known that left- and right-adjunction both occur. One example of adjuncts linearised after their heads will actually be discussed in this thesis: chapter 3 shows that specifiers/adjuncts[9] in the Arabic noun phrase consistently follow their heads.

Furthermore, if we wish to determine whether there is a universal ordering, we must also take the domain of Head-Comp orderings into account. Here, Kayne admits that there is no clear predominance of Head-Comp over Comp-Head. The fairly common occurrence of OV structures in languages is proof of this, as is the existence of languages with sentence-final complementisers and sentence markers, and languages with phrase-final determiners.

This means that the only elements that support the idea that there is a universal underlying order (and of Spec-Head-Comp being that universal order) are subjects and fronted elements.[10] The fact that subjects and fronted elements are predominantly clause-initial is something that must be explained, but Kayne generalises the linearisation properties of what is essentially a very specific set of specifier-head structures to all specifier-head structures and also to head-complement structures. However, given the fact that these types of specifiers have specific (semantic) properties that other specifiers and complements do not share,[11] we cannot be sure that this generalisation is a valid step to take. In other words, it is not at all certain that the behaviour of this group of specifiers is sufficient reason to assume that there is a universal underlying order, let alone that this order is Spec-Head-Comp.

Kayne nonetheless assumes a universal Spec-Head-Comp ordering, and he also tries to give a rationale for it. Let us examine his account, to see if it makes a stronger case than the empirical evidence. The argument goes as follows. Kayne states:

> Recall from chapter 1 that the asymmetric c-command relation is significantly similar to the dominance relation (both are locally linear). Associated with the dominance relation on phrase markers is a "root node" that has the property of dominating every node in the phrase marker (except itself). In the usual phrase marker, no node has the property of asymmetrically c-commanding every node except itself. I would like to propose bringing asymmetric c-command and dominance more into parallel by postulating an abstract node A for every phrase marker, with the property that A asymmetrically c-commands every other node. This abstract node should be thought of as being adjoined to the root node.
> (Kayne 1994, p. 36)

Obviously, it would be in some way attractive to make dominance and asymmetric c-command parallel in this way, but there is really no compelling reason, conceptual

[9] Like Kayne, I will argue that specifiers and adjuncts should be treated the same.

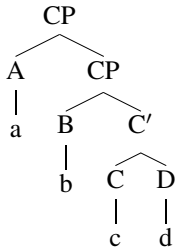
[10] For example, clausal orders in which the object precedes the subject, i.e. OVS, OSV and VOS, are quite rare, even though none of them is completely unattested.

[11] That is, elements in specifier positions in the CP domain often have specific semantic features, such as [+Q], [+TOPIC], etc.

or otherwise, to do so. In itself, there is no reason why the asymmetric c-command relation should even *have* the property that there is some element that asymmetrically c-commands all others. The dominance relation has this property because of its nature: trees are defined in such a way that there will be one node that dominates all others. (Asymmetric) c-command is defined in such a way that it does not. Introducing it artificially is undesirable.

But let us for the moment follow Kayne's argument. Kayne argues that this node A dominates an abstract terminal element a. There are two options for the location of this node a in the phrase: either a will appear at the beginning of the phrase (in a Spec-Head-Comp ordering), or it will appear at the end of the phrase (in a Comp-Head-Spec ordering). Kayne argues that it must necessarily appear at the beginning, because of the following. Assume a tree with the following structure:[12]

(10)



Now Kayne says: "Let us think of the string of terminals as being associated with a string of time slots. (...) Let me then make the further claim that what is paired with each time slot is not simply the corresponding terminal, but the substring of terminals ending with that terminal (i.e., the substring produced up to that time)." (Kayne 1994, p. 37)

The notion "time slot" is not defined, but let us examine what this move gives us. For a Spec-Head-Comp ordering, it gives the following:

slot	string
1	a
2	ab
3	abc
4	abcd

For a Comp-Head-Spec ordering, the following would result:

slot	string
1	d
2	dc
3	dcb
4	dcba

[12] Kayne does not actually give the tree, he just gives the string of terminals that is associated with it, but I find that the argument is better explained with the tree.

The argument then goes that the Spec-Head-Comp ordering is preferred, because the element a precedes all terminals in *each* time slot, whereas in the Comp-Head-Spec ordering it follows all elements *only* in the *last* time slot. Kayne says (emphasis mine): "*If* the abstract root node for asymmetric c-command needs to be mapped (..) to a corresponding abstract "root node" for terminals, and *if* the root node for terminals must be in some fixed relation to every terminal in every substring, then (...) the fixed relation must be 'precedes'." *ibid.*

Obviously, *if* both the conditions that Kayne formulates here hold, then 'precede' must be preferred over 'succeed'. But both conditions are stipulated: it is not clear what they are based on, or why they should hold. Even worse, it is not even clear what they mean, because the notions on which they are based, the abstract root node A and its terminal element a, and the notion 'time slot', are undefined.

Summarising, it turns out that the idea that UG prescribes a universal Spec-Head-Comp ordering rather than a Comp-Head-Spec ordering is a stipulation. Kayne's attempt to give a rationale for this ordering does not hold up. Therefore, we must conclude that Kayne's attempt to reduce linear ordering to asymmetric c-command is problematic. The LCA may provide a linear ordering of terminal elements, but it does not say whether that ordering is precedence or subsequence.

Moreover, the empirical basis for the assumption that there should be a universal underlying order is also problematic: the word-order phenomena that we find across languages are too varied to support it. For these reasons, we are forced to reject the LCA.

2.3.2 Practical implementation

In spite of the fact that Kayne's theory has this essential shortcoming, the idea that Spec-Head-Comp ordering is universal has been widely accepted in syntactic research. The application of this idea to languages that are predominantly Spec-Head-Comp, such as English and the Romance languages, has been fairly successful, but serious problems are encountered when such an antisymmetric approach is applied to languages with a (partially) OV structure, such as Dutch or Japanese.

In order to account for structures in which the complement of a head is linearised before that head, an antisymmetric approach must assume that the complement has moved to a position higher than the head. This complement must move to the specifier position of some head. It is not possible to move to an adjunct position, because specifiers *are* adjuncts, and because a head can never have more than one specifier/adjunct.[13] For this reason, each movement that takes place introduces the need for an additional head.

However, there is only a very limited number of heads — and therefore landing sites — available in a standard tree. In many cases, this limited number of landing sites has proven to be too limited. In order to account for various word order phenomena, antisymmetric analyses often require more movements than there are heads available.

[13] If a head has two specifiers, non-terminals in either specifier will asymmetrically c-command non-terminals in the other. As a result, no linear order can be specified between them.

The standard solution for this problem is to assume that there are in fact more heads in the functional complex than hitherto assumed. The problem with this is that these heads are usually not associated with any specific feature or morphological element which would justify positing their existence.

For example, the C head is assumed to exist because many languages fill this position with some overt element. The existence of such elements justifies the adoption of the C head. This justification is often absent for a head that is posited in antisymmetric approaches. The head is merely there in order to provide a target for movement. This lack of justification is reflected in the names of such heads, which are often quite meaningless. Common names are W, X, Y or Z, and the projections that they head are WP, XP, YP or ZP.

A good example of this practice can be found in Shlonsky (2000), which I will discuss in chapter 3. Shlonsky posits the existence of an XP, YP, ZP, and also of 1P, 2P, 3P and even 22P. Another even more extreme example can be found in Koopman & Szabolcsi (2000), who argue that the clause contains an unlimited number of LPs, which function as targets for movement.

It should be pointed out that it is sometimes argued that there are indeed many more functional heads in the clause than hitherto assumed. Cinque (1999) for example examines the positioning of adverbs in a large variety of languages, and finds that there are remarkable similarities in the ordering relations of combinations of adverbs across languages. He argues that this can only be explained if we assume that the clause actually contains a large number of functional heads, basically one for each category of adverbs. An adverb can then only occur in the specifier of the head of its own category. Similar conclusions are reached on the basis of research into fronting positions, e.g. by Szabolcsi (1997). Like Cinque, Szabolcsi argues that the functional structure in the clause is much richer than previously assumed.

It is sometimes argued that once we understand more about the syntactic processes that are going on in clauses, it will turn out that the undesignated heads that are often assumed in antisymmetric approaches are in fact the heads that this research into adverb placement and fronting reveals. This, however, is problematic for two reasons. First, Cinque and Szabolcsi argue that the heads are providing specifier positions for specific categories of adverbs or fronted elements. It is not likely, then, that those specifier positions can also be occupied by other elements that are not adverbs, and that do not have the proper semantic feature to be licensed in such a position. Second, the heads that Cinque and Szabolcsi posit often do not have any justification in the form of morphological material. The only reason for positing them is the assumption that the observed ordering phenomena cannot be accounted for in semantic terms, and therefore have to be syntactic. This would be a valid reason for positing these heads, *if* it is really the case that the phenomena cannot be explained in semantic terms. Recent research, however, shows that at least some of these ordering phenomena certainly can be accounted for in semantics (Nilsen 2003). Future research may well show that this is possible for all such adverb and fronting phenomena. In other words, it is not at all certain that we really need to assume the existence of so many adverb-hosting and fronting heads in the clause.

If these heads that are posited to provide targets for movement are really undesig-

nated, they violate the principle of Full Interpretation (Chomsky 1986b), which states that the syntactic computation cannot introduce or contain any elements that are only needed to facilitate the syntactic computation. As such, these heads are undesirable.[14]

Not only are the heads that are targets for movement undesirable, the movements themselves are as well. Like the heads, the movements are not motivated in any way other than by the desire to account for a given word order. Traditionally, the assumption that a certain element moves is motivated by the observation that the element in question is positioned in some position X, while it is interpreted as if it were positioned in position Y. For example, a *wh*-element that is fronted from object position is interpreted as an object but has moved to a higher position because of the interrogative feature on it. For a lot of the movements that are posited in antisymmetric approaches, however, no such motivation can be given.[15]

Applying the LCA as developed by Kayne in the analysis of language data leads to a situation where it becomes inevitable to posit undesignated heads and movements motivated only by the need to account for the ordering. If it were the case that the LCA itself expressed a fundamental truth about Universal Grammar, we could accept these effects as following inevitably from something known to be true. However, we have seen that the LCA cannot give us a linear ordering of a set of terminal elements in a tree, which means we cannot accept it as a fundamental truth about UG. Add to that the fact that using the LCA as the theoretical basis for analyses of syntactic structures leads to an undesirable proliferation of functional structure,[16] and we see that we cannot use the LCA as the basis of our syntactic analyses. It is both theoretically and practically inadequate.

[14]Note that the LCA also violates the principle of Full Interpretation: the LCA maps a relation that is defined for non-terminals to a relation that is defined for terminals. This means that the non-terminals, i.e. elements such as X′, XP and also X, which in Kayne's theory is different from x, must be present in the structure. None of these elements have a semantic interpretation, however.

[15]It should be noted that there is at least one movement that is known to take place but for which no semantic rationale has been discovered (yet). This is the movement of the subject to spec,TP. There is quite a lot of evidence that the subject is generated in a lower position and then moves to its surface position. Such evidence includes quantifier float, the subject position in existential clauses, the position of subjects with quirky (i.e. non-nominative) case, etc. This movement is usually ascribed to the so-called EPP feature, although it is not very clear what this EPP feature is. There is no obvious semantic motivation for it, which means it is a violation of the principle of Full Interpretation. However, the empirical evidence for this particular movement is very strong, which is sufficient reason to posit the EPP feature, on the assumption that we will be able to give an explanation for it in the future. The existence of one case that appears not to adhere to Full Interpretation does not justify abandoning that notion altogether. In other words, the existence of one — obvious — case of semantically unmotivated movement does not mean we can posit a large number of unmotivated movements, if such a step would mean abandoning an important guiding principle of syntactic research.

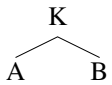
[16]This proliferation of functional structure follows from the LCA because it is very rigid in nature, not because the current implementation stipulates a Spec-Head-Comp ordering. Even if we were to stipulate that the LCA forces a Comp-Head-Spec ordering, we would still end up with a similar proliferation. A universal Comp-Head-Spec would be just as inadequate to describe order variation as a universal Spec-Head-Comp is.

2.4 Recursive Linearisation

2.4.1 Linearisation as a PF procedure

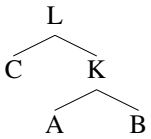
As discussed in chapter 1, Chomsky (1995, 1998, 1999) argues for a different conception of trees. Chomsky argues that the only operations that the computational system needs are Merge and Agree. Merge is the operation that builds trees. It takes two elements and combines them into a larger structure. For example, Merge can take the two elements A and B, and form a larger structure K(A,B), which can be represented as a tree:

(11)



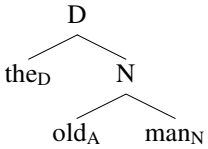
The elements that Merge operates on can themselves be the output of previous applications of Merge. In other words, they can be compound elements. For example, we can take the element K(A,B) and merge it with another element C, to form L(C,K(A,B)):

(12)



In Chomsky's view, the elements that Merge combines are lexical items. So for example, Merge can operate on old_{A} and man_{N} to form N(old, man), which can then be combined with the_{D} to form D(the, N(old, man)):

(13)



It is important to see that the trees thus formed are radically different from the trees that Kayne (1994) bases the LCA on. A tree as in (13) could never be linearised with the LCA because the only non-terminal nodes it contains are D and N, neither of which asymmetrically c-commands the other. In Kayne's theory, asymmetric c-command is *only* defined over non-terminal nodes, which means there is insufficient structure in (13) to linearise it according to the LCA.

Chomsky, however, intentionally abolishes X-bar theory, which is what Kayne's LCA is based on,[17] and furthermore, he takes another rather radical step. He argues that the structures that Merge creates are not linearly ordered (Chomsky 1995, p. 334). In a structure K(A,B), no order is defined between A and B. In other words, K(A,B) is the same as K(B,A). This move is motivated by the idea that at LF only the dominance

[17] Or rather, Kayne argues that the LCA explains why the X-bar schema holds.

relations are of importance. The linear ordering is not relevant to LF. Linear ordering is only relevant at PF, where the phonological representation of a phrase is computed.

In other words, Universal Grammar does not need to specify any universal linear ordering. What we need is a procedure at PF that derives the linear order from the tree structure. What we need to do, then, is to find a procedure that does this with the least amount of additional mechanisms.

When Merge creates a structure, the resulting structure is compound. The elements from which this compound structure is formed can be compound themselves, but they can also be simplex, i.e. not formed by Merge.[18] Such simplex elements are the terminal elements of the tree, which in a linguistic tree are the elements that carry phonological features and must be spelled out.

Linearising a tree, then, can be thought of as searching the tree for terminal elements to spell out. The most efficient way to search a tree is to do it recursively: the search procedure starts with the root node, and then applies itself to each of the subnodes. Traditionally, there are several methods for searching (binary) trees, breadth-first and depth-first being the most common. Of these, only depth-first search is compatible with standard syntactic assumptions.[19]

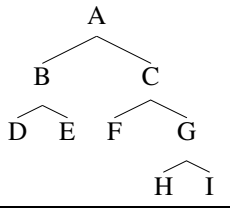
The search procedure that linearises a tree consists of the following steps:

(14) linearise tree T:
   a. if T is a terminal, spell out T.
   b. otherwise, take subnodes A and B of T.
   c. linearise A then B, or B then A.

In step (c), the procedure has to make a choice: it has to decide whether to spell out the two subnodes in the order A-B or in the order B-A. The choice which of the two subnodes to spell out first has a direct effect on the order in which the terminal elements are found, and consequently a direct effect on the order in which the terminal elements end up in the linear string.

To see how the procedure works, let us look at an example tree:

(15)



[18] It is conceivable that they are compound elements on another level, that is, formed by some operation other than Merge (e.g. in the lexicon). In the trees under consideration, however, they count as simplex elements because they are not formed by Merge.

[19] The difference between the two search methods comes down to this: breadth-first searches a tree layer by layer, whereas depth-first searches a tree branch by branch. As a result, a breadth-first search would yield linear orderings in which all the terminal elements in a specific *layer* end up adjacent, whereas depth-first search yields linear orderings in which all the terminal elements in a specific *branch* end up adjacent. Syntactic trees are built up on the assumption that the latter is the case. For some discussion, see appendix B.

In (15), the terminal elements are D, E, F, H and I. When the search procedure is at the top node A, the decision whether B has to be searched before or after C directly influences the order in which the elements are found: if B is searched first, D and E will be found and linearised before F, H and I. If C is searched first, F, H and I will be linearised before D and E.

In order for the procedure to work, it will need a way to make the decision which of two subnodes to linearise first. The linearisation procedure (which I will call RLin, for 'Recursive Linearisation') should be able to make this decision locally, that is, without using any information that is not available at the node that it is processing. All the information that is available during the syntactic computation is also available when RLin starts processing the tree, because it is not possible to strip any information off the tree before it is sent to PF.[20] This means that the information in (16) is all available to RLin when it processes a node K(A,B):

(16) a. the label of K
b. which of the two nodes A and B projects
c. whether the projecting node is a terminal element (i.e. a head)
d. whether the non-projecting node is selected or not

In a bare phrase structure approach, the label of a node K is simply one of the features of K. The features of K are available to RLin, which means that its label is, too. This gives RLin (16a). Because RLin has to split up K into its subnodes, the features of these subnodes are also available. Comparing the label of K to the labels of A and B allows RLin to determine which of the two nodes is the projecting node. This gives (16b).

A node that is a terminal element is a node that has phonological features, and it is furthermore a node that is not compound. Let us say that it is the presence of phonological features that enables RLin to determine whether a subnode is a terminal element or not.[21] This gives (16c).

In order to obtain the information in (16d), that is, in order to see whether a node is selected, RLin will have to compare the features on both nodes. If there are features of a certain category that match, the non-projecting node has been selected by the projecting node. If not, the non-projecting node is an adjunct. The matching features can be of three categories: categorial features, agreement features or theta features (in the sense of Reinhart 2002.)

Categorial features are the features that mediate c-selection: a head D selects for a complement N (or any of the other heads in the noun phrase, which all have some categorial feature that distinguishes them as nominal categories), and a head T selects for v. If two heads have matching agreement features, the two heads have been merged as the result of an Agree relation: T agrees with the subject of the clause and then attracts it. I assume that a similar relation holds between heads in the CP domain and

[20] Any procedure that would strip off information that is not necessary at PF is in fact a part of PF, because it is performing an operation that is essentially relevant to PF.

[21] This means that when a node projects, it does not pass on its phonological features to the newly created node, which is a reasonable assumption.

their specifiers, e.g. C attracting a *wh*-phrase. If two nodes share theta-features, the projecting head s-selects its sister.[22]

We see that in order for RLin to determine the information in (16), it only needs to examine the features of the node K and of its subnodes A and B. The next question that we need to answer is how RLin uses this information in order to decide which of the two subnodes of a node K(A,B) it is to linearise first.

As noted in section 2.3.1, Kayne observes that spec,CP and spec,TP, that is, fronted elements and subjects, generally precede their heads. Subjects in spec,TP agree with T and are therefore selected elements in the sense described above. We can say the same for fronted elements if we adopt the common assumption that their fronting is mediated by features. For example, a *wh*-element is fronted because it has a *wh*-feature which it shares with the head of the specifier position to which it moves. We can implement Kayne's observation in RLin by saying that a selected element is always linearised first. That is, if in a node K(A,B) B is the projecting node, B is not a head and A is selected by B, A is linearised first.

Unlike the ordering of selected specifiers and their heads, it is an established fact that Head-Complement combinations *are* subject to variation. The best known example is the distinction between VO and OV languages (Greenberg 1966); other examples of this variation are the distinction between clause-initial and clause-final complementisers and force markers, and between prenominal and postnominal determiners. This variation is usually captured in the so-called head parameter, which determines whether the head precedes or follows its complement.[23] In the current model, we could implement this head parameter as a parameter for RLin. We could say that if in a node K(A,B), A is the projecting node and A is a head, RLin uses the setting of the head parameter in order to determine which subnode to linearise first.

However, if we do this, we run into a problem: complements are always selected. They are either c-selected, e.g. D selecting N or C selecting T, or they are s-selected by a lexical head, i.e. V selecting its object. That means that for this node K(A,B) not only (16c) is relevant, but also (16d). We have just established that selected elements are always linearised first. This means that a head parameter would not be relevant at all: the selected element, that is, the complement, must be linearised first.

One way to solve this problem is to say that the head parameter always overrides the requirement that selected elements are linearised first. However, we can also do away with the head parameter and instead say that heads, like selected elements, are always linearised first. This means that when RLin encounters a head-complement node, it must fulfil two requirements: there is a requirement that the selected element (i.e. the complement) be linearised first, and there is a competing requirement that the head be linearised first.

In other words, the two requirements clash. We can resolve this clash by saying that these two requirements are ordered with respect to each other. The variability that we observe in Head-Comp orderings can then be accounted for by saying that

[22] Reinhart (2002) develops a system in which theta roles are abandoned and replaced by sets of binary features consisting of the features [±change of state] and [±mental state].

[23] In an antisymmetric approach, the head parameter determines whether the complement moves over the head to a higher position or not.

the relative ordering of these two principles is parameterised. In categories that have a head-complement order the head requirement is stronger, in categories that have a complement-head order, the selection requirement is stronger.

So we see that we can describe the linearisation of a structure [Spec [Head Comp]] with two ordered principles. We have a principle H which requires that heads are linearised first, and we have a principle S which requires that selected elements are linearised first. If we have the ordering $S > H$ we obtain a linear order of Spec-Comp-Head, if the ordering is $H > S$, we obtain Spec-Head-Comp.

We know that within one language, the head parameter can have different values for different categories. For example, in Dutch, V follows its complement, whereas C and D precede it. This means that the relative ordering of S and H can be specified independently for different categories. This is potentially problematic, because syntactic studies show that the number of (functional) categories in language can be quite extensive. A clause does not only have C, T, v and V, there is also evidence for projections like Asp, Neg, Top, Focus and perhaps others. In the noun phrase, we have apart from D and N also categories such as K, Poss, Num and Gen.

A simple calculation reveals that if we have six categories in the noun phrase, and each category can specify its own ordering of S and H, we obtain 2^6 = 64 different word orders. However, in any one language, many of these orders will not be distinguishable for the simple reason that most languages do not overtly mark all of these heads. Therefore, we can limit the group of heads that can specify their own ordering of S and H to those heads that are overtly marked.

More specifically, only those heads that have an *independent* morphological form can specify their own ordering for S and H, whereas a head that is marked by only an affix, cannot. Instead, such a head must adopt the settings of the head to which it affixes, i.e. its complement.[24] For example, in section 3.4.2 we will see that in Arabic only D and N project independent forms. The other heads that are projected, Poss and Num, only occur as affixes on N. This means that the noun can specify an ordering for S and H, and that this ordering applies not only to N, but also to Num and Poss. D, on the other hand, could in principle determine its own setting.[25]

One advantage of limiting the set of heads that can specify their own settings to those heads with independent morphological form is that a child acquiring its mother tongue will be able to determine quickly for which heads it must set the ordering, and furthermore, it can determine quickly what the ordering must be. Once a child has determined that for example X is a head and YP its complement, it only needs to look at the linear order in which X and YP occur to determine the relative ordering of S and H.[26]

We can even go one step further and say that the ordering $H > S$ is the default order. That is, a child will assume that every category has the ordering $H > S$, unless

[24] Or, more precisely, only an independent morphological form can determine an ordering for S and H, and this ordering applies to each syntactic head that is represented in that form.

[25] We know, however, that both D and N have the ordering $H > S$, because both heads precede their complements.

[26] Obviously, the linear order of X and YP must occur with a certain consistency. A single occurrence of OV does not guarantee that the language is indeed OV.

it finds clear evidence in the primary linguistic data for the opposite order. Only if the primary linguistic data shows that a specific head consistently follows its complement will the language-learning child switch the ordering to $S > H$. This accounts for the observation that Head-Comp orderings occur more than Comp-Head orderings (Greenberg 1966), and that newly created languages, e.g. pidgins and creoles, have a strong preference for VO orders (Holm 2000). Because a child that forms a creole out of a pidgin does not hear any consistent evidence for OV structures, it will use the default setting, which yields a VO structure. At the same time, we can understand how it is possible for OV languages to exist and to be stable over long periods of time: because OV structures occur abundantly, a child acquiring the language has sufficient evidence to switch the order from its default.

The two principles S and H only cover heads and selected elements. Non-selected specifiers, i.e. adjuncts, do not fall under either of them, which means they cannot be linearised with them. As we have seen, there is variation in the position in which adjuncts are linearised: we have both left adjunction (e.g. prenominal adjectives) and right adjunction (e.g. postnominal adjectives). The most straightforward way to deal with adjuncts is to say that their linearisation is parameterised. That is, if RLin encounters a node K(A,B) in which A is the projecting node and B is not selected, the order in which A and B are linearised is determined by a parameter, which I will call the adjunct parameter.

The considerations concerning the learnability of the relative ordering of S and H also apply to the adjunct parameter. We must limit the ability to specify a value for the adjunct parameter to independent morphological forms.[27]

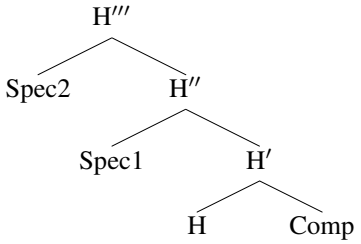
Summarising, we see that RLin only needs two parameters to account for the ordering variability in language: one parameter to linearise adjuncts, and one to order the two principles S and H. All the other information that RLin needs in order to linearise a tree structure is available in that tree structure itself. In the remainder of this section, I will demonstrate how RLin works on an abstract tree. In the following chapters, I will show how RLin accounts for word order in the Arabic noun phrase.

[27]The adjunct parameter may well have a default value just like the parameter that orders H and S. What this default value is would depend on the relative frequency of left- and right-adjunction. I am not aware that either of these is more frequent than the other, so I will not take a position on this.

2.4.2 A demonstration

To demonstrate how the procedure works, take the following tree:

(17)



(17) contains a tree that is a projection of the head H. H is merged with three different elements, a complement and two specifiers. Let us first consider the case where both specifiers of H are selected.[28] RLin will start at the top node of the tree, H‴. This node has two subnodes, H″ and Spec2, of which Spec2 is selected. Principle S says that a selected element must be linearised first, so RLin will first spell out everything in Spec2 and then continue on with H″. H″ is treated in the same way: Spec1 is selected and must therefore be linearised before H′.

When RLin gets to H′, it finds that both principles S and H apply: H is a head, to which principle H applies, and Comp is a selected element, to which principle S applies. So depending on the relative ordering of the two principles, we either get H-Comp or Comp-H.

This means that when Spec1 and Spec2 are both selected, we can derive the following linearisations:

(18) a. Spec2 Spec1 H Comp
b. Spec2 Spec1 Comp H

If the two specifiers in (17) are not selected, the adjunct parameter starts playing a role. If the adjunct parameter is set to adjuncts first, the results that obtain are the same as in (18): the specifiers are linearised first, which means they will appear before H and Comp, and in the order specified.

However, if the adjunct parameter is set to adjuncts second, we get very different results. Now, when RLin starts out with H‴, it will first linearise H″. In H″, it will do the same: the adjunct is to be linearised last, which means H′ must be processed first. In H′, the relative ordering of the two principles S and H determine the order of H and Comp.

After H′ has been linearised, we either have H-Comp or Comp-H. At this point, RLin has completed the first part of the linearisation of the node H″. The second part is the linearisation of Spec1. As a result, everything in Spec1 will be spelled out after H-Comp or Comp-H. When RLin has spelled out Spec1, it has finished the linearisation of H″, which also completes the linearisation of the first subnode of H‴.

[28] This is not a very common situation in language, but we can think of languages such as Bulgarian that have multiple *wh*-movement.

The second subnode, Spec2, will be linearised next. The two resulting orders are in (19):

(19) a. H Comp Spec1 Spec2
b. Comp H Spec1 Spec2

The interesting thing to note is that the (non-selected) specifiers in this case are linearised in the reverse order with respect to (18). In (18), Spec2 precedes Spec1, but in (19), Spec1 precedes Spec2. In other words, a setting of adjunct second yields mirror-image orders.

2.4.3 Derivation by phase

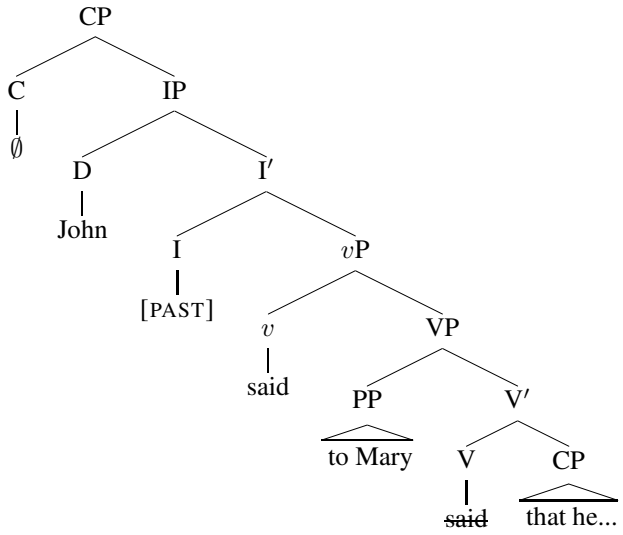
As noted in chapter 1, Chomsky argues that derivation takes place phase by phase, that is, spell-out is applied to a phase as soon as it is ready. In principle, RLin would be compatible with a system in which spell-out only applies to the entire tree after it has been built. To take up the example of the previous chapter:

(20) John said to Mary that he was going to leave her

As explained, Chomsky argues that the embedded CP *that he was going to leave her* is built and spelled out independently from the matrix CP. RLin would be compatible with an approach in which the structure is built up in its totality, i.e. with the embedded and the matrix CP in one large syntactic tree, which is then linearised by RLin.

However, RLin is also compatible with the derivation by phase approach. Since a phase is a piece of hierarchical structure, it can be linearised by RLin in the same way that a larger syntactic tree can be. Therefore, I will say that RLin does indeed apply to phases. This means that we must consider a phase that has been spelled out as equivalent to a terminal element, at least from the point of view of RLin. Take, for example, the linearisation of the matrix CP in (20):

(21)



The embedded CP is merged into this structure, which means that RLin will at some point encounter it. But because the CP has already been spelled out, RLin no longer needs to process it. Instead, it has to treat it in the same way that it treats the terminal elements *John* and *said*.[29]

2.5 A conceptual comparison

In this chapter, I have presented RLin as an alternative to the more or less standard linearisation approach as advocated by Kayne (1994). It may be useful to give a quick comparison of both proposals, to see where the merits of each lie.

The idea that hierarchical tree structures are defined in such a way that the linear structure can be read off of them is conceptually attractive. However, as we have seen, Kayne cannot make this claim without making the stipulation that the universal underlying word order is Spec-Head-Comp. This is a stipulation because the LCA only predicts that specifiers and complements appear on opposite sides of the head. It does not say on which side the specifier and the complement are. Nonetheless, if one accepts the idea that tree structures indirectly (i.e. through the LCA) define a universal linear order, then the assumption that this order is Spec-Head-Comp rather than Comp-Head-Spec makes sense, because we do indeed see more Spec-Head and Head-Comp structures in human language than we see Head-Spec and Comp-Head.

The core idea of RLin is that hierarchical tree structures do *not* define linear ordering relations, not even in an indirect way. In this, it differs fundamentally from antisymmetric approaches. The argument is that linear ordering is determined at PF. This position forces one to posit principles and parameters that operate at PF. In other

[29]The same reasoning applies to the PP *to Mary*, which is (presumably) also a phase that has already been spelled out. And in fact, if we consider DPs to be phases, *John* must be considered a phase as well, to which spell-out has already applied.

words, we must assume that not only core syntax is organised with principles and parameters, but the PF component of language is, as well. The PF component in this model is the place where parameters such as the head parameter are implemented.[30]

Another — apparent — advantage of the antisymmetric approach advocated by Kayne is that the LCA allows us to derive the basic properties of X-bar structure, such as the restriction that nodes can have at most two subnodes, the restriction that a head can only have one complement and one specifier, and the restriction that each X-bar structure must have exactly one head. However, this advantage disappears when one abandons the X-bar schema, as Chomsky does. If we do not have an X-bar schema, there is of course no need for it to be reduced to a more principal notion.

If one adopts a minimalist approach, with bare phrase structure and the principle of Full Interpretation, the LCA even becomes quite problematic, because it is far from minimalist. The LCA crucially relies on the existence of non-terminal nodes such as XP, X′ and X.[31] These elements all violate the principle of Full Interpretation, because they only exist for the syntactic computation. They do not have any semantic value, and should therefore not be present in the syntactic structure.

RLin, on the other hand, is compatible with a bare phrase structure approach and with the principle of Full Interpretation. RLin operates on tree structures created by Merge as described by Chomsky (1995), and it does not need to introduce any elements into the hierarchical tree structure that are only needed for linearisation. It only requires the presence of syntactic features that we must assume to be present for independent reasons.

Another effect of the LCA is that by ascribing all word order effects to movements taking place in core syntax, the structures that arrive at LF will be rather different from one language to the next. For example, a language that has the order D-N will have a structure $[_{\mathrm{D}}$ D N] arriving at LF, whereas a language that has the order N-D will have a structure $[_{\mathrm{X}}$ N $[_{\mathrm{D}}$ D t]] arriving at LF. This poses a problem for the standard assumption that the syntactic structure of a phrase is mapped onto a semantic structure. We would need to account for the fact that very different syntactic structures map onto the same semantic structure. One way to do this would be to argue for the possibility of reconstruction: the structure $[_{\mathrm{X}}$ N $[_{\mathrm{D}}$ D t]] can be reconstructed to the original $[_{\mathrm{D}}$ D N] when it arrives at LF. This, however, raises the question why the movement takes place in core syntax at all.

RLin has the advantage that it does not use movement to account for word order effects, which means that the syntactic structures that arrive at LF will be more uniform across languages. Both a language that has the order D-N and a language that has the order N-D has the syntactic structure $[_{\mathrm{D}}$ D N] (or $[_{\mathrm{D}}$ N D], given that order is not defined) arriving at LF.

A final thing that must be mentioned is that antisymmetric analyses often also

[30] In the current implementation of RLin, the head parameter exists as the ordering parameter of the two principles S and H. Note that the question whether such parameters should indeed be PF parameters or whether they are parameters of core syntax is to a certain extent an empirical issue. If for example the head parameter is a PF parameter, then we predict that it does not interact with any principle or parameter of core syntax. At most, it can interact with other PF principles or parameters.

[31] Note that in Kayne's approach, X is a non-terminal node dominating the terminal node x.

attempt to eliminate head movement and covert movement. They try to reduce all movement to overt movement of XPs. Personally, I do not really see the need for the elimination of head movement,[32] but if one wants to pursue this course, there is in principle no reason why it cannot be combined with an RLin approach to linearisation. RLin operates on (binary) tree structures of any size and complexity, which means it can operate equally well on trees in which all head movement has been replaced by phrasal movement. The same is true of covert movement: the tree structures that are produced when all movement is overt will be more complex than when one allows covert movement, but RLin will be able to linearise those just as well.

[32]Chomsky (1998) claims that head movement only has effects on word order, never effects on the semantics of a phrase, and he argues that it therefore must take place at PF, not in core syntax.

3

Non-event nominals

In much research on noun phrases, event nominals are the starting point, presumably because they resemble verb phrases to some extent. However, event nominals are derived nominals, which means that they do not represent the noun phrase in its most primitive form. For that, we have to look at nominals that are not eventive.

With the term *non-event nominals* I refer to basic nouns such as *table, car, person*. Included also are the nouns that Grimshaw (1990) calls "simplex event nominals": nouns that do refer to an event, but that according to Grimshaw's definitions do not have argument and event structure. This latter category can be confusing, however, because event nominals are often ambiguous between a complex event and a result reading. Therefore I will not use such words in the examples in my discussion.

The reason for taking these non-event nominals as a starting point, rather than deverbal nouns, lies in the fact that deverbal nouns have at the same time nominal and verbal properties. Studies such as those of Abney (1987) and for Arabic Fassi Fehri (1993) have shown that deverbal nouns such as gerunds and in Arabic masdars start out as verbal projections which are at some point in the derivation "transformed" into nouns. As a result, their properties are not solely nominal in character. Rather, they present a mixture of nominal and verbal properties. Non-event nominals, on the other hand, can be argued to be "purely" nominal, because they do not have such a verbal base. If we examine those first, we can then move on to deverbal nouns, to see which of their properties are nominal, and which of their properties are verbal.

In section 3.1, I discuss the basic structure of the Arabic noun phrase. First, I show how a noun in Arabic licenses a genitive complement. I argue that the ability of nouns to license a genitive complement is on a par with a verb's ability to license nominative and accusative arguments, and that it is mediated by a functional head. Following that, I discuss which other functional heads we must posit in the Arabic noun phrase.

In section 3.3 I elaborate the model thus developed, by discussing how modifiers other than genitives fit into it. I show that such modifiers occur in two types: prenominal modifiers are heads, whereas postnominal modifiers are full projections.

The discussion leads up to the section 3.4, where I discuss the relative ordering of all the elements in the noun phrase. I discuss the existing antisymmetric approaches to the linearisation of the Semitic/Arabic noun phrase, showing how they suffer from the problems that I discussed in chapter 2. After this discussion, I work out the linearisation procedure that I introduced in chapter 2 and show how it can account for the data discussed in this chapter.

3.1 Noun complementation

3.1.1 Genitive-marked complements of nouns

Noun phrases often contain additional elements apart from the head noun. For example, a noun can usually take syntactic dependents, such as prepositional phrases or genitive/possessive complements. If a noun in Arabic takes a genitive complement, a special construction is used, which is commonly known as the *construct state* (See also Ritter 1991. In this construction, the genitive complement follows the head noun directly, and it takes genitive case:

(1) sayyār-at-u -l-raǧul-i
car-F-NOM the-man-GEN
'the man's car'

The head noun of the construction in (1) is *sayyāra* 'car', its genitive complement is *raǧul* 'man'. The order of the constituents in the construction is always *head noun – genitive modifier*, and the complement noun is assigned genitive case.

Case is generally divided into two types. There is structural case, which is assigned to a certain syntactic position, regardless of the theta role of the noun occupying that position. Structural case is assigned by a functional head. The other type of case is inherent case, which is assigned on the basis of the semantic function of the noun that receives it. The assigner of inherent case is a lexical head.

The typical examples of structural case are nominative and accusative. These are assigned by T and v respectively, without any reference to the theta role of the noun that receives them. A common example of inherent case is the dative, which is assigned to a noun that has a receiver theta role. Other inherent cases are the instrumental, the locative, etc.[1]

Given that there are two types of case, the question is raised which type genitive case belongs to. In (1), the genitive modifier expresses the possessor of the head noun, but the theta role of the genitive noun is not necessarily that of possessor. Lindauer (1995) shows that the German genitive can have a large number of thematic roles. This property is not specific for German. Other languages that have a genitive case generally allow the same roles. Arabic is no different in this respect. (2) shows some of Lindauer's examples adapted to Arabic:

[1] Prepositions are presumably much like inherent case, in that they carry meaning in themselves.

(2) a. bayt-u ’ab-ī
house-NOM father.GEN-my
‘my father’s house’ (possessive)

b. ḥall-u -l-ṭālib-i
solution-NOM the-student
‘the student’s solution’ (subject)

c. ḥall-u -l-muškilat-i
solution-NOM the-problem-GEN
‘the solution of (to) the problem’ (object)

d. ḥağm-u -l-ġurfat-i
size-NOM the-room-GEN
‘the size of the room’ (property-carrier)

e. ᶜamal-u -l-šāᶜir-i
work-NOM the-poet-GEN
‘the poet’s work’ (agent)

f. madrasat-u ’aẖ-ī
school-NOM brother.GEN-my
‘my brother’s school’ (belonging-to)

As for example Longobardi (1995) and De Wit (1997) point out, the fact that the genitive noun can have any of a number of theta roles shows that genitive is a structural case, like nominative and accusative, and not an inherent case, which is linked to a specific theta role.[2]

Standard Arabic actually provides some further evidence that supports this conclusion. Arabic has three cases: nominative, genitive and accusative. There is no dative or any other case that is undisputedly inherent. If one looks at the properties of the three cases, the genitive does not stand out in any way that would warrant the conclusion that it is of a different nature than the other two cases.

First of all, morphologically, the genitive does not stand out. Case in Standard Arabic is indicated by suffixes, in most instances short vowels: *-u* for nominative, *-i* for genitive and *-a* for accusative. Dual and plural suffixes indicate case with different suffixes, but the genitive is in essence the same in those suffixes as well.[3] Furthermore, when looking at the syntactic properties, one can say that the case that a noun has in Arabic is always determined by its syntactic position. A (free-standing) topic always has nominative case,[4] as does a subject. An object always has accusative case, and

[2]De Wit (1997) argues that the noun phrase has two structural case positions. In her view, a noun phrase such as *Caesar’s destruction of the city* shows a nominative-like case on *Caesar* and genitive on *the city*. If, however, there would be two cases in the noun phrase, we would see more evidence of it. One would expect that languages with morphological case systems would actually have two morphological case forms in the noun phrase, which is not borne out by the facts. Even in situations where two such cases in the noun phrase would be useful, e.g. with nominalisations where both the subject and the object are expressed, languages with morphological case use the genitive for one of the arguments, and employ a different strategy for the other. For example, Arabic uses accusative or a preposition to license the object, and Russian uses an adjectival-like form for the subject (De Wit 1997).

[3]In fact, the dual and plural suffixes only have one combined form for genitive and accusative. See appendix A for details.

[4]Provided that its place in the clause is taken by a resumptive pronoun. An somewhat archaic option is to topicalise an object without using a resumptive pronoun. In that case, the topic takes accusative case.

so do adjuncts of any kind. A noun that is the complement of another noun receives genitive case, as do complements of prepositions (many of which are derived from nouns).

In all instances, case is assigned based on the structural position of the noun, independent of the noun's theta role. For topics, this is obvious. As for the subject of a verb, it can be an agent, but also a cause, experiencer, patient, etc. The object cannot be agent or cause, but other non-causative roles are open to it: experiencer, patient, etc. The genitive can express ownership, but also a series of other roles, as we have seen above.

We conclude that Arabic only has structural case, and that the genitive is a structural case, like nominative and accusative. This means that nouns, like verbs, are able to license an argument. The functional shell of the verb creates two positions in which arguments can be licensed. In the same way, the functional shell of the noun phrase creates a position in which an argument can be licensed. Structural case is the formal property through which this licensing takes place.[5]

Chomsky (1995) argues that nominative and accusative are assigned by a functional head in an agree relation. Nominative is assigned by T, accusative is assigned by v. Since genitive is a structural case, like nominative and accusative, it must be assigned by a functional head as well. The existence of such a head, usually termed Poss, has been argued for by Valois (1991), Delsing (1993), Szabolcsi (1994), Longobardi (1995), and others.

This Poss head is a projection of the feature POSS, which indicates whether the head noun in the construction has a syntactic dependent.[6] It is often pointed out that the feature POSS resembles the feature TENSE on the head T. TENSE has the value ±FINITE, and if TENSE is +FINITE, it has an additional set of φ-features, which are unvalued at the onset of the derivation. Similarly, POSS has the value ±POSS, and, as data from various languages show, the possessive feature of a noun can be associated with φ-features. In a language like Hungarian, for example, the possessive marker agrees with the possessor:

(3) a fiú kalap-ja
the boy.NOM hat-POSS.3SG
'the boy's hat'
(Szabolcsi 1994)

Furthermore, it is also possible to introduce the topic with a topic marker. In such a case, the topic marker assigns accusative case to the topic.

[5]There are two differences between the argument-licensing properties of nouns and verbs. A verb can license two arguments (even three, in the case of double object constructions), whereas the noun can only license one. This, I assume, is due to v, which not only licenses an argument, but also introduces another theta role. Presumably, v extends the predicate in some way. (E.g. it is sometimes argued to be a transitiviser.) The noun phrase does not have such a functional head. The second difference is that verbs assign specific theta roles to their arguments, whereas the theta role of a noun's argument depends on the semantics and pragmatics of the phrase at hand and is usually not determined by the noun itself (except in the case of relational nouns, such as *brother, sister, friend* etc.)

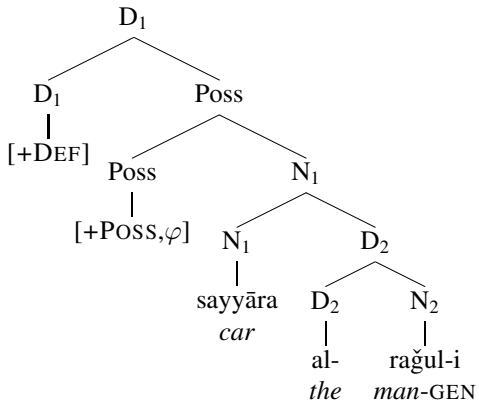
[6]It should be noted that although this head is called 'Poss', the relation that exists between the head noun and its complement is not necessarily one of possession, as made clear above.

The head noun in (3) is *kalap(-ja)* '(his) hat'. The suffix *-ja* indicates that the noun has a possessor, which is *fiú* 'boy', but this suffix not only indicates possession, it is also marked for 3rd person singular, agreeing with the possessor. Similar patterns are found in languages such as Yup'ik, Maya and Turkish (see Abney 1987).

The agreement in φ-features that Poss shows is significant. In the clause, T establishes an agree relation with the subject in order to value its unvalued φ-features. In this process, nominative case is assigned to the subject. Because Poss shows a similar agreement in φ-features, we can say that the same process is taking place: Poss has unvalued φ-features which need to be valued. For this reason, Poss establishes an agree relation with the complement of the noun, in which the complement noun is assigned genitive case.

These considerations give us the following — preliminary — tree for the Arabic construction:

(4)



In this tree, the head Poss has the value [+POSS] plus a set of unvalued φ-features. Because the φ-features are unvalued, Poss is active and will try to value these features. It will do this by probing its (c-command) domain for an active match. It finds a match in the complement DP *al-rağul*, which is active because it has an unvalued CASE feature.[7] A match is established, and the unvalued features on both sides are valued.

The fact that the φ-features do not show up on the head noun in Arabic is not problematic. The φ-features of the clausal head T are not marked overtly in many languages. We can therefore simply say that the φ-features of Poss in Arabic are covert.

3.1.2 Definiteness inheritance

Arabic nouns are marked for the feature DEFINITENESS. That is, a noun in Arabic will take either a definite determiner or an indefiniteness marker. The definite determiner

[7] I have put the genitive marked noun in the complement position of the head noun, not in the specifier position of some functional projection. Although nothing in the present discussion really hinges on this, I believe comp,NP is the most likely position, because the genitive noun is not an external argument. The external argument of the head noun is *R*, as Higginbotham (1985) and Zwarts (1992) argue.

is a prefix *al-*, whereas the indefiniteness marker is a suffix *-n*. These two elements are in complementary distribution, as shown in (5c):

(5) a. rağul-u-n
man-NOM-INDEF
'a man'
b. al-rağul-u
the-man-NOM
'the man'
c. *al-rağul-u-n
the-man-NOM-INDEF

Adjectives that modify a noun agree with that noun in definiteness. That is, the adjective receives either a definite article or an indefiniteness marker, just like the noun:

(6) a. bayt-u-n kabīr-u-n
house-NOM-INDEF large-NOM-INDEF
'a large house'
b. al-bayt-u -l-kabīr-u
the-house-NOM the-large-NOM
'the large house'

In both examples of (6), the adjective *kabīr* 'large' has the same definiteness marking as the noun *bayt* 'house'. This is what is known as definiteness agreement in Arabic.

One of the interesting properties of the genitive construction is that its head noun is not marked for definiteness. It has neither the definite article, nor the indefiniteness suffix. Consider again the example given in (1), repeated here as (7):

(7) a. sayyār-at-u -l-rağul-i
car-F-NOM the-man-GEN
'the man's car'
b. (*al)-sayyār-at-u -l-rağul-i
(the)-car-F-NOM the-man-GEN
'the man's car'
c. sayyār-at-u-(*n) -l-rağul-i
car-F-NOM-(INDEF) the-man-GEN
'the man's car'

As one can see, the head noun *sayyāra* cannot take the definite article *al-*, nor can it take the indefiniteness suffix *-n*. Instead, it must remain unmarked for the feature DEF.

At first sight, it might seem that the head noun of a genitive construction does not have a DEF feature. This is not the case, however. When a modifying adjective is added to a noun that has a genitive complement, the adjective *is* marked for definiteness:

(8) a. bayt-u -l-raǧul-i -l-ḥaǧariyy-u
house-NOM the-man-GEN the-stone-NOM
'the man's stone house'
b. bayt-u raǧul-i-n ḥaǧariyy-u-n
house-NOM man-GEN-INDEF stone-NOM-INDEF
'a man's stone house'

In (8), the adjective *haǧariyy* '(made of) stone' modifies the head noun *bayt* 'house'. Note that an adjective that modifies the head noun of a genitive construction does not follow directly after the head noun. Instead, it follows after the genitive noun. Although this is the same location that an adjective modifying the genitive noun would appear in,[8] it is obvious that *haǧariyy* 'stone' in (8) modifies *bayt* 'house', not *al-raǧul* 'the man': first of all, the phrase refers to a stone house, not to a stone man. Secondly, the adjective not only agrees in definiteness, it also agrees in case. And in (8), the case on *haǧariyy* 'stone' is nominative, like the case on *bayt* 'house', but unlike the case on *al-raǧul* 'the man', which has genitive case.[9]

When we look at the definiteness markings on the adjective in both examples of (8), we see an interesting phenomenon: the adjective, although it obviously modifies the head noun *bayt*, has the same definiteness marking as the genitive noun *raǧul* 'man'. Because the adjective modifies the head noun, and because it clearly agrees with it in case, we cannot say that the adjective also agrees with the genitive noun. There is no precedent for such a "double" agreement process, and it is not clear how it could be implemented.

In other words, the examples in (8) show that the head noun *bayt* 'house' does have a definiteness feature, even though it is not marked for it. Furthermore, it is clear that the definiteness feature that the head noun has is the same as the definiteness feature of the genitive noun. This phenomenon is commonly known as *definiteness inheritance*: the head noun of a genitive construction inherits the DEF feature of its complement. Obviously, we need to account for this.

Given that we want to keep our grammar as minimalist as possible, we prefer to account for it with the operations that we already have in our grammar. This means that definiteness inheritance must be the result of an Agree operation: the DEF feature enters the derivation unvalued, and in the course of the derivation it is valued through Agree. Unfortunately, if we look at the tree above in (4) we see no obvious way in which this can take place.

There is another fact that is of importance here. Whether D enters the derivation with a valued or with an unvalued DEF feature depends on the value of the POSS feature: if Poss has the value [–POSS], i.e. when the head noun does not have a genitive complement, DEF enters the derivation already valued, and appears either as *al-* or as *-n*. It is only when Poss has the value [+POSS] that DEF is unvalued at the onset of the derivation, and the noun in question appears without either *al-* or *-n*.

[8] If both the head noun and the genitive noun are modified by an adjective, the adjective of the genitive noun comes first.

[9] Adjectives also agree in gender and number. See chapter 4 for details.

We can account for both of these facts if we say that we do not have two separate heads D and Poss in the construct state. Instead, we have a combined D/Poss head, which projects both the DEF *and* the POSS feature. The idea that one single head can be a projection of two features is taken from Giorgi & Pianesi (1997). The theory of Giorgi & Pianesi says that all syntactic features present in a structure can in principle project a functional head. Projection takes place in a specific order. That is, if two features X and Y are ordered as X $>$ Y, then Y can only project before X projects. Once X has projected, Y can no longer project.

It is not necessary, however, for all features to project a separate head. A feature can remain unprojected. Furthermore, Giorgi & Pianesi argue that functional categories can be what they call *syncretic*. That is, two (or possibly even more) syntactic features can be projected onto one single head. This is only possible for two "consecutive" heads, that is, two heads which are projected one after the other.[10]

As a typical example of a syncretic head, Giorgi & Pianesi mention the ending *-a* in an Italian adjective like *bella*. The ending *-a* expresses both singular and feminine, which are both independent grammatical features and could therefore in principle project independent heads.

A special type of syncretic category is the so-called *hybrid* category. A hybrid category is a syncretic category with the additional property that the value of either feature is fixed if the other feature has a specific value. As an example, Giorgi & Pianesi give the English Agr/T category. Agreement and tense in English project as a syncretic category, because they are expressed together as one element. But moreover, the category is a hybrid one, because if the agreement takes the value 3rd person singular, the tense feature is fixed on present tense.

The example may seem a little odd, because one could argue that the past tense ending *-ed* in *he walked* is marked for third person singular and for past tense. But in Giorgi & Pianesi's (1997) model, such an ending is only marked for past tense, and not for agreement, because the ending *-ed* occurs in all persons. The only ending that is explicitly marked for (third person) agreement is *-s*, which is also explicitly marked for tense. Agr/T is a hybrid category because if Agr is explicitly valued for a specific person, T is automatically valued for present tense.

The heads D and Poss are consecutive in the sense that D directly dominates Poss. Or, more properly, the DEF feature is projected immediately after the POSS feature. What I will argue is that in the construction under consideration, these two features are projected syncretically. That is, in the construct state, the functional shell of the noun phrase in Arabic contains one functional head that has both a DEF and a POSS feature.

But this is not all. In Arabic, the features DEF and POSS are not only projected syncretically, they also form a hybrid category. If POSS has the value [+POSS], DEF is forced to remain unvalued, at least at the point of lexical selection.[11] Because it

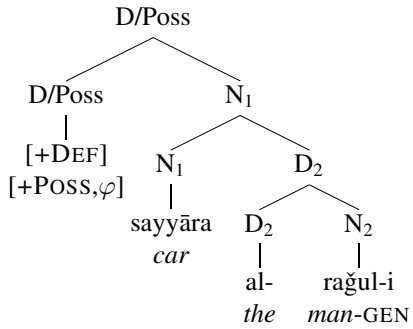
[10] A similar idea is also expressed by Reuland (1990), who argues that the V and I heads in the Dutch clause project as a single head, i.e. that V-to-I movement takes place "in the lexicon", as Reuland calls it.

[11] Note that Giorgi & Pianesi only talk about the possibility of one feature fixing the value of the other, not of one feature forcing the other feature to remain unvalued. However, Giorgi & Pianesi do not make use of a valued feature system, which means that in their model features always have a value: they cannot

enters the derivation unvalued, it has to be valued during the derivation.[12]

In other words, the structure of the Arabic genitive construction is not as given in the tree in (4). Rather, it is as in (9), with a hybrid D/Poss head:

(9)



With this tree, we find a straightforward explanation for the definiteness inheritance effect. D/Poss is a hybrid category: if POSS has the value [+POSS], it forces DEF to remain unvalued. In this way, we account for the fact that DEF remains unvalued only when POSS has the value [+POSS].

Because DEF is unvalued, it must be valued during the derivation. In the configuration of (9), the DEF feature will automatically be valued when Poss probes for a match to value its φ-features. As I explained in chapter 1, Chomsky (1995) claims that when an active head finds a match, all the unvalued features on both the probe and the goal will be valued, to the extent possible.[13] In the current situation, the goal is the complement DP. This DP has a valued DEF feature, which means that the DEF feature of the hybrid D/Poss head is automatically valued. So we see that definiteness inheritance is linked to genitive case assignment: the process that assigns genitive case has as a side effect that the head noun inherits the definiteness of its complement.

Fassi Fehri (1999), however, argues that definiteness inheritance and genitive assignment should not be linked. He claims that there are instances of genitive constructions in which there is no definiteness inheritance. One such case would be given by (10), in which a construct state noun functions as a predicate:

(10) hāḏā ʾaḥ-ī
this brother-my
'this is my brother'
(Fassi Fehri 1999, p. 129)

be unvalued. In the current model, we must extend the notion of 'hybrid category' to include the possibility that one feature forces the other to remain unvalued rather than fixing its value. This does not seem a problematic step to take.

[12]We can think of such a hybrid head in the following way. Chomsky (1995) argues that functional heads are lexical elements. We can say that there are two Poss heads in the lexicon: one with a [+POSS] feature and one with a [–POSS] feature. The [+POSS] head has an additional unvalued DEF feature. Once [+POSS] has been selected, it is no longer necessary or even possible to select a D head, because the numeration already contains a DEF feature. Take the lexical items *man* and *men* as an analogy: both these elements have the same semantic features MAN, but the second has an additional [NUM: PL] feature.

[13]That is to say, an unvalued feature will be valued if the other element has the same feature *with* a value.

Fassi Fehri claims that construct state nominals can be used as predicates, just like indefinite nouns can, and argues that when they are, definiteness inheritance apparently does not take place. In (10), the noun *'aḫ-* 'brother' is a noun in construct state, the genitive modifier here is the pronoun suffix *-ī* 'my'. The form *'aḫ-ī* 'my brother' is used as a sentential predicate here. Fassi Fehri argues that only indefinite nouns can be used as predicates, and that consequently, *'aḫ-ī* in (10) must be indefinite. But I would argue that definiteness inheritance does take place in (10). There is no restriction against definite nouns appearing in a predicate position:

(11) hāḏā huwa -l-bayt-u
this it the-house-NOM
'this is the house'

In (11), an additional pronoun *huwa* is added in order to separate the subject and the predicate.[14] Structures like these are not typical predicational structures. Rather, they are equational in nature: two elements are equated, one of which is often, but not necessarily, a deictic element. It could well be that their structure differs from true predicational sentences, but even if it does not, there is obviously no reason to assume that (10) differs in any way from (11). Hence, (10) does not provide an example of a genitive construction without definiteness inheritance.

Fassi Fehri (1999) claims that definiteness inheritance is also absent in partitive constructions like the following:

(12) 'aḥad-u -l-riǧāl-i raǧaᶜa
one-NOM the-men-GEN returned
'one of the men returned'

Fassi Fehri claims that the head noun of this construction, *'aḥad* 'one' is not definite. He notes that the head noun cannot be modified by an adjective, whereas the modifier noun can:

(13) a. *'aḥad-u -l-riǧāl-i -l-ṭawīl-u raǧaᶜa
one-NOM the-men-GEN -the-tall-NOM returned
intended to mean: 'one of the tall men returned'
b. 'aḥad-u -l-riǧāl-i -l-ṭiwāl-i raǧaᶜa
one-NOM the-men-GEN the-tall.PL returned
'one of the tall men returned'

However, there is a mismatch in (13a): the adjective *ṭawīl* 'tall' modifies *'aḥad* 'one', which is not an element that can be modified by an adjective. Compare, for example, the equivalent English expression in (14):

(14) a. *tall many/many tall of the men
b. many of the tall men

[14]Without it, the phrase could easily be understood as meaning *this house*.

In (14a), an attempt is made to have the adjective *tall* modify *many*. In either order, this is not possible. The type of *many* is not such that it allows adjectival modification. The element *'aḥad* 'one' in the Arabic example is similar. Therefore, the ungrammaticality of (13a) is not caused by the lack of definite inheritance, but because of a type mismatch. In other words, there is no reason to assume that (13) does not involve definiteness inheritance.

Fassi Fehri (1999) also claims that certain types of deverbal nouns show no definiteness inheritance:

(15) a. 'iḏā^{c}at-u -l-ḫabar-i ^{c}amd-an ḫaṭa'-un
broadcasting.F-NOM the-news-GEN deliberately mistake-NOM
'deliberately broadcasting the news is a mistake'

b. *'iḏā^{c}at-u -l-ḫabar-i -l-mutasarri^{c}at-u ^{c}amd-an
broadcasting.F-NOM the-news-GEN the-hasty.F-NOM deliberately
ḫaṭa'-un
mistake-NOM
intended to mean: 'the hasty broadcasting of the news deliberately is a mistake'

(15a) contains an example with a deverbal noun, *'iḏā^{c}a* 'broadcasting', which is modified by an adverbial *^{c}amdan* 'deliberately'. The genitive complement *al-ḫabar* 'the message' is the object of 'broadcasting'. (15b) contains the same clause, but now the deverbal noun is also modified by an adjective, *mutasarri^{c}* 'hasty'. This phrase is ungrammatical.

Fassi Fehri argues that the ungrammaticality of (15b) shows that the head noun *'iḏā^{c}a* 'broadcasting' is not definite. This would be caused by the lack of definiteness inheritance. As a result, the adjective cannot agree with the head noun in definiteness, leading to ungrammaticality. There are two problems with this argument, however. First, the phrase with the adjective becomes grammatical when the adverb *^{c}amdan* is left out:

(16) 'iḏā^{c}at-u -l-ḫabar-i -l-mutasarri^{c}at-u ḫaṭa'-un
broadcasting.F-NOM the-news-GEN the-hasty.F-NOM mistake-NOM
'the hasty broadcast of the news is a mistake'

What this shows is that the ungrammaticality of (15b) is not due to the presumed absence of definiteness inheritance.[15] The second problem with Fassi Fehri's argument is that there is no restriction against a definiteness feature on a (complex) event nominal. Take the following example:

[15]Rather, it seems that the combination of an adjective and an adverb modifying the same noun is ungrammatical in Arabic. Generally, adverbs modify verbs, adjectives or other adverbs. They can also modify what Grimshaw (1990) calls "complex event nominals". As I argue in chapter 5, deverbal nouns in Arabic can be both complex event and simplex event/result nominals. If we assume that in (16), *'iḏā^{c}a* is a result nominal, whereas in (15), it is a complex event nominal, we explain the distribution of the adjective and the adverb.

(17) yuṭālibu min-nī -l-ḏahāb-a bi sur^{c}at-in
he.demands from-me the-going-ACC with speed-GEN
'he demands that I go quickly'

Here, the deverbal noun *al-ḏahāb* 'going' has the value of a subclause, and is furthermore modified by an adverbial. Both facts indicate that the deverbal noun is a complex event nominal.[16] Nonetheless, the nominal has a definite article. This is in fact the normal construction if the event nominal is not modified by a genitive-marked object or subject.

The cases for which Fassi Fehri (1999) claims that they show that definiteness inheritance takes place independently from genitive assignment do not hold. The facts still point to the conclusion that although definiteness inheritance and genitive assignment are two different processes, they do occur in conjunction. The D/Poss head is not marked for definiteness, but the goal with which this head agrees is, either because it has a definiteness feature of its own, or because it inherited one. In this constellation, only one thing can happen: when the probe values the case feature of the goal, its own definiteness feature is valued by the definiteness feature of the goal.

3.2 Derivation of nominal forms

So far, I have focused on the features DEF and POSS. However, Standard Arabic nouns are also marked for number and case, and feminine words are usually marked for gender. In this section, I will examine these other features in more detail in order to see if Giorgi & Pianesi's (1997) model can account for them.

3.2.1 Case, number and gender

The CASE feature on Arabic nouns can have three values: nominative, genitive and accusative. The NUMBER feature can also have three values: singular, dual and plural. GENDER has only two values: masculine and feminine. If we take Giorgi & Pianesi's model seriously, it follows that these features, CASE, NUMBER and GENDER can project heads.

The K head

The idea that CASE projects an independent head has already been put forward by Abney (1987), who argues for the presence of a K head.[17] Evidence to support this idea is found in languages that have case markers that appear as independent morphological elements. Bittner & Hale (1996) quote a few examples, of which I will give one. The following phrase is from Khasi:

[16] See chapter 5 for details.
[17] Abney chooses the letter K to indicate the CASE projection, because C is obviously taken.

(18) ka la yo"i *ya* 'uu khlaa
she PST see ACC the tiger
'she saw the tiger'

Bittner & Hale consider the element *ya* in (18) to be a case marker. That is, *ya* is not a preposition that licenses the argument as a last resort. Rather, it is a case element heading a case projection. Abney argues that the K projection dominates DP, which finds confirmation in (18), where the case element *ya* precedes the determiner, which in turn precedes the noun.

The Num head

Cross-linguistically, there is ample evidence that NUMBER can project a head Num. Take the Turkish example in (19):

(19) kitap-lar-ı
book-PL-POSS.3SG
'his books'
(Kornfilt 1997)

NUMBER appears as a separate suffix in (19), which provides evidence for a Num projection. There is also evidence for its existence in other languages: Ritter (1991) argues for a Num head based on Hebrew data, Szabolcsi (1994) also assumes a Num projection in Hungarian, and Bernstein (2001) argues that Num surfaces as a clitic in Walloon. All in all, the existence of a Num head is well established.

The Gen head

The idea that GENDER can project an independent head has been put forward by Picallo (1991). Bernstein (1993) argues that GenP should be replaced with a more general Word Marker Phrase, or WMP, which does not indicate gender per se but a nominal category in general.[18]

Ritter (1993) opposes the idea of a head Gen. She argues that the gender feature does not project independently, but is present on one of the other heads in the noun phrase. In Hebrew, for example, she argues that GENDER is present on the noun stem, whereas in Romance languages it is located on the head Num. For the moment, however, I will assume that GENDER can project an independent head Gen. At the end of the discussion, I will show how Ritter's proposal can be modified to fit the current analysis.

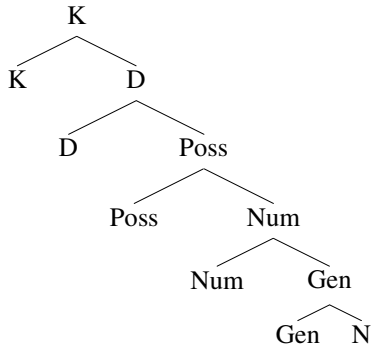
[18] Given the fact that there are languages, most notably Bantu languages, that have a much more elaborate categorisation of nouns, the term 'gender' is not entirely accurate, so replacing it with something like 'word marker' makes sense. However, because Arabic makes the same gender distinction that Romance languages do (i.e., a distinction in masculine and feminine), I will continue to use the more familiar term 'gender'.

3.2.2 Possible word forms

The preceding discussion shows that we have five functional heads in the noun phrase dominating the lexical projection N: K, D, Poss, Num and Gen. Each head in the noun phrase can be associated with a morpheme. Our model makes very clear predictions about the ways these morphemes can be combined into word forms. Let us see what these predictions are.

The standard assumption is that the features project in the order as indicated. We therefore have the tree in (20):

(20)



The fact that these heads project in a fixed order suggests that we will find that the morphological elements that represent these features have a fixed order as well. Assuming that heads precede their complements,[19] we would expect the morphological elements to appear in the order CASE-DEF-POSS-NUM-GEN-Noun.

It is possible to create other orders. If head movement takes place, the order of the elements will change. In Arabic, head movement can result in two distinct morphological processes: head adjunction and something we might call 'head merger'. In the case of head merger, the morphological elements of the two heads are discontinuous, and they merge to form a complete phonological form. For example, the form *riğāl* 'men' is composed of the stem consonants *r-ğ-l* and the vocalic plural marker *-i-ā-*.

Assuming that head adjunction takes place on the left,[20] the tree in (20) gives us several possibilities. If no head movement takes place, we predict the following order of morphemes:

[19] In the linearisation procedure that I outlined in chapter 2, this is not a given. As I will show, however, in Arabic heads do indeed precede their complements.

[20] If the linear ordering of syntactic structures is not fixed by UG, as I argue, then one may argue the same for head adjunction. In other words, we must allow for the possibility of head adjunction to take place on the right. It is safe to say, however, that head adjunction in Arabic takes place on the left, because Arabic does not have affixes that are prefixes. All affixes are either discontinuous, or they are suffixes. The verbal system does make use of person prefixes, but there is interference with gender and number: the first person prefixes are marked for number, and the third person singular prefix is marked for gender. These person markers always combine with a specific vowel pattern in the verb stem and with a suffix that indicates mood and in some persons also gender and number. Because the features gender and number are distributed over pre- and suffixes, we can argue that the person prefixes in these verb forms are not separate affixes, but are part of a circumfix or discontinuous affix.

(21) CASE-DEF-POSS-NUM-GEN-N

No head movement has taken place in (21), and consequently we expect there to be no affixes. If any of the elements CASE, DEF, POSS, NUM or GEN is marked with a morpheme, we expect this morpheme to be independent.

If head movement does take place, say from N to Num, we expect (22):

(22) CASE-DEF-POSS-NUM-[N-GEN]

Here, the combination [N-GEN] is either Noun-suffix or a merged form expressing Noun and GENDER. Further movements are also possible:

(23) a. CASE-DEF-POSS-[N-GEN-NUM]
b. CASE-DEF-[N-GEN-NUM-POSS]
c. CASE-[N-GEN-NUM-POSS-DEF]
d. [N-GEN-NUM-POSS-DEF-CASE]

These are the possibilities that head movement gives us. It should be noted that head movement cannot skip a head. This means that configurations such as in (24) are not possible:

(24) *CASE-DEF-[N-POSS]-NUM-GEN

That is, a configuration in which N has moved to Poss, skipping Num and Gen, is ungrammatical, and we expect it not to occur.

Another possibility is that two features are projected syncretically. For example, if POSS and NUM are projected syncretically, we would get:

(25) CASE-DEF-[POSS.NUM]-GEN-N

The difference that we expect between syncretic projection and head movement is that in the case of head movement we can still discern two morphemes, possibly discontinuous, whereas in the case of syncretic projection, there is only one morpheme that expresses two features. So in (25), POSS and NUM would be combined into one inseparable element.

3.2.3 Morphological markers in the Arabic noun phrase

All the features discussed in the previous section can be marked on Arabic nouns. Since the theory makes clear predictions about the way in which the features can be marked, we must see whether the Arabic noun phrase behaves as expected. When we examine how the features are expressed on Arabic nouns, it turns out that one feature is problematic: CASE. I will therefore propose to treat this feature separately.

Number markers

NUMBER in Arabic is generally not marked with a specific ending, but by the vowel pattern that is applied to the noun:

(26) a. al-kitāb; al-kutub
the-book; the-books
b. al-raǧul; al-riǧāl
the-man; the-men

In (26a), the vowel pattern *-i-ā-* indicates singular, whereas the pattern *-u-u-* indicates plural. In (26b), the patterns are different, but the principle is the same: the pattern *-a-u-* indicates singular, while *-i-ā-* indicates plural.[21] Plurals formed this way are called 'broken plurals'.

Although such broken plurals are very wide-spread in Arabic, there are also number suffixes in Arabic. The dual is always marked with a suffix, and there are also two plural suffixes, generally known as 'sound plurals':[22]

(27) a. al-kitāb; al-kitāb-āni
the-book; the-book-DUAL
'the book; the two books'
b. al-mu^{c}allim; al-mu^{c}allim-ūna
the-teacher; the-teacher-M.PL
'the teacher; the teachers'
c. al-mu^{c}allim-a(t); al-mu^{c}allim-āt
the-teacher-F; the-teacher-F.PL
'the (female) teacher; the (female) teachers'

(27a) shows the dual ending *-āni*. This ending can be applied to any count noun to form a dual. (27b) shows the so-called masculine sound plural *-ūna*, which is used for words that refer to male persons. (Although it is not the case that all nouns referring to a male person take this plural ending. Such nouns may also have a broken plural, cf. *riǧāl* 'men'.) (27c) shows the so-called feminine sound plural, which is applied to nouns referring to female persons, and also to quite a number of inanimate nouns.[23]

The Num head is the obvious place for the generation of the number marking.[24] In other words, Num is projected independently in Arabic, and we must conclude that N-to-Num movement takes place. In the case of broken plurals, we can say that the vowel pattern is generated in Num, whereas the consonantal root is generated in N

[21]Note that the same pattern may indicate singular in one noun but plural in another: *-i-ā-* marks singular in *kitāb* 'book' and plural in *riǧāl* 'men'. See also appendix A for a more elaborate discussion of the morphological system of Arabic.

[22]A word usually has only one plural form, either a broken or a sound plural, although there are words with more than one possible plural form. It should be noted that the sound plural suffixes cannot be applied freely.

[23]Inanimate nouns that take the sound feminine plural ending are not necessarily feminine. For example, the word *qiṭār* 'train' is masculine, but one of its plural forms is *qiṭār-āt*.

[24]Note that I do not assume that Num is the position of numerals. Num is the locus of number features, that is, singular, dual or plural. Numerals, on the other hand, are lexical elements, presumably of some N-like category, which have their own projections. The distinction becomes clear if we consider the fact that Arabic cardinals over 10 require the counted noun to be singular, which shows that the NUMBER feature and the cardinality of the noun are two distinct properties. For a discussion of numerals, see section 3.3.1.

and that N-to-Num movement combines the two.[25] The number suffixes *-āni, -ūna* and *-āt* are also generated in Num, after which N moves to Num and left-adjoins to the ending.

The Poss marker *-n*V

Markers for the feature POSS are quite rare in Arabic. The only case in which POSS is overtly marked is on the dual and masculine plural endings *-āni* and *-ūna*. These elements consist of a long vowel which indicates NUMBER and an additional suffix *-n*V (where V is either *-i* or *-a*) that indicates [–POSS]. To see how this works, take a look at (28):

(28) a. mu^{c}allim-ū-na
teacher-PL.NOM-UNPOSS
'teachers' (indefinite)

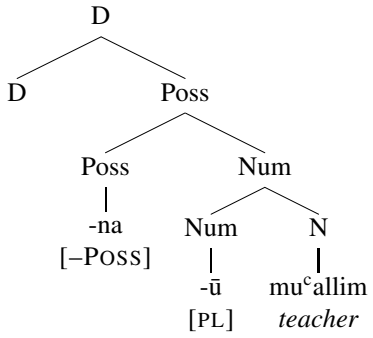
b. al-mu^{c}allim-ū-na
the-teacher-PL.NOM.UNPOSS
'the teachers' (definite)

c. mu^{c}allim-ū (-l-madrasat-i)
teacher-PL.NOM (the-school-GEN)
'the teacher of (the school)' (construct state)

What we see here is that the form with the suffix *-n*V appears in the indefinite in (28a) and in the definite in (28b). In other words, the element *-n*V appears in the non-possessed forms. In the possessed form, the construct state in (28c), it is dropped. The dual ending *-āni* works in the same way: in construct state, *-ni* is dropped and the ending becomes *-ā*.

The position of this element is rather easy to account for. Because it marks a non-possessed form, i.e. a form with the feature [–POSS], we must say that *-n*V is generated in the position of the Poss head. The noun, which has already moved to Num, moves on to Poss and adjoins to it. That is, the tree starts out as in (29):

(29)



[25] This assumption means that we must consider the singular pattern as a default pattern, because it is also present in the sound plurals and in the dual. This can be achieved with a system of Distributed Morphology (Halle & Marantz 1993): the root is generated in N and moves to Num if it is attracted by a number affix. If at the moment of spell-out the root is found to lack a vowel pattern, a default one is inserted.

Now, N moves to Num and left-adjoins to it to form *mu^{c}allim-ū*. After that, the N-Num complex moves to Poss and left-adjoins to it, to form *mu^{c}allim-ū-na* 'teachers'.

Gender markers

Arabic nouns are either masculine or feminine. Masculine nouns are not marked for gender, feminine nouns usually have the ending *-at*, although there are quite a few feminine nouns that do not have this ending. Furthermore, the plural endings are marked for gender: *-ū(na)* is masculine and *-āt* is feminine. As I said earlier, however, the masculine plural ending *ū(na)* only occurs on nouns referring to male persons, whereas the feminine plural ending *-āt* appears on nouns referring to female persons but also on quite a number of inanimate nouns, which are not necessarily feminine. In other words, it would seem that the gender markings on the plural endings are not syntactic but rather refer to a semantic or pragmatic gender distinction.

Note, however, that inanimate plural nouns in Arabic are treated as feminine singular. That is, they trigger feminine singular agreement on demonstratives, verbs and adjectives. For this to be possible, inanimate nouns must carry a syntactic [GEN: F] feature. Given this fact, we can say that inanimate nouns that are masculine in the singular and that form their plural with *-āt* are marked for feminine with this ending. In other words, the gender features on the plural suffixes *are* grammatical features. And since the only nouns that are masculine in the plural are animate nouns, the ending *-ū(na)* does not occur on inanimate nouns: the ending would incorrectly mark inanimate plurals as masculine.

The question that we must ask is whether GEN projects an independent head, and if so, in which cases it does so. The fact that the plural suffixes are marked for gender indicate that no Gen head is projected when these suffixes are present. Instead, we must consider *-ū(na)* and *-āt* as syncretic heads, projecting both NUMBER and GENDER.

In the singular and with broken plurals, masculine gender is never marked explicitly. This means that [GEN: M] never projects a Gen head. At first sight, the feminine singular seems to be different. The feminine marker *-at*[26] is an explicit marker of gender. However, this ending *only* occurs on singular nouns.[27] Therefore, we can say that *-at* is also a syncretic head, projecting both the features [NUM: SG] and [GEN: F].

In other words, GEN never projects an independent head in Arabic. Instead, it is present either on the noun stem or on Num. It seems then that Arabic corroborates Ritter's (1993) argument that GEN is always carried by another head, and does not project a head of its own. Below, however, I will argue that this conclusion is most likely too strong.

[26]Note that this marker has a short vowel, in which it differs from the feminine plural suffix, which has a long vowel. The two endings also differ in the strength of the phoneme /t/: the /t/ in the plural suffix *-āt* is never dropped, whereas the /t/ in *-at* is dropped before a pause or at the end of a clause.

[27]There are some exceptions. Some nouns referring to male persons have a plural that ends in *-at*, such as *sayyid* 'gentleman', which has a plural of *sādat* 'gentlemen'. However, in such cases the ending *-at* does not indicate feminine gender: such nouns are masculine. The ending *-at* is therefore not a projection of Gen, but simply a part of the plural marking.

Markers for definiteness

There are two markers for the feature DEF in Arabic: the definite article *al-*, which is prefixed to the noun, and the indefiniteness suffix *-n*. The indefiniteness suffix is the last suffix that a noun can take, that is, there is no suffix that can follow after *-n*. Most nouns can take *-n*, except nouns with a dual or masculine plural ending:[28]

(30) a. mu^{c}allim-ūna(*-n)
teacher-PL-(*INDEF)
'teachers'
b. mu^{c}allim-āni(*-n)
teacher-DU-(*INDEF)
'two teachers'

It is sometimes claimed that indefinite noun phrases are not DPs but NumPs. The indefinite article would then appear in the position of the Num head, which would account for the fact that indefinite articles are often derived from the numeral for *one*. This claim may hold for certain languages, but it is quite problematic for Arabic. First, the indefiniteness marker in Arabic *-n* is not derived from the numeral for 'one', which is *wāḥid*.[29] Second, the indefiniteness marker also occurs on plural nouns:

(31) a. riğāl-u-n
men-NOM-INDEF
'men'
b. mu^{c}allim-āt-u-n
teacher-F.PL-NOM-INDEF
'(female) teachers'

If *-n* were generated in Num, we would have no position for the plural vowel pattern *-i-ā-* or the plural marker *-āt* to be generated in. We must therefore conclude that *-n* is not generated in Num, but in D. With *-n* being generated in D, we can account for its position in a straightforward manner. The noun moves to D, left-adjoining to *-n*.[30]

As said before, definiteness is marked with the prefix *al-*. At first sight, the fact that *al-* is a prefix seems to contradict the argument that affixes in Arabic are always suffixes or merge with the noun stem. However, the definite determiner has some properties that suggest that from a morphological point of view it is a different kind of element. All the elements discussed so far show what we may call morphological interference: the form that a morphological marker takes can be influenced by a number of factors. For example, whether the plural feature is expressed with a plural suffix or

[28] And also a category of so-called diptotic nouns, which I discuss below in section 3.2.4.

[29] Moreover, I argued in footnote 24 that Num is not the position where numerals are found.

[30] Fassi Fehri (1993) argues that the indefiniteness marker *-n* and the suffix *-n*V that I analysed as a marker of [−POSS] are one and the same element. Although they may historically be related, it is obvious that they must be distinguished, simply because *-n* does not appear on definite nouns, whereas *-n*V does. The present analysis accounts for this difference.

with an altered vowel pattern depends on the specific noun;[31] whether [–POSS] is expressed with *-n*V depends on the number marker; whether indefiniteness is expressed with the suffix *-n* or remains unexpressed morphologically, is determined by the form of the number marker: if the noun has a dual or masculine plural ending, indefiniteness in not expressed morphologically;[32] and the head on which GEN is expressed depends on the value of the NUM feature and on whether the noun in question has an overt feminine marker or not.

We can account for this interference by saying that when a head moves and adjoins to or merges with another head, the morphological elements that express the features on those heads can influence each other in arbitrary ways. For example, there is no reason why [–POSS] should be expressed with *-n*V only in the context of a dual or masculine plural suffix other than that this is lexically specified. Similarly, the nouns that take a masculine or feminine plural ending rather than a broken plural are also lexically specified.[33] Similar interference is seen in the indefiniteness marker.[34]

The definite determiner, on the other hand, does not show such interference. Its form is always *al-*. The *-l-* does assimilate to the following consonant if this is an alveolar consonant, but this is a phonological process, not a morphological or syntactic one. Obviously, the fact that *al-* does not undergo any morphological interference does not prove that it is not an affix that is attached to the noun after head movement. There is no reason why an affix should be subject to morphological interference.

Even though it is not solid evidence, we may nonetheless take it as an indication that the noun does not move to D if D carries the feature [+DEF]. If we make this assumption, we can account for the fact that *al-* is prenominal, whereas the other morphemes are all suffixes or merged affixes. The prenominal position of *al-* follows from the fact that the noun is in the complement of D. Since a head always precedes its complement in Arabic,[35] the definite determiner precedes the noun. The fact that *al-* is always attached to the noun is accounted for by the fact that the determiner and the noun always end up adjacent in the linear string.[36]

There is another fact that suggests that N does indeed not move to D in Arabic. Longobardi (1994) argues that an empty D head receives a default existential interpretation. In cases where this existential interpretation is not desirable, D has to be

[31] Some nouns even have more than one possible plural form, and there are also plural forms that combine a broken plural with the ending *-āt*. E.g. the noun *qiṭār* 'train' has as plural forms *qiṭār-āt, quṭur* and *quṭur-āt*.

[32] Furthermore, there is a class of nouns in which indefiniteness is expressed in yet another way. See section 3.2.4 for details.

[33] This is not to say that this lexical specification is completely arbitrary. There are certainly generalisations. For example, participles used as nouns that refer to human beings usually get plural endings rather than broken plurals: the noun *muᶜallim* 'teacher' is a participle, and takes the plural *muᶜallim-ūna*. Such generalisations are never absolute however: *ᶜāmil* 'worker', which is also a participle, has a broken plural *ᶜummāl*.

[34] It should be noted that this kind of interference is different from the agreement in case that some languages show, such as Classical Greek in (34) below. Morphological interference is often arbitrary, and can involve any kind of morphological or semantic category of the elements in question. Agreement can only take place in a predictable manner and only involves syntactic features.

[35] I will argue for this in section 3.4.

[36] In section 3.4 I will show how we derive this fact.

filled. This can be done by generating an overt determiner in D, or by moving N to D. Generic nouns do not have an existential interpretation, which means that D must be filled. In Romance, this is generally achieved by generating an overt determiner. Longobardi argues that in Germanic, generic nouns move to the D position at LF. If Arabic nouns move to D in general, we would predict that there is no need to generate an overt determiner in D with generic nouns. However, in Arabic, such nouns generally appear with a definite determiner:

(32) al-kuḥūl-u ḍārr-un bi -l-ṣiḥḥat-i
the-alcohol-NOM damaging-NOM with the-health-GEN
'alcohol is bad for one's health'

The D position of *al-kuḥūl* is filled by *al-* in (32). This means that N-movement is apparently not available to cancel the default existential interpretation. We conclude, then, that N-to-D movement does not take place.

3.2.4 Case as a "roaming" feature

So far, we have seen how we can account for the morphological markers of most of the features in the Arabic noun phrase. Each feature has its own position in the functional structure, and nominal forms are created by head adjunction or head merger. One feature we have not yet discussed is the feature CASE. When we look at the way CASE is marked on Arabic nouns, we see that there is no clear system:

(33) a. al-raǧul-ā-ni
the-man-DUAL.NOM-UNPOSS
DEF-N-[NUM.CASE]-POSS
'the two men'
b. riǧāl-u-n
man.PL-NOM-INDEF
[N.NUM]-CASE-DEF
'men'
c. ṣaḥrā'-u
desert.SG.F-NOM.INDEF
[N.NUM.GEN]-[CASE.DEF]
'a desert'

In (33), the line below the gloss indicates the order in which the features are expressed. Features that are expressed on one morpheme have been grouped together with square brackets. In (33a), CASE is marked on the number ending. That is, the suffix *-āni* not only indicates dual and non-possessed, but also nominative. The genitive/accusative form is *-ayni*. The same is true for the masculine plural ending *-ūna*, of which the oblique form is *-īna*. In other words, the dual and masculine plural markers actually carry three features: one for NUMBER, one for CASE, and one for POSS. In Giorgi & Pianesi's (1997) theory they could therefore be considered syncretic categories, if NUM and CASE were to project consecutively. Unfortunately, they do not:

the features POSS and DEF project after NUM, but before CASE. We obviously do not have a syncretic head of four features, because the feature [+DEF] projects independently as *al-*. Therefore, we have a problem: we cannot explain why CASE appears on the number suffix.

(33b) also poses a problem. CASE is marked with a separate suffix, but the position it appears in is problematic. I have argued that the indefiniteness suffix *-n* is generated in D and that N moves to D and adjoins to the suffix *-n*. This means that any suffix that appears between the noun stem and *-n* must have adjoined *before* N moves to D. In other words, the case suffix must have been generated in a head dominated by D. But the general assumption is that K dominates D, not the other way around. Therefore, the order N-CASE-DEF should not occur.

We are forced to the conclusion that the CASE feature in Arabic behaves differently from the other features. Let us say that CASE in Arabic is a "roaming" feature: it does not project a head of its own, but is instead present on another head, and it finds morphological expression when the derivation is spelled out. In a system of distributed morphology (see, e.g., Chomsky 1998, and Halle & Marantz 1993), we can say that at spell-out, the correct morphological form of N is chosen from the lexicon. The form that is chosen reflects the features on N as closely as possible.

What I will argue specifically is that CASE in Arabic is present as a feature on the noun. When a functional head takes the noun as its complement, it inherits all features of the noun, including the case feature. Because the case feature is now present on that head, it can be morphologically expressed on it. This analysis is supported by the fact that in some languages case is expressed in more than one position. Take the following example from Classical Greek:

(34) ho anēr; tou andros
the.NOM man.NOM; the.GEN man.GEN
'the man; the man's'

In Classical Greek, case shows up on the definite article and on the noun. There does not seem to be a straightforward explanation for this if we assume that the case feature is present on a separate K head. On that assumption, we would have to explain how the case feature can come to be expressed on both D and N. D could perhaps be argued to move to K, or K and D could project a syncretic head, but that does not account for the case ending on the noun. With the current proposal, we can account for it: the case feature does not project an independent head, but it is present on N and inherited by D.

Obviously, it is not the case that any head inherits the features of its complement. There must be a 'boundary' across which features are not inherited. The most natural boundary that comes to mind is the phase. Within a phase, features of the complement are inherited, but once the phase level has been reached, this inheritance stops. In other words, a DP will inherit the features of the noun it dominates, but the VP that takes this noun as a complement does not.

In Arabic, the same thing is taking place: CASE is present on N and is inherited by the other heads in the noun phrase.[37] At spell-out, the feature is present in the feature matrix, and can be spelled out on any head. In most cases, it is spelled out on the noun with one of the suffixes *-u, -i* or *-a*.[38] In the case of the masculine sound plural and the dual ending, it is spelled out on Num.

In the case of the feminine sound plural ending, CASE is spelled out on Num as well. This suffix has the form *-āt*, which is in itself not marked for case, but it takes a different ending for the accusative: *-i* instead of the usual *-a*. This variant accusative ending only occurs on the ending *āt*, something we account for if we say that case is on the Num head, together with the suffix *-āt*.[39]

There is another type of case marker. There are some nouns that do not take the indefiniteness suffix to mark indefiniteness. Instead, they take different case endings. Such nouns are traditionally known as diptotic noun.[40] (35a) is the common case, in which the noun takes *-n* to mark indefiniteness. (35b) is the exceptional case, where the case markings change:

(35) a. mu^{c}allim-u-n; mu^{c}allim-i-n; mu^{c}allim-a-n
teacher-NOM-INDEF; teacher-GEN-INDEF; teacher-ACC-INDEF

a′. al-mu^{c}allim-u; al-mu^{c}allim-i; al-mu^{c}allim-a
the-teacher-NOM; the-teacher-GEN; the-teacher-ACC

b. ṣaḥrā’-u; ṣaḥrā’-a; ṣaḥrā’-a
desert-NOM.INDEF; desert-GEN.INDEF; desert-ACC.INDEF

b′. al-ṣaḥrā’-u; al-ṣaḥrā’-i; al-ṣaḥrā’-a
the-desert-NOM; the-desert-GEN; the-desert-ACC

As one can see, diptotic nouns have a different genitive ending when they are indefinite. Although there is only one case ending that differs, I will argue that the case endings actually mark indefiniteness in (35b): what we see as case markers are in fact indefinite markers that are also marked for case. In other words, the endings *-u, -a, -a* in (35b) are projections of D which have inherited the CASE feature from the noun.

[37] This is not to say that CASE cannot project its own head. In certain languages, it does. It is just that it does not do so in Arabic.

[38] This means that these suffixes are not adjoined to the noun in syntax, but in the lexicon.

[39] I analyse the plural and dual suffix *-ūna* and *-āni* as being composed of a morpheme *-ū* or *-ā* indicating both CASE and NUMBER followed by a morpheme *n*V indicating [−POSS]. Traditionally, Arabic grammarians argued that the vocalic element *-ū* is composed of a case marker *-u* and a number marker in the form of a glide (i.e., the semi-vowel /w/), which effectively lengthens the vowel. If one were to adopt this analysis, one might argue that the case marker *-i* for accusative, which I claim only occurs on the feminine plural suffix, also occurs on the masculine plural suffix: the accusative form of the masculine plural suffix is *-ī(na)*, which on this analysis is composed of a case marker *-i* and also a glide, in this case /j/, indicating plural. Although the analysis would not be a problem for the theory that I develop here, I will not adopt it, because it is rather problematic for the dual. For the dual one would have to argue the exact opposite: the vowel *-a* would be the number marker rather than the case marker, and the glide (/ʔ/ for the nominative and /j/ for the oblique forms) is the case marker, rather than the number marker.

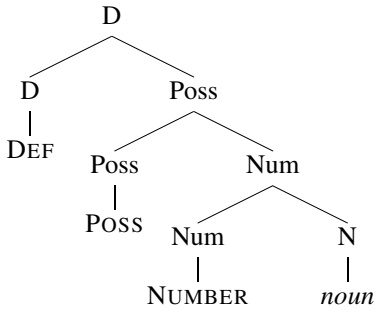
[40] They are called ‘diptotic’ because they only have two distinct case endings when they are indefinite: *-u* and *-a*. This contrasts with the so-called triptotic nouns, which have three case endings: *-u, -i* and *-a*.

With this, we have established an account that includes all the suffixes and merged affixes on the Arabic noun. All the features except for CASE and GEN project heads. In most cases, N moves to Num to pick up its NUMBER marker. It can move on to Poss and D, if these heads contain morphological elements that require such movement. CASE is present as a feature in the noun phrase, but it does not project independently. Instead, it is present on N and inherited by the heads that dominate it. GEN only projects a syncretic head with Num, or it is covertly present on N. Ritter (1993) argues that GEN never projects a Gen head, but given the fact that CASE does project a separate head in certain languages, we cannot conclude that GEN never does simply on the basis that it does not do so in Romance and Semitic. All we can say is that Gen is not projected independently in these languages.

3.2.5 Summary

In this section, I have developed the following basic structure for the Arabic noun phrase:

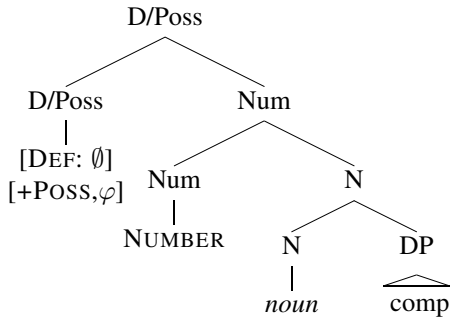
(36)



In this tree, D projects the definiteness feature, and Poss projects the possessive feature. If the noun does not have a complement, and Poss consequently has the feature [–POSS] and no φ-features, Poss and D are not projected syncretically, but scattered. In this case, Poss does not have an overt reflex except in the dual and masculine plural suffixes.

The head noun of the DP can take another DP as a complement. This DP is licensed by the Poss head. In this case, Poss has the feature [+POSS] and it also has a set of unvalued φ-features. Furthermore, it projects a hybrid category together with D:

(37)



D in this case has an unvalued DEF feature, which means it has to be valued during the derivation. Because D/Poss is a hybrid category, this valuation can take place in the same Agree operation that values the φ-features of Poss. This Agree operation also values the CASE feature of the complement noun, which is assigned genitive.

The fact that the assignment of genitive takes place in the manner described means that genitive is a structural case, just like nominative and accusative. This is corroborated by the fact that the theta role of the genitive complement of the noun is not fixed, but can vary, like the theta role of the nominative and accusative marked arguments of the verb.

Apart from D and Poss, NUMBER also projects a separate head. This head hosts the cardinality feature of the noun. The CASE feature does not project an independent head, in spite of the fact that case is overtly marked on Arabic nouns. Instead, CASE is a roaming feature: it is present on the noun and is passed on to the heads that dominate N, up to the phase level. As such, it can be expressed morphologically on any head.

In the noun phrase, there is N-to-Num movement. Whether N moves on to Poss and D depends on the morphological elements that express the features. Most of these elements are affixal in nature and require head movement of N. The exception to this is the determiner *al-*, which is an independent element that cliticises onto the noun and does not require N-to-D movement.

The only question that remains open for the moment is whether N-Num moves on to D/Poss in the construct state, i.e. if N has a complement that is licensed by the syncretic head D/Poss. This question will not be answered until chapter 5.

3.3 Modifiers in the noun phrase

Noun phrases do not only contain (possessive) complements. They can of course also contain a number of other modifiers, such as adjectives, numerals, quantifiers and relative clauses.[41] As Fassi Fehri (1999) shows, most of these modifiers (with the exception of relative clauses) can occur both before the head noun and after it in Arabic, although the unmarked position differs from modifier to modifier. There is a remarkable difference between the postnominal and the prenominal modifiers,

[41] I will use the term "modifier" in a pre-theoretic sense, referring to any constituent in the noun phrase besides the head noun.

however: prenominal modifiers are heads that take the noun as a complement, whereas postnominal modifiers are full projections that appear in specifier positions in the noun phrase. In this section, I discuss each of these modifiers, and present evidence showing that they are heads when they occur prenominally, and specifiers when they occur postnominally.

3.3.1 Adjectives

The most common position for adjectives is after the noun. Adjectives in that position agree with the noun in gender, number, case and definiteness. Plurals of inanimate nouns are always treated as feminine singular, which means that they trigger feminine singular agreement on adjectives.[42] This is demonstrated in the following examples:

(38) a. al-sayyār-at-u -l-ḥamrā'-u
the-car-F-NOM the-red.F-NOM
'the red car'
b. al-kitāb-u -l-'aḥmar-u
the-book.M-NOM the-red.M-NOM
'the red book'
c. al-kutub-u -l-ḥamrā'-u
the-books.M.PL-NOM the-red.F.SG-NOM
'the red books'

(38) shows two singular nouns modified by the adjective *'aḥmar* 'red'. *Sayyāra* 'car' in (38a) is feminine, and the adjective has the corresponding feminine form, whereas *kitāb* 'book' in (38b) is masculine, with the adjective in masculine form. (38c) has an example of an inanimate plural. The adjective here takes feminine singular form, whereas the noun *kitāb* is masculine.

Animate plurals take plural form of the adjective. There are distinct plural forms for masculine and feminine. Consider the following examples:

(39) a. al-raǧul-u -l-ṭawīl-u
the-man-NOM the-tall.SG-NOM
'the tall man'
b. al-riǧāl-u -l-ṭiwāl-u
the-men-NOM the-tall.PL-NOM
'the tall men'
c. al-nisā'-u -l-ṭawīl-āt-u
the-women-NOM the-tall-F.PL-NOM
'the tall women'
d. al-hunūd-u -l-ḥumr-u
the-Indians-NOM the-red.PL-NOM
'the American Indians'

[42] They also trigger feminine singular agreement on other modifiers, and on verbs.

Fassi Fehri (1999) argues that adjectives can also occur prenominally:[43]

(40) a. ’ukinnu la-hu wāfir-a -l-iḥtirām-i
I.entertain for-him abundant-ACC the-respect-GEN
‘I have much respect for him’
b. qara’tu ǧadīd-a -l-kutub-i
I.read new.SG.M-ACC the-books-GEN
‘I read the new (of the) books’
c. yaḥduṯu hāḏā fī muḫtalif-i -l-mayādīn-i
happens this in various.SG.M-GEN fields-G
‘this happens in various fields’
(Fassi Fehri 1999)

When looking at the construction in (39) and (40), the first thing to note is that it is not at all similar to the structure of prenominal adjectives in Germanic languages. There is no agreement between the adjective and the noun: the adjective has a default masculine singular form, regardless of the number and gender of the noun:

[43]It should be noted that this construction is not equivalent in meaning to a construction in which the adjective follows the noun. As El-Ayoubi et al. (2001) say, the meaning of the adjective usually changes when it is used prenominally:

> “(...) Gegenüber der attributiven Verbindung des relativen Adjektivs mit dem Kernnomen bringt die Stellung im Vorfeld einen höheren Grad der relativen Eigenschaft in bewertender Weise zum Ausdruck, sofern nicht überhaupt eine deutlichere Bedeutungsverschiebung eintritt.”
> (El-Ayoubi et al. 2001, p. 157)

> “(...) The relative adjective’s positioning in the prenominal field of the head noun expresses a higher grade of the relative property, compared to the attributive connection, unless it causes a more distinct shift of meaning altogether.”

Ayoubi et al. give the following examples to illustrate the point:

(i) a. ma^{c}a ǧazīl-i -l-šukr-i
with abundant-GEN the-thanks-GEN
‘with the greatest (of) thanks’
a′. šukr-an ǧazīl-an
thanks-ACC abundant-ACC
‘many thanks’
b. sābiq-u ’inḏār-i-n
preceding-NOM warning-GEN-INDEF
‘a fore-warning’ (lit. ‘the preceding of a warning’)
b′. ’inḏār-u-n sābiq-u-n
warning-NOM-INDEF preceding-NOM-INDEF
‘a previous warning’ (lit. ‘a preceding warning’)
(SASG p. 157)

The examples in (i) show a prenominal adjective construction and its postnominal counterpart. As can be seen, the meaning of the prenominal and the postnominal constructions differ subtly. In (ia) the prenominal adjective expresses a higher grade than the more usual postnominal one in (ia′). Furthermore, (ib) refers to a forewarning, whereas (ib′) simply refers to a previous warning.

(41) a. wāfir-u -l-ṯiq-at-i
abundant.M.NOM the-trust-F-GEN
'much trust'
b. ǧadīd-u -l-kutub-i
new.SG.M-NOM the-books-GEN
'the new books'

The noun *ṯiqa* 'trust' in (41a) is feminine, but the prenominal adjective is still masculine. In (41b), the noun *al-kutub* 'the books' is an inanimate plural, which means it triggers feminine singular agreement, but the adjective *ǧadīd* 'new' is still masculine singular. So we see that there is no agreement between a prenominal adjective and the noun.

Another thing to note is that the adjective in these constructions always appears in construct state: it is not marked for definiteness. Furthermore, the noun always has genitive case. The case that is assigned from outside the phrase shows up on the adjective, not on the noun.

In other words, the construction is a standard genitive construction, with the noun in the position of genitive complement and the adjective in the position of the head. This conclusion is supported by the fact that prenominal adjectives cannot take any other genitive complement. For example, postnominal adjectives allow a construction as in (42):

(42) raǧul-un kabīr-u -l-sinn-i
man-NOM large-NOM the-age-GEN
'an old man'

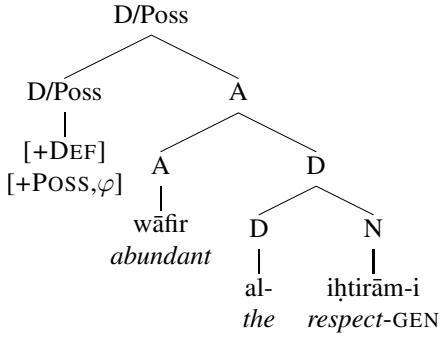
In (42), the adjective *kabīr* 'large' takes the noun *al-sinn* 'the age' as syntactic complement, forming the collocation 'large of age', meaning 'old'.[44] This complex, however, can only be used postnominally. A structure like (43) is impossible:

(43) *kabīr-u -l-sinn-i raǧul-un/in
large-NOM the-age-GEN man-NOM/GEN

We must conclude, then, that prenominal adjectives in Arabic are in fact the heads of the construction, and the nouns are their complements. The two occur in a standard genitive construction:

[44]Note that this structure is different in meaning from the one encountered above in (40), although superficially, they are much alike. The difference is essentially that the structure of (40) has the distribution of a noun phrase, whereas the structure in (42) has the distribution of an adjective phrase. See chapter 4 for a discussion of the structure in (42).

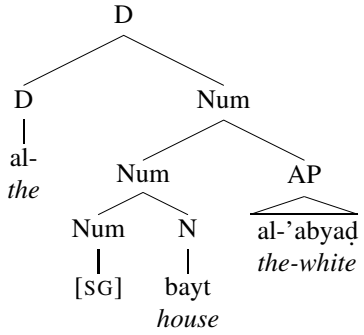
(44)



This tree is identical to the tree in (9), except for the fact that the lexical head of the phrase is not a noun but an adjective.[45] What this means, then, is that in the case of prenominal adjectives, the DP is not a projection of the noun, but of the adjective.[46]

Postnominal adjectives are different. Unlike prenominal adjectives, they agree fully, as we have seen above, and they are not heads. Instead, they are full projections of their own: postnominal adjectives are maximal projections, and must therefore be located in specifier positions in the noun phrase:[47]

(45)



As one can see, the construction in (45) is markedly different from the prenominal adjective in (44). The DP is now not a projection of the adjective, but of the noun, and the adjective is now an adjunct in the noun phrase.[48]

[45]In chapter 4 I will analyse the adjective phrase in more detail. It will turn out that the structure of the noun phrase and the adjective phrase are very similar, but not entirely identical. In this respect, (44) is not entirely accurate, but the differences that exist are not relevant to the point at hand.

[46]This also explains the fact that prenominal adjectives often have a different meaning, as remarked in footnote 44.

[47]Actually, Kayne (1994) proposes a different analysis for adjectives, in which they are not specifiers of the noun, but are still maximal projections. I do not adopt this analysis, for reasons I will discuss in chapter 4.

[48]Throughout this thesis, I will assume that adjectives are adjoined to the Num projection, although it is quite likely that at least some adjectives adjoin to other projections. Nothing in the discussion really hinges on this, so I will not go into this question here. For ease of exposition, I have adjoined the AP to the right, because the adjective appears after the noun in Arabic. As I explained in chapter 2, in a bare phrase structure approach such as the one I assume, trees are not linearly ordered, which means that simply adjoining the AP to the right cannot explain the linear ordering of the phrase. In section 3.4.2 I will show how we derive

3.3.2 Numerals

Numerals, both cardinals and ordinals, show a pattern that is very similar to adjectives. Take the following examples:

(46) a. al-dars-u -l-rābi^{c}-u
the-lesson-NOM the-fourth-NOM
'the fourth lesson'
b. al-buyūt-u -l-ḫams-at-u
the-houses-NOM the-five-F-NOM
'the five houses'

In (46a), the noun *dars* 'lesson' is modified by the ordinal *rābi^{c}* 'fourth'. As in the case of the adjective, we see full agreement in gender, number, case and definiteness. With the cardinal in (46b), the matter is basically the same, although a cardinal like *five* is of course inherently plural.[49]

Both cardinals and ordinals can also appear prenominally:

(47) a. ṯālit-u marr-at-i-n
third.M.NOM time-F-GEN-INDEF
'the third time'
b. 'arba^{c}-at-u kutub-in
four-F.NOM books.M.GEN
'four books'

With the ordinal in (47a) we see the pattern that is familiar from adjectives. The ordinal takes the default masculine singular form, and the noun takes genitive case. The case of the ordinal is assigned from outside the phrase. Unlike the prenominal adjective, however, the noun is indefinite in this construction.[50]

The example in (47b) shows the prenominal cardinal. Note that the meaning of this construction differs from the postnominal cardinal in (46b). In fact, (46b), with the postnominal cardinal, is the only way to express '*the* five books': the prenominal cardinal is incompatible with a definite reading, which means that the noun in (47b) cannot have an article.

The important thing to note is that the numerals, like adjectives, are heads that take the noun as a complement when they are prenominal, whereas postnominally, they are maximal projections that stand in specifier positions.

For ordinals, this point is quite easy to make. They essentially behave like adjectives: prenominally, they show the by now familiar genitive construction, with the ordinal in construct state and the noun in the genitive. There is no agreement between the two elements. Postnominally, the ordinal agrees in the familiar way. Although

the correct linear ordering.

[49]It should be noted that the feminine form of the cardinal is *not* the result of the fact that inanimate plurals are treated as feminine singular. Cardinals from 3 through 10 (and the units in 13–19) show a strange phenomenon called *polarisation*. If the modified noun is masculine, the cardinal takes feminine form, and vice versa. Because the singular *bayt* 'house' is masculine, the cardinal is feminine.

[50]This is remarkable, especially when one considers the fact that the meaning of the phrase as a whole is always definite: (47a) means '*the* third time', not '*a* third time'.

there is no evidence similar to that in (42) for adjectives, it is safe to assume that postnominal ordinals are maximal projections and appear in specifier positions, like adjectives.

For the cardinal, it is somewhat more difficult to make the point that it falls in the same pattern as the adjective and the ordinal. The prenominal cardinal is presumably not a normal genitive structure,[51] but we do see that the cardinal in (47b) imposes its requirements on the noun: the cardinal determines the case of the noun, which is genitive in this case. Furthermore, the case assigned to a noun phrase containing a prenominal cardinal shows up on the cardinal.[52] Given the standard assumption that a noun is a syntactic dependent of the element from which it receives case, and that the syntactic head of a noun phrase is the element that carries the case that is assigned to the noun phrase, we can conclude that the prenominal cardinal in (47b) is the syntactic head of the structure and the noun is a complement of it.

For the postnominal cardinal, we can make largely the same point that we made for the postnominal ordinal: it agrees with the head noun in the familiar way, which means that the cardinal is a maximal projection in a specifier position of the noun phrase.[53]

3.3.3 Quantifiers

Quantifiers[54] often appear prenominally, and take the noun as a complement, in construct state:[55]

(48) kull-u -l-ṭullāb-i
all-NOM the-students-GEN
'all (the) students'

Again, we see the by now familiar structure: the quantifier takes case according to the position of the noun phrase in the sentence, and the noun itself invariably takes genitive case. Furthermore, the quantifier does not have a definite article, nor does it

[51] Cardinals in Arabic (and in Semitic languages in general) show a rather complicated pattern. The cardinals 3–9 take an indefinite genitive plural complement, but cardinals between 11 and 99 do not take a genitive plural noun but an accusative singular, and hundreds, thousands etc. take a genitive singular. Because of this variety, it is difficult to determine what the exact structure is.

[52] If it shows up at all: the cardinals 11 and 13–19 are invariable.

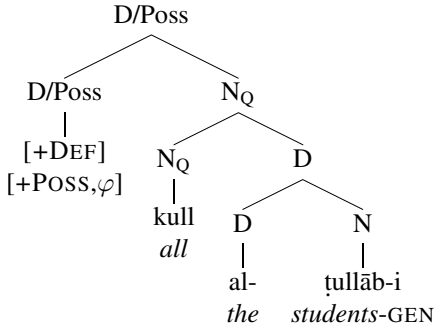
[53] It should be noted that the prenominal cardinal also shows agreement, or rather polarisation, with the noun. I will not take this agreement as evidence of XP-hood, because it is limited to agreement in gender. The agreement that shows up on postnominal numerals is agreement in gender, case and definiteness. (Number being excluded for obvious reasons.) In chapter 4, I will show that agreement in definiteness and case is a different process from agreement in φ-features, and that it can only take place if the agreeing element is a full projection. The fact that prenominal cardinals lack agreement in definiteness and case is a strong indication that they are not specifiers. I assume that the polarisation that we see is basically agreement in φ-features which is part of the case assignment process, as we have seen for T and Poss. See also section 3.3.3.

[54] Note that I use the term 'quantifier' rather loosely here, to refer to such elements as *most, all, some, few, each, every* etc. That is, I do not use it in a type-semantic sense, in which 'quantifier' refers to an expression of a specific type.

[55] Some quantifiers can also occur in different constructions, such as *qalīl min* 'a little of'. I will not discuss these here.

have an indefiniteness marker. The tree structure of (48) is rather straightforward, and we have seen it already for the normal construct state and the prenominal adjectives:

(49)

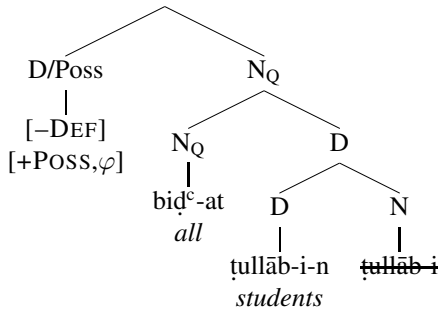


As in the case of the prenominal adjective, the prenominal quantifier *kull* in (49) does not agree with the noun in any way. However, there are some quantifiers, notably *biḍ^c* 'some, several', that show polarisation in gender:

(50) a. bid^c-at-u ṭullāb-i-n
some-F-NOM students.M-GEN-INDEF
'some students'

b. bid^c-u ṭālib-āt-i-n
some-M-NOM student-F.PL-GEN-INDEF
'some (female) students'

c.



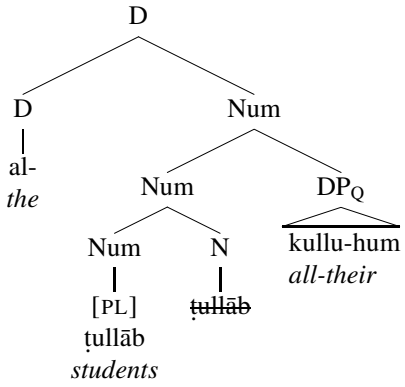
In (50a), the quantifier *bid^{c}at* 'some' has a feminine form, whereas the quantified noun *ṭullāb* is masculine. In (50c), matters are reversed: the quantifier has a masculine form, whereas the quantified noun is feminine. This agreement process is easily explained: D/Poss has a set of unvalued φ-features. In most cases, these φ-features do not show up overtly on the head noun, but apparently in the case at hand, the gender feature does.[56]

[56]Why the gender feature should polarise, instead of simply agree, is not clear. A trivial solution would be to say that what looks like the feminine form, because of the suffix *-at*, is actually the masculine form, and vice versa.

Some of the quantifiers that appear prenominally can also appear postnominally. In that case they take a suffix pronoun that refers back to the head noun:[57]

(51) a. al-ṭullāb-u kull-u-hum
the-students-NOM all-NOM-their
lit. 'the students, all of them'

b.



Roughly, (51a) means 'the students, the totality of them'. It has the same value as the phrase in (48), but its structure is different: the quantifier phrase is a full projection, not a head, and appears in adjoined (i.e. specifier) position.

3.3.4 Demonstratives

The noun-phrase modifiers discussed so far all have very similar structures: prenominally, they are heads heading their own projection, in which the noun is a complement. Postnominally, they are in adjunction (i.e. specifier) positions in the noun phrase of the noun they modify.

Demonstratives in Standard Arabic also occur both prenominally and postnominally. There are, however, some essential differences between demonstratives and the other modifiers, which means that we have to analyse them in a different way. The most important difference is that prenominal demonstratives agree with the head noun in number and gender:

(52) a. hāḏā -l-bayt-u
this.M the-house.M-NOM
'this house'

b. hāḏihi -l-sayyār-at-u
this.F the-car-F-NOM
'this car'

[57] Note that although the translation uses an apposition structure, this is presumably not the structure in Arabic, since the intonational pattern is different.

Furthermore, the dual of the proximity demonstrative also agrees in case:[58]

(53) a. hāḏ-āni -l-bayt-āni
these.M-DU.NOM the-house.M-DU.NOM
'these two houses' (NOM)
b. hāḏ-ayni -l-bayt-ayni
these.M-DU.GEN/ACC the-house.M-DU.GEN/ACC
'these two houses' (GEN/ACC)

In (53a) the noun *bayt* 'house' has dual number and nominative case, in (53b) it also has dual number, but genitive or accusative case.[59] The demonstrative *hāḏā* agrees with the noun in number, gender and case, as is shown by these examples.

With the other modifiers that I discussed above, the fact that the modifier and the noun do not agree in case is evidence for the structure that I proposed: the modifier takes the noun as a complement. Because the demonstrative shows agreement in case, we are obviously dealing with a different construction here. With the modifiers discussed above, I argued that they are specifiers if they are postnominal, because postnominal modifiers can project fully and because they show agreement not only in φ-features but also in case and definiteness. The fact that the prenominal demonstrative shows case agreement therefore suggests that it is a specifier as well.

There is a problem with this analysis, however. A prenominal demonstrative requires the following noun to have the definite determiner *al-*. If *al-* is absent, e.g. with the head noun of a genitive construction, the demonstrative cannot be used prenominally:

(54) hāḏihi sayyār-at-u -l-raǧul-i
this.F car-F-NOM the-man-GEN
*'this car of the man'

(54) is ungrammatical in the intended reading.[60] The only way to get the intended reading is to put the demonstrative after the genitive construction:

(55) sayyār-at-u -l-raǧul-i hāḏihi
car-F-NOM the-man-GEN this.F
'this car of the man' (lit. 'the man's this car')

In (55), the demonstrative modifies the head noun *sayyāra* 'car', not the genitive noun *raǧul* 'man', as can be seen from the gender agreement. In other words, a postnominal demonstrative is located in a similar position as a postnominal adjective: both follow the genitive modifier, even if they modify the head noun of a genitive construction. Let us say that the postnominal demonstrative is an adjoined specifier, just like the postnominal adjective.

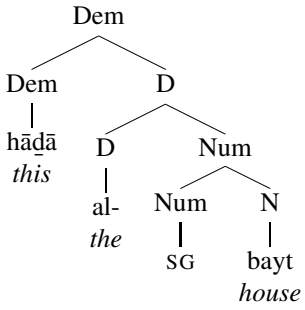
[58]The reason that the singular and plural demonstratives do not agree in case is probably phonological in nature. The demonstratives end in vowels, and case endings are vowels as well. Nouns that end in a (long) vowel, such as the proper name *mūsā* 'Moses' or a noun like *daᶜwā* 'claim' do not take case endings, either.

[59]There is one ending for both cases.

[60](54) does have a grammatical reading. It can be construed as a clause meaning 'this is the man's car'.

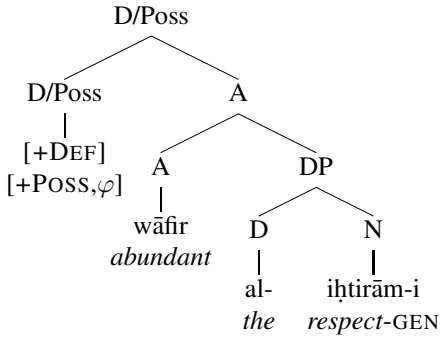
If the postnominal demonstrative is a specifier, it seems problematic to claim that the prenominal demonstrative is, too. Instead, we can say that the prenominal demonstrative is like the other prenominal modifiers: it is a head that takes the noun as a complement. The difference with the other modifiers is that the demonstrative is part of the same phase as the noun phrase, whereas the other prenominal modifiers, adjectives, numerals and quantifiers, are part of a different phase. The structure of the prenominal demonstrative is (56):

(56)



Compare this to the prenominal adjective structure of (44), repeated here as (57):

(57)



Unlike the demonstrative, the adjective heads a DP of its own. The complement DP in (57) is a phase, and the adjectival DP, which is the D/Poss projection, is a different phase. In (56), however, the DemP is simply an extra layer in the DP phase.

The agreement that shows up on prenominal demonstratives can be accounted for straightforwardly in this analysis. Remember that we concluded in section 3.2.4 that there is feature inheritance inside a phase. The demonstrative in (56) gets its features for case, gender and number through this mechanism from its complement D.

3.3.5 Relative clauses

Relative clauses in Arabic only occur postnominally. A relative clause is marked with a special marker *allaḏī*, which agrees with the head noun in gender, number and, in the case of the dual, also in case. The relative clause contains a resumptive pronoun that refers back to the head noun:

(58) al-kitāb-āni -llaḏāni katabtu-humā
the-books-DUAL.NOM REL.DUAL.NOM.M I.wrote-them
'the two books that I wrote'

In (58), the relative marker *allaḏāni* is masculine, dual and nominative. The resumptive pronoun is the object pronoun suffix *-humā*. Note that the case of the relative marker is different from the case of the resumptive pronoun in the clause. The resumptive pronoun in (58) has accusative case, whereas the relative marker has nominative.

Kayne (1994) assumes that a relative clause is a CP complement of D. The head of the relative clause (the noun that is modified) is base generated inside the CP (or rather, IP) and moves out to spec,CP. In cases where the relative marker is a *wh*-element, Kayne assumes that *wh*+noun moves out of CP to spec,CP, and that the noun then moves to spec,*wh*.

Kayne himself notes that case is problematic for this proposal. In relative clauses with a *wh* relative marker, the head and the *wh*-element do not have to have the same case. That is strange considering two facts. First, the *wh*-noun complex starts out as one single phrase, which is case marked in the relative clause. Because it is a single phrase, one would expect it to have one case. Second, when the noun moves to spec,*wh*, it is in a spec-head relation with the *wh*. In an antisymmetric framework, one might expect spec-head agreement in such a configuration, which would copy the case of *wh* onto the noun.[61]

Another problem has not been noted by Kayne. Many languages, including Arabic, have a resumptive pronoun in the relative clause. The status of this resumptive pronoun is unclear in the analysis that Kayne adopts. There are some proposals (e.g., Pesetsky & Torrego 2001) that argue that a trace of movement can be spelled out as a resumptive pronoun, which could in principle solve this problem. However, one would still have to explain why in relative clauses this spell-out is obligatory, whereas it is impossible in many other instances of movement in Arabic. And of course, if such a solution is adopted, the question why the resumptive pronoun can have a case that is different from its antecedent becomes even more difficult to answer, because it contradicts the common assumption (cf. Chomsky 1986a) that all the elements in a chain must carry the same case.

Because Kayne's analysis faces these problems, I will adopt an alternative analysis which says that relative clauses are in specifier positions in the noun phrase, just like the other postnominal modifiers that we discussed. This analysis is supported by the fact that relative clauses show the same phenomena that the other postnominal modifiers show. For example, they show agreement: the relative marker agrees with the head noun. It agrees in number and gender, and the dual relative marker also agrees with it in case, as I remarked earlier.

In fact, the relative marker also shows agreement in definiteness. It is only present if the head noun is definite. If the head noun is indefinite, the relative clause follows it directly (but after any adjectives), without any marker at all. In this respect, too,

[61] Kayne assumes that the case assigned to *wh* is only passed on to the noun if this noun does not move further. If it does, the noun can be case-marked by another case assigner later. This, however, is hardly a satisfying analysis, given standard assumptions on case marking.

the relative clause is identical to adjectives, which also show definiteness agreement. In chapter 4 I will present a detailed analysis of adjectives, which shows that they are indeed very similar to relative clauses.

We see, then, that there is a fundamental difference between prenominal and postnominal modifiers in the Arabic noun phrase. Prenominal modifiers head a projection in which the noun is a complement of the lexical element (i.e. of the modifier in question). Postnominal modifiers on the other hand are adjoined to the noun they modify. (More precisely, to one of the functional projections in the noun phrase in which they are modifiers.)

This means that the structures with prenominal modifiers and those with postnominal modifiers are derived from different base structures, a conclusion that goes against the analysis that Shlonsky (2000) presents, which is inspired in this respect by analyses such as those of Cinque 1994, and which argues that both pre- and postnominal modifiers in the (Semitic) noun phrase are in specifier positions. The data clearly show that this is not the case in Arabic.

3.3.6 Combinations of modifiers

When we look at ordering effects of combinations of more than one modifier, we often see that there is a preferred ordering. For example, if a (prenominal) ordinal and cardinal are combined in Arabic, we see that the order ordinal-cardinal is preferred:

(59) ’awwal-u ḫams-i muḥāḍarāt-in
first.M-NOM five.M-GEN lectures-GEN
‘the first five lectures’
(Fassi Fehri 1999, p. 113-4)

Similarly, if a quantifier is combined with a demonstrative or a cardinal, the quantifier comes first:

(60) a. kull-u hāḏā -l-kalām-i
all-NOM this the-talk-GEN
‘all this talk’
b. kull-u ṯalāṯat-i riğāl-i-n
all-NOM three-GEN men-GEN-INDEF
‘every three men’
(Fassi Fehri 1999, p. 113-4)

As I explained in the previous section, noun-phrase modifiers in Arabic can also occur postnominally. If we look at the ordering properties of postnominal modifiers, we see that the modifiers have the opposite order:[62]

[62]It should be noted that the exact meaning of the postnominal constructions can differ from the prenominal constructions. For example, the example in (59) in indefinite, whereas the postnominal counterpart in (61a) is definite. The same difference exists between (60b) and (61b).

(61) a. al-muḥāḍarāt-u -l-ẖams-u -l-'ūlā
the-lectures-NOM the-five.M-NOM the-first.F-GEN
'the first five lectures'
b. al-riğāl-u -l-'arbaᶜūna kull-u-hum
the-men-NOM the-forty.NOM all-NOM-them
'all forty men'

It should be noted that certain combinations are not possible. A prenominal demonstrative cannot be combined with a prenominal numeral or adjective, for the simple reason that a prenominal demonstrative must be immediately followed by the definite article *al-*. Furthermore, prenominal adjectives usually have such specific meanings that they are often not compatible with other modifiers. Postnominal adjectives can be combined with other modifiers. As (62) shows, they normally appear first after the noun:

(62) al-kutub-u -l-faransiyyat-u -l-ẖamsat-u -l-'ūlā
the-books-NOM the-French-NOM the-five-NOM the-first
'the first five French books'

Fassi Fehri (1999) summarises the order of modifiers in the Arabic noun phrase as in (63):

(63) a. Q-Dem-Ord-Card-Adj-(Det)-N-(Gen)
b. (Det)-N-(Gen)-Adj-Card-Ord-Dem-Q-Rel

(63a) gives the order of prenominal modifiers, (63b) gives the order of postnominal modifiers. The (Det)-N-(Gen) complex is a fixed combination that cannot be split. Det and Gen are in parentheses, because they are in complementary distribution. The modifiers Q, Dem, Ord, Card and Adj can occur on both sides, appear in a fixed order, and show so-called mirror image effects. That is, the order in which they occur postnominally is the reverse of the prenominal order. Lastly, relative clauses can only appear postnominally and follow any other modifiers.

3.3.7 Combinations of adjectives

It is a well known phenomenon that when more than one adjective appears inside a noun phrase, they show specific ordering preferences, just like combinations of modifiers. Take the English example in (64):

(64) a. a beautiful gold watch
b. ??a gold beautiful watch

There are many such preferred orders. Interestingly enough, these ordering preferences are very stable cross-linguistically. A list of many of these preferences is given by Shlonsky (2000) and the references cited therein. Shlonsky also notes that the preferred order of adjectives in Hebrew and Arabic is the opposite of the order in English. The example in (65), which is the Arabic equivalent of (64), illustrates this:

(65) a. sāᶜat-un ḏahabiyyat-un ğamīlat-un
watch-NOM gold-NOM beautiful-NOM
b. ??sāᶜat-un ğamīlat-un ḏahabiyyat-un
watch-NOM beautiful-NOM gold-NOM

The ordering of combinations of adjectives is linked to the semantic categories in which adjectives can be divided. In other words, it is not so much the case that specific adjectives are ordered with respect to each other. Rather, the categories to which they belong are ordered. Laenzlinger (2000) gives the following (simplified) structure, cited from Sproat & Shih (1988, 1991), to which I have added the category *age* based on Shlonsky's list:[63]

(66) quantity > quality > size > age > form > colour > nationality

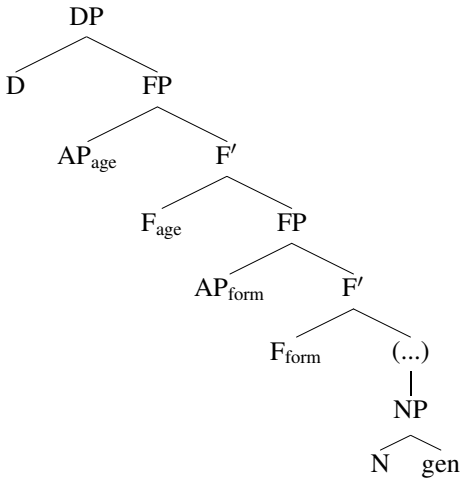
(66) means that quantity denoting adjectives in English will precede quality denoting adjectives, which in turn will precede size denoting adjectives, etc. Obviously, the order in Arabic will be the reverse.

The ordering that we see here is a hierarchical ordering. A quantity-denoting adjective is not just ordered linearly before or after a quality-denoting adjective, the two adjectives are in a hierarchical ordering. For example, in the phrases in (64) and (65), the two adjectives *beautiful* and *gold* are not on the same level, each modifying the noun *watch*. Rather, the adjective *beautiful* modifies the combination *gold watch*. In other words, a *beautiful gold watch* is not a watch that is both beautiful and made of gold, it is a gold watch that is beautiful. This hierarchical ordering of the adjectives is reflected in the analysis: the lower adjective *gold* is adjoined to the noun first, and the higher adjective *beautiful* is adjoined to the structure thus formed. In this way, the higher adjective takes scope over the lower one.

The fact that the ordering preferences are so strong cross-linguistically, even across languages that are not related, has led to the idea that the ordering relations are part of syntax. For example, Cinque (1994) and Laenzlinger (2000) argue that adjectives appear in the specifier positions of a series of designated functional heads. For each category of adjectives there is a functional head with a specifier position in which the adjective phrase can be generated. So for a phrase with an age and a form denoting adjective, we would have the structure in (67):

[63] Although they do not note it, Sproat & Shih must assume that the category "nationality" is a subcategory of some other, as yet unidentified category. It seems unlikely that UG would have accommodated a category of nationality adjectives, since nationalities are a relatively recent human invention.

(67)



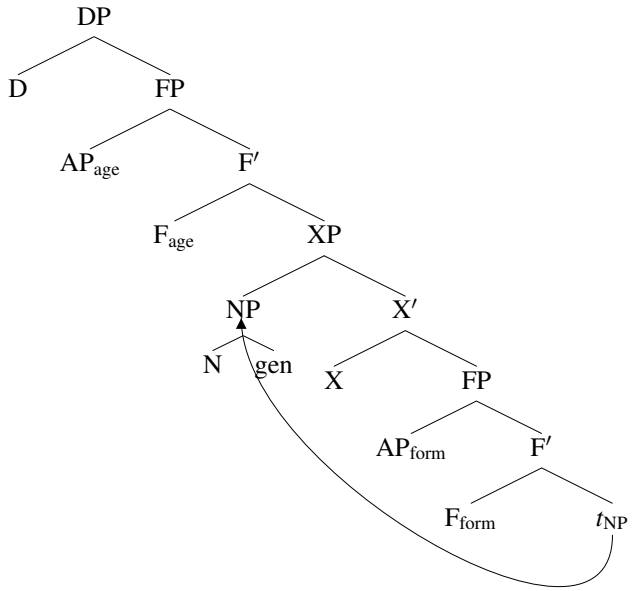
This tree contains an FP with the feature FORM, which takes a form-denoting adjective phrase as its specifier. Dominating that FP, there is another FP with an AGE feature, which takes an age-denoting adjective as its specifier. The idea is that F_{age} selects for F_{form} (and F_{form} for F_{colour} etc.), which means that the various FPs always end up in the same order, and consequently that the adjectives that they take as their specifier will do the same.

In the account of Cinque (1994) and Laenzlinger (2000), the tree structure in (67) is the basic universal structure. Both Cinque and Laenzlinger work within an antisymmetric framework and therefore assume that the tree in (67) is linearly ordered as D - AP_{age} - AP_{form} - NP. As (68) demonstrates, this is the English order:

(68) an old round hat

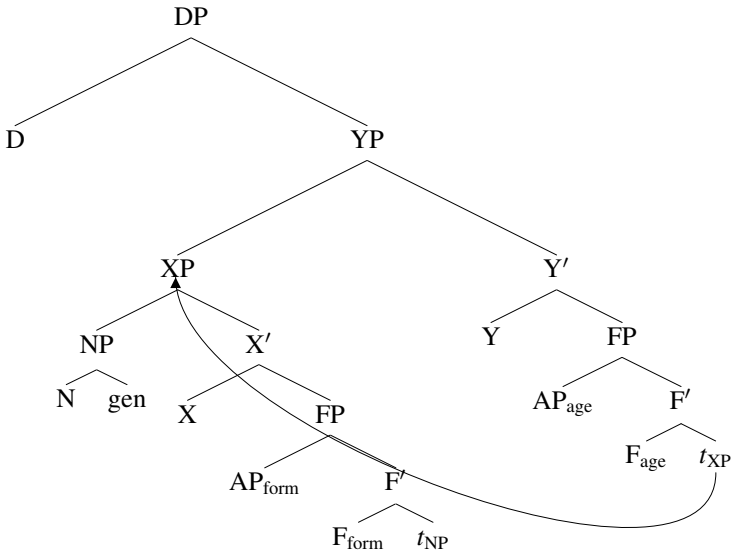
In order to account for the fact that languages such as Hebrew and Arabic have the reverse order of (68), Cinque (1994, 1996) argues that the NP can move up in the tree. First, the NP moves to the specifier position of some undesignated head X, which is located just above FP_{form}:

(69)



In the next step, this XP moves to the specifier of a functional projection above FP_{age}:

(70)



Linearising (70) according to the LCA results in the order in (71), which is the correct order for Arabic:

(71) D N Gen AP_{form} AP_{age}

The analysis outlined by Cinque is an example of the problem that I discussed in chapter 2. Cinque needs to assume the existence of two undesignated functional heads XP and YP which provide landing sites for movement.

Apart from the need to posit undesignated heads, there is also a more serious problem with the analysis. Cinque argues that the relative ordering of adjectives is the result of the hierarchical ordering of a series of functional heads. In other words, the explanation for the adjective orderings is sought in syntax. A syntactic ordering, however, puts too many restrictions on the system. It is a well known fact, for example, that adjectives *can* occur in variant orderings:

(72) a. the old round hat
b. the ROUND old hat

Sproat & Shih (1988) and Scott (1998) point out that the variant in (72b) is grammatical if the first adjective bears focus stress and if the second adjective and the noun form a discourse-relevant category. Suppose, for example, that we have a small collection of hats, two of which are old. One of these is round, the other is more oval-shaped. When in such a situation one is asked to pass "the old hat", one would have to know which of the two is meant. One could ask something like "which one?", to which an answer such as "the ROUND old hat" could be appropriate, with due stress on *round*.

Sproat & Shih's point that the second adjective and the noun must form a discourse-relevant category is very important. Above, I said that the adjectives in the 'standard' order in (72a) are hierarchically ordered. *Old* modifies not just *hat*, it modifies the complex *round hat*. What Sproat & Shih say is that in the variant ordering of (72b) *round* modifies not just *hat*, but the complex *old hat*. This means that the most straightforward way of forming (72b) is to adjoin *old* to *hat* and then adjoin *round* to the phrase *old hat*. It is unlikely that (72b) is formed on the basis of (72a), with *round* moving to a higher position.

The idea that the relative ordering of adjectives is a syntactic phenomenon faces another problem. Two adjectives in a noun phrase can be of the same category and still require an ordering. Take the example in (73) (Ø. Nilsen, p.c.):

(73) a. a visible invisible star
b. an invisible visible star

The two adjectives *visible* and *invisible* obviously belong to the same category: one is simply the negation of the other. If syntax were to provide a specifier position for each category of adjectives, the two adjectives in (73) would obviously compete for the same position. In the analysis proposed by Cinque (1994), it is not impossible for two adjectives of the same category to appear in one noun phrase, but the only way to achieve it is by conjoining the adjectives, either with *and* or asyndetically, separated by a comma. There is no room for two hierarchically ordered adjectives of the same category, because there is only one FP of that category in the projection. If two adjectives of the same category are conjoined, we see that there is no preferred order:

(74) a. a beautiful, sweet girl
b. a sweet, beautiful girl

Neither of the two orders in (74) is marked in comparison to the other. The examples in (73), however, are obviously not conjoined. Conjoining them with *and* results in inconsistency:

(75) a. #a visible and invisible star
b. #an invisible and visible star

In other words, the adjectives in the examples in (73) are hierarchically ordered. As a result, the two phrases have distinct interpretations. The (a) example of (73) denotes a star that is invisible under normal circumstances, but is now temporarily visible, say, due to some natural phenomenon. The (b) example denotes the opposite: a star that is normally visible, but is now temporarily invisible.

The examples show that the relative ordering of two adjectives is a semantic matter. Adjectives that describe a more intrinsic property are adjoined lower in the structure. In many cases, the most intrinsic of several properties of a noun is fixed, which means that out of context two adjectives will show a preferred order, as we see in (72). But in a restricted context, the situation could be different. This is the case in the dialogue sketched above: within the limited context of the dialogue, the 'oldness' property of the hats has already been established, and thus becomes fixed. The shape of the desired hat, however, is yet to be determined. As a result, *old* takes the position closest to the noun. The same thing happens in the examples of (73): the order of the two adjectives *visible* and *invisible* is determined by the semantics of the phrase.[64]

[64]In determining what it means for a property to be 'more intrinsic' than others, the distinction between individual-level and stage-level predicates is important. Individual-level properties are less intrinsic than stage-level properties. If we look at the ordering of adjective types given above, some more refinement immediately presents itself:

(i) quantity > quality > size > age > form > colour > nationality

Closest to the noun are absolute adjectives such as those of form, colour and nationality. Preceding those are relative adjectives, such as those of quality and size. Relative adjectives are less intrinsic than absolute ones because they depend on a comparison to other elements, not just on the referent itself. Furthermore, if we look at the relative adjectives, we see that adjectives of quality, which often carry a large degree of subjectivity, precede adjectives of size, which express a more objective measure.

As for the relative ordering of the absolute adjectives, it seems likely that the intrinsicness of nationality has to do with the fact that it is a human invention: it does not depend on the physical form of the world, but on our organisation of it. We could argue that human-imposed properties are more intrinsic than 'natural' properties, because we are in a sense more 'sure' of them, exactly because we invented them.

Quantity adjectives precede all other adjectives. This is not surprising, since quantity, or cardinality, is a highly variable property. In fact, quantity is not even a property of the noun itself, but only of the referent of the noun phrase as a whole. That is, in *two cars*, the cardinality is not a property of the nominal predicate *car*, but of the actual referent in the universe of discourse, which happens to be formed of two cars.

There is of course a lot more that can be said about the semantic categories of adjectives. These few remarks are not intended as an account of the hierarchical ordering of the various categories, but they do show that such an account is feasible.

3.4 Linearisation

In the previous section, I have argued that the Arabic noun phrase can contain two types of modifiers: heads that take the noun as complement, and specifiers that are predicated of the noun. What we see is that head-modifiers appear before the modified noun, and that specifier-modifiers appear after the modified noun. In this section, I discuss the linearisation of the noun phrase. First, I discuss the existing antisymmetric approaches, and show that they suffer from the problems discussed in chapter 2. After that, I show how RLin can account for the ordering facts.

3.4.1 Antisymmetric approaches

The fact that head-modifiers precede the noun is not surprising under the assumption that the universal (underlying) order of phrase structure is spec-head-comp, as Kayne (1994) does. In that order, heads precede their complements. However, given such an order, we would not expect that the specifier-modifiers *follow* the noun. They would be expected to precede it.

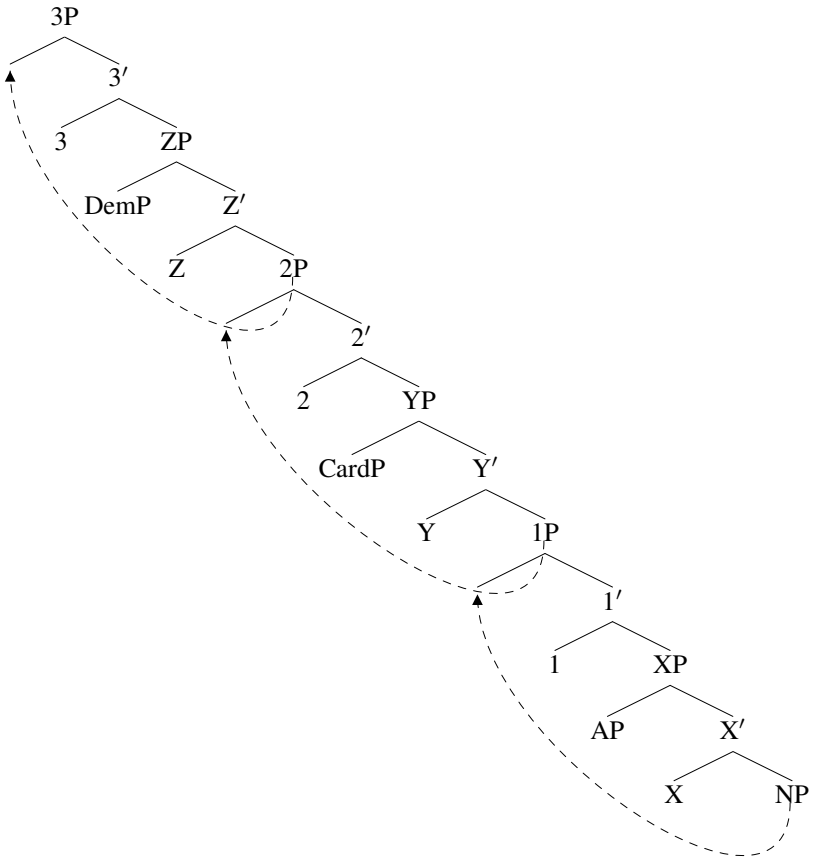
Intuitively, the fact that the specifier-modifiers in Arabic follow the noun is related to the fact that they appear in mirror-image orders, as we saw in the previous section. Shlonsky (2000) tries to account for this intuition with an antisymmetric approach. Although he mainly discusses Hebrew, he claims that the analysis works for Arabic as well. He argues that there is no head movement in the Semitic noun phrase,[65] and that instead, all word order variation is obtained through the interplay of two operations: pied-piping and remnant movement.

Pied-piping is what we have seen in (69) and (70); it derives the postnominal positionings. Remnant movement is movement of a constituent that contains a trace.[66] Shlonsky starts out with the tree in (76). I have indicated the movements that will take place:

[65] Although at some point, he does propose head movement of Dem to some higher projection, leaving it unclear what this projection is, and why the movement should take place.

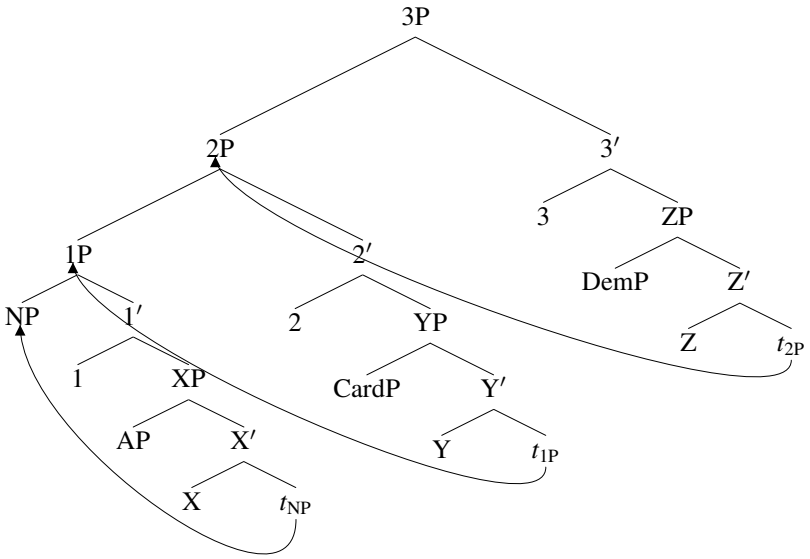
[66] Note that this has the effect of moving a trace out of the c-command domain of its antecedent, yielding a configuration in which the trace is no longer c-commanded by its antecedent. This may seem odd, but there are clear examples of such configurations, e.g. in structures in Dutch or German in which the VP has been topicalised without the object: the object first moves out of the VP through scrambling, and the VP, with the trace, then moves to spec,CP.

(76)



The noun phrase contains a number of undesignated functional heads (indicated with XP, YP and ZP) that house the modifier projections in their specifiers, and there are a number of undesignated functional heads (indicated with 1P, 2P and 3P) which provide landing sites for movement. In the pied-piping derivation, NP now moves to spec,1P. In the following step, XP is supposed to move, but XP pied-pipes the phrase it is contained in, 1P. As a result, 1P ends up in spec,2P. The process repeats itself, by moving 2P to spec,3P. The resulting tree is the following:

(77)

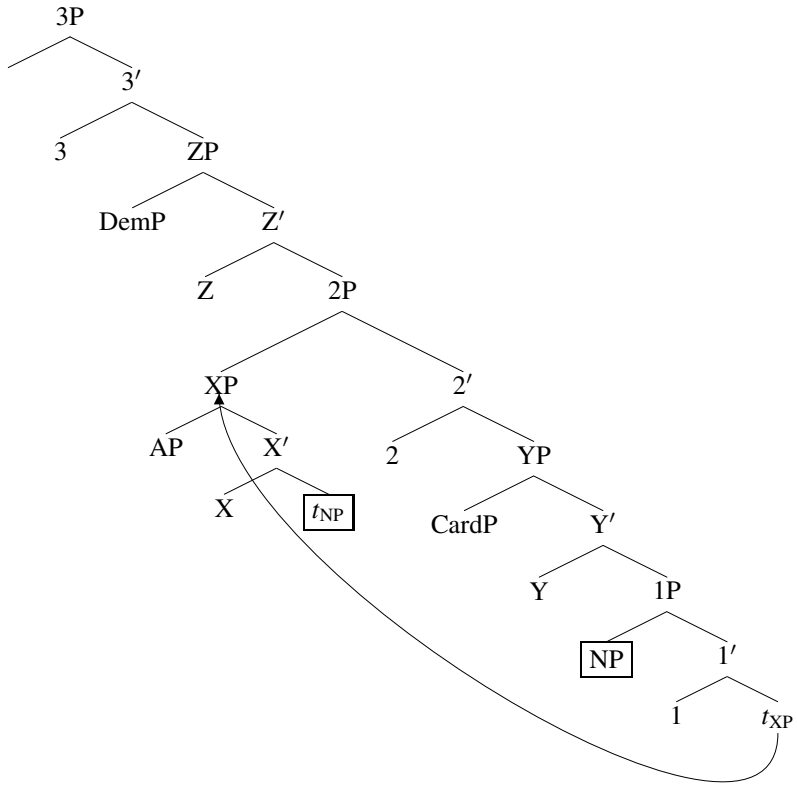


When linearised, this tree produces the following order, which has the modifiers in postnominal position:

(78) NP AP CardP DemP

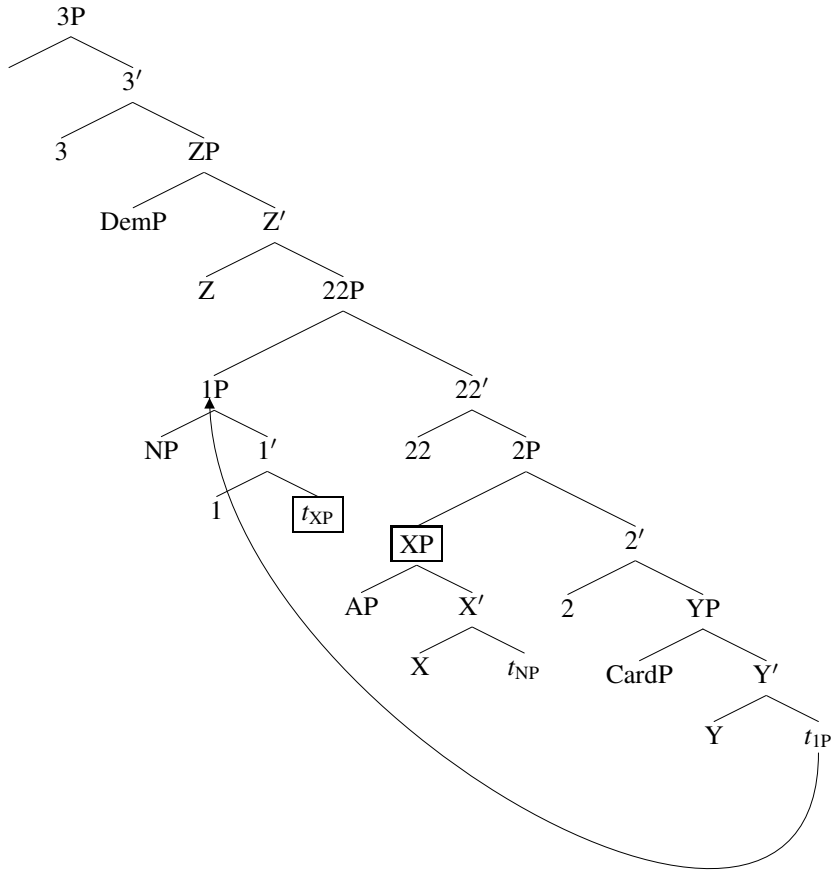
In the remnant-movement structure, the NP also moves to spec,1P, but when XP moves to spec,2P, it moves without pied-piping 1P. As a result, the trace of NP, which is inside XP, moves to a position above NP. This creates the following tree:

(79)



After this movement, Shlonsky must assume that 1P moves (again positioning a trace above its antecedent, in this case the trace of XP) to a position below ZP, which he calls position 22P:

(80)



At this point, we have an order in which the demonstrative is prenominal:

(81) Dem NP AP CardP

Shlonsky basically develops an analysis in which there are two ways to derive a certain order: both the pied-piping approach and the remnant-movement approach yield the same orders.

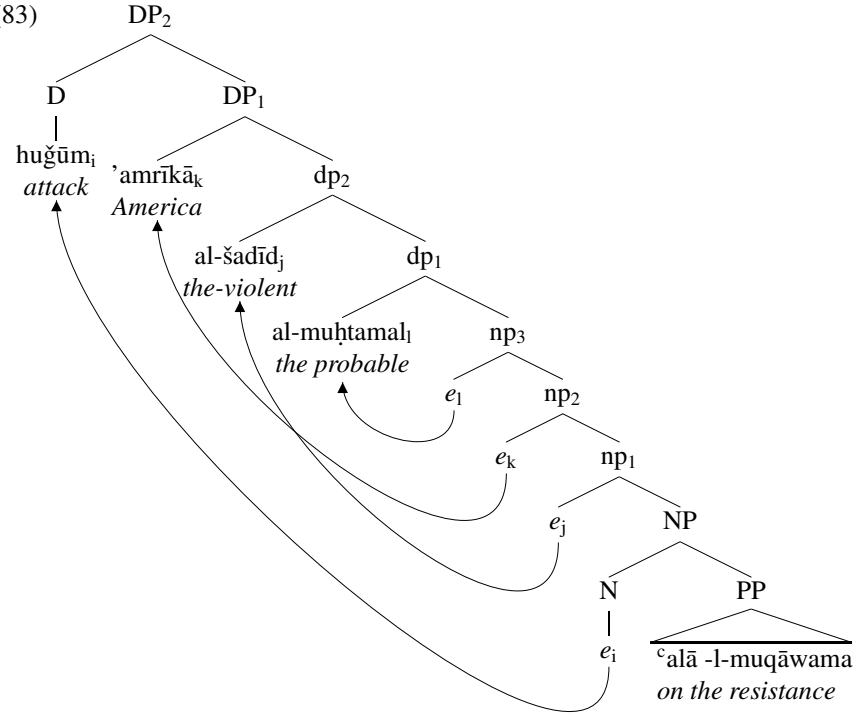
However, a more important problem is that Shlonsky cannot account for the fact that prenominal modifiers are of a different nature than postnominal ones. In Shlonsky's analysis, prenominal modifiers are simply those modifiers that the noun does not pass while moving up in the tree. This would imply that prenominal and postnominal modifiers are exactly the same, which is not the case, as I have shown in section 3.3. Fassi Fehri (1999), who discusses Standard Arabic, argues for a very different approach. In his analysis, APs can move independently. Take the following example:

(82) huğūm-u ’amrīkā -l-šadīd-u -l-muḥtamal-u ᶜalā
attack-NOM America.GEN the-violent-NOM the-probable-NOM on
-l-muqāwamat-i
the-resistance-GEN
‘the probable violent attack on the resistance by the US’

(82) contains an example of a deverbal noun, *huğūm* ‘attack’ with *’amrīkā* ‘America’ as its subject and the PP *ᶜalā -l-muqāwama* as the object. The noun is also modified by two adjectives, which appear in mirror image order.

The tree structure that Fassi Fehri proposes is the following:

(83)



What happens here is that the adjectives and the genitive are generated as specifiers of a series of *np*’s. The adjectives then move to the specifier positions of a series of *dp*’s to be licensed. These movements take place from top to bottom, that is, the highest adjective moves first, which creates the mirror-image effect. Then the genitive moves to the specifier position of D_1.[67] Lastly, N moves to the highest D_2.

Fassi Fehri’s analysis is problematic for several reasons. First, although the analysis is claimed to be an antisymmetric one, the trees that Fassi Fehri proposes do not comply with the LCA. It follows from the LCA that each head must have a complement and a specifier, and that each projection must have a unique head. However, in

[67] Fassi Fehri explicitly states that the genitive complement moves to the specifier of DP_1. The head D_1 is not indicated, and I am not sure where to locate it.

(83), the small *dp*'s and *np*'s do not have heads, and as already remarked, DP_1 does not seem to have a head, either.

A second problem with Fassi Fehri's approach is that the movements of the adjectives and genitive modifiers violate minimality. The adjective that is highest in the tree moves first. When the second adjective moves, under any interpretation of minimality, the first adjective should intervene. Fassi Fehri explains this by saying that the small *np*'s somehow form a minimal domain, because they are further projections of NP. But if this is the case, there does not seem to be any way to force the highest adjective to move first.

Fassi Fehri's tree is also rather inconsistent. The *np*'s are located above NP, (which seems to indicate that they are a way to formalise multiple specifiers in an antisymmetric framework), but the *dp*'s are located *below* DP. Locating them above DP would obviously result in the wrong word order, but it is not clear why the various projections are located where they are.

More seriously, however, is the fact that Fassi Fehri's analysis abandons the intuition that the mirror-image order of the adjectives is linked to their postnominal position. In (83), the adjectives move independently of the noun, and there is no real reason to assume that both the noun movement and the adjective movements *must* take place. It seems perfectly possible for the adjectives to move without the noun moving, which would yield an order with prenominal adjectives in mirror-image order. This, however, is not a configuration that is known to exist in any language.

3.4.2 The recursive approach

In section 3.3.6 I have given the following schematic orderings for Arabic noun phrases:

(84) a. Q-Dem-Ord-Card-Adj-(Det)-N-(Gen)
b. (Det)-N-(Gen)-Adj-Card-Ord-Dem-Q-Rel

As I have argued, the prenominal modifiers in (84a) are all heads taking the next modifier or the noun as a complement. The postnominal modifiers in (84b), on the other hand, are specifiers, except for Gen, which is a complement of N. We see, then, that the order of the Arabic noun phrase is best characterised as follows:

(85) head comp spec

That is, a head always precedes its complement, and a specifier always follows its head (and consequently the complement of that head).

Because heads precede their complements, we know that the ordering of the principles S and H must be $H > S$. The head principle is stronger than the selection principle, which will give the head precedence in the linearisation, resulting in the head-complement order.

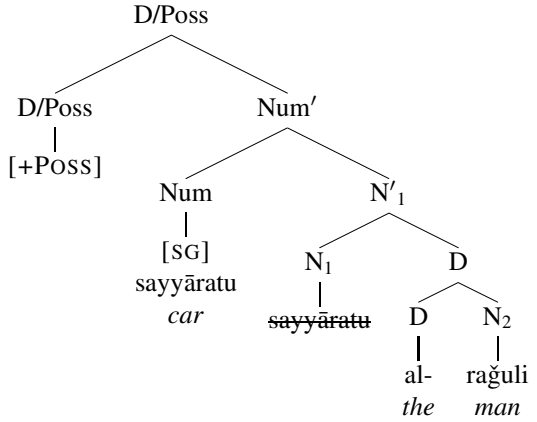
The specifiers relevant to the linearisation are the specifiers in (84b): the postnominal adjective, numerals, demonstratives, quantifiers and relative clauses. As we have seen in section 3.3, all of these elements are non-selected specifiers, or adjuncts. This means that the adjunct parameter must have the setting adjunct second. In the previ-

ous section, we saw that the setting adjunct second yields mirror-image orders, which is correct for Arabic. As the ordering in (84) shows, the postnominal modifiers show mirror-image orders.

To demonstrate, let us consider some examples. It should be noted that I will always draw trees with specifiers preceding their heads, and with heads preceding their complements. The reason for this is twofold. First, on the practical side, trees of this type take up less room on the page, and they are easier to read.[68] From a more theoretical point of view, the trees that I work with are not linearly ordered, they only represent hierarchical structures. For that reason, it does not matter how we draw a tree. I will emphasise this by explicitly drawing all trees in the same manner and by giving the intended linearisation when required.

Let us first see how the procedure works for a basic genitive construction in Arabic, as we have analysed it above: a combined D/Poss head taking Num as its complement, and N the complement of Num. The possessor of the noun is the complement of N. Furthermore, N moves to Num:

(86)



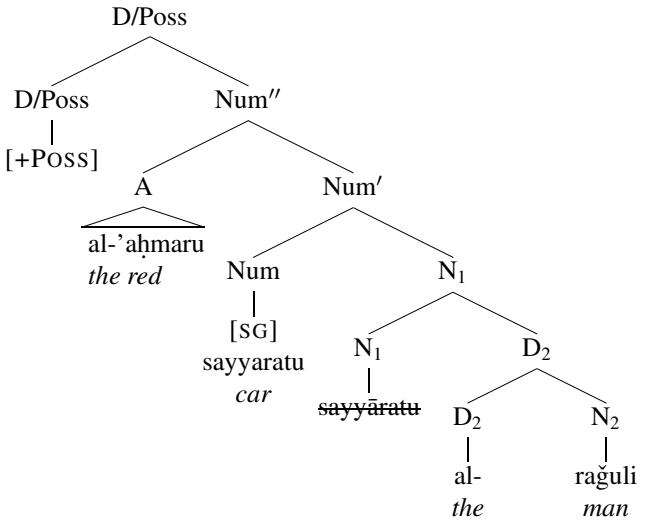
If we linearise (86) with $H > S$, D/Poss is spelled out first, because it is a head, which has precedence over its complement. After that, the sister node Num′ is linearised. In Num′, [SG] is spelled out first, together with the noun *sayyāra*, which has moved to it. Next, N′$_1$ will be linearised. In N′$_1$, the trace of the noun *sayyāra*, being the head, is spelled out first, followed by the D *al-rağuli*. The resulting order is the following:

(87) +POSS sayyārat-u.SG ~~sayyārat-a~~ al-rağul-i
car-NOM the-man-GEN
'the man's car'

As we know from section 3.1, this is the correct word order. Now let us see what happens if we modify the head noun with an adjective. If we assume that the adjective is a specifier of Num, we get the following:

[68] And, for that matter, easier to construct in LaTeX.

(88)



If we linearise this with adjunct-second the adjective phrase will be spelled out last. The reason for this is obvious: because of the parameter setting adjunct-second, RLin will first linearise the projecting node. So when RLin reaches Num″, the projecting node Num′ will be linearised *before* the non-projecting node, that is, the adjective phrase. The resulting linear structure is the following:

(89) +POSS sayyārat-u.SG ~~sayyārat-a~~ al-rağul-i -l-’aḥmar-u
car-NOM the-man-GEN the-red-NOM
‘the man’s red car’

As we see, RLin linearises the adjective phrase that modifies the head noun of a genitive construction *after* the genitive noun. We know from section 3.1 that this is the correct order.

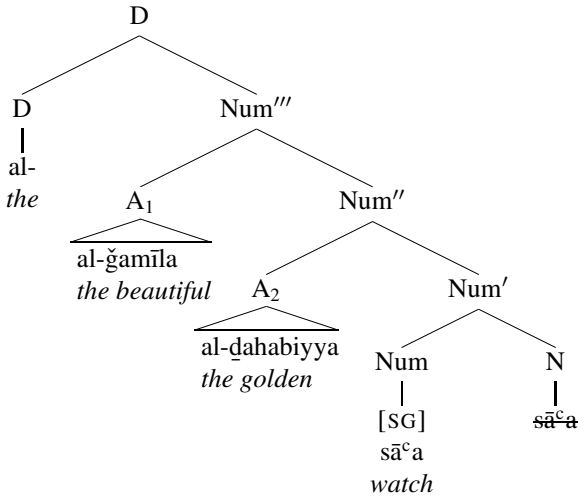
Now let us see how the linearisation of a noun phrase with two adjectives works. Take the following example:[69]

(90) al-sāᶜa -l-ḏahabiyya -l-ğamīla
the-watch the-gold the-beautiful
‘the beautiful gold watch’

[69] I have omitted the case endings here.

The tree structure of this noun phrase is the following:

(91)



The noun phrase is linearised with $H > S$ and with adjunct-second. That means that at the highest level, the node D, which is a head, is spelled out first. In the sister node of D, Num‴, we see that the subnode A_1 is not selected, which means that RLin will linearise Num″ first. In Num″, the same thing happens: the adjective A_2 is not selected, which means that Num′ will be linearised first. In Num′, the ordering $H > S$ forces the linearisation of the head before the complement, which means that Num, with the noun *sāᶜa* in it, is spelled out first, followed by N. Up to this point, we have spelled out the head D and the noun:

(92) al- sāᶜa.SG ~~sāᶜa~~
the watch

Num′ is one of the subnodes of Num″, which means that we have now completed part of the linearisation of Num″. We must now linearise the other subnode, A_2. As a result, the adjective *ḏahabiyya* is spelled out. This concludes the linearisation of Num″. Since Num″ is a subnode of Num‴, we have also concluded the linearisation of the first subnode of Num‴. So we must now linearise its second subnode, A_1. Omitting the trace of the noun, the order that results has the adjectives in a mirror image order:

(93) al- sāᶜa.SG al-ḏahabiyya al-ǧamīla
the watch the-gold the-beautiful

It is possible in a genitive construction that both the head noun and the genitive noun are accompanied by an adjective. In such instances, the adjective that modifies the genitive noun will directly follow it, and the adjective that modifies the head noun follows after that adjective. In other words, we have the order N-Gen-A_{Gen}-A_N. This construction is only used when there is little danger of confusion, e.g. when the nouns

have a different gender or number. Two examples of this construction are given in (94):

(94) a. mustawā ṭullāb-i -l-ṯānawiyy-at-i -l-ᶜāmm-at-i
level.M-NOM students-GEN the-secondary-F-GEN the-general-F-GEN
-l-hazīl-u ǧiddan
the-skinny.M-NOM very
'the very poor results of the students of general secondary school'

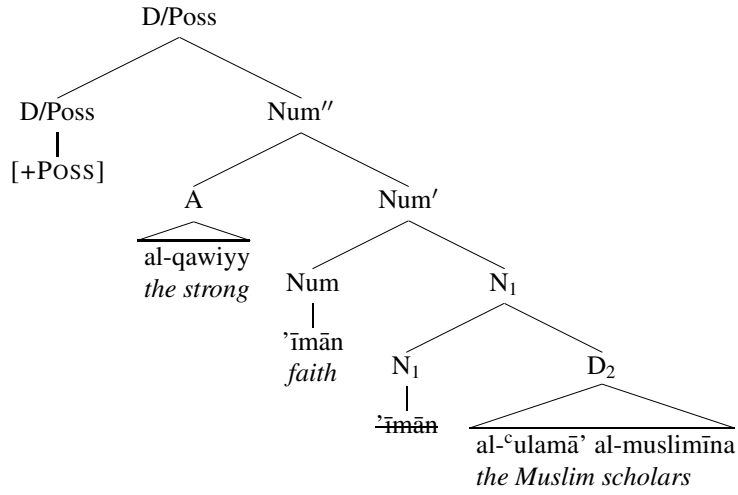
b. 'īmān-u -l-ᶜulamā'-i -l-muslim-īna -l-qawiyy-u
faith-NOM the-scholars-GEN the-muslim-PL.GEN the-strong.SG-NOM
'the strong faith of the Muslim scholars'

In (94a), the adjective *ᶜāmma* 'general' modifies the noun *ṯānawiyya* 'secondary school', whereas the adjective *hazīl* 'skinny, meagre' modifies the head noun *mustawā* 'level'.[70] This is confirmed by the form of the adjectives: *ᶜāmma* 'general' is feminine, like the noun *ṯānawiyya* 'secondary school', whereas *hazīl* 'skinny' is masculine, like *mustawā* 'level'.

The same thing happens in (94b), where the singular head noun *'īmān* 'faith' is modified by the second adjective *qawiyy* 'strong', which is singular as well, whereas the genitive noun *ᶜulamā'* 'scholars', which is plural, is modified by the first adjective *muslimīna* 'Muslim, Islamic', also a plural form.

RLin actually derives this order without any problems:

(95)



[70]The genitive construction is in fact a double-layer construction: the genitive modifier of *mustawā* 'level' is *ṭullāb-u -l-ṯānawiyya* 'students of secondary school', which is itself a genitive construction with the head noun *ṭullāb* 'students' and the genitive noun *ṯānawiyya* 'secondary school'.

The complement of the noun *ʾīmān* 'faith' is linearised as a whole, directly after the heads are spelled out. Because the adjective *muslimīna* 'Muslim' is part of this complement, it will appear directly after *ʿulamāʾ* 'scholars'. The adjective *qawiyy* 'strong' which modifies *ʾīmān* 'faith' is linearised after the noun and its complement, just like in the case of (88). As a result, the ordering indicated in (94b) ensues.

3.4.3 The English noun phrase

In order to show how the model proposed here can derive cross-linguistic variation in word order, I will now take a look at English. As I will demonstrate, the English noun phrase resembles the Arabic noun phrase to a great extent. We can describe the differences that we observe with the parameters that I have introduced.

Let us first look at the Saxon genitive. Like the Arabic construct state, only the possessor noun in a Saxon genitive construction has a determiner. The possessed noun does not:

(96) the man's (*the) car

Furthermore, it has been noted (e.g. Grimshaw 1990) that the definiteness of the noun phrase depends on the definiteness of the possessor noun:

(97) a. *there is the man's shirt on the chair
b. there is a man's shirt on the chair

In (97), a Saxon genitive appears in an existential *there*-construction, which requires an indefinite. Only (97b), the example in which the possessor noun is indefinite, is grammatical, which means that *the man's shirt* must be definite, and that *a man's shirt* must be indefinite. So we conclude that the Saxon genitive has definiteness inheritance, just like the Arabic construct state.

This means that the Saxon genitive, like the construct state, has a syncretic D/Poss head. Unlike Arabic, this D/Poss head has an overt reflex in the clitic element *-'s*. As has often been pointed out, the *-'s* in the Saxon genitive is a clitic element, not a suffix. Two relevant examples are in (98):

(98) a. the Queen of England's horses
b. (i) John's car
(ii) John and Mary's car

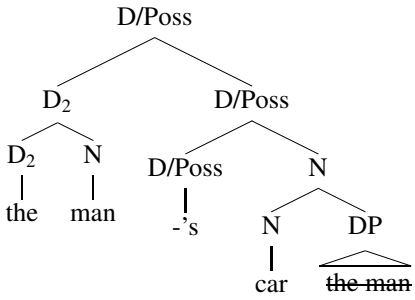
If *-'s* were a suffix, we would expect it to appear on the head noun of the possessor DP. In (98a), however, the *-'s* attaches to *England*, which is not the head noun of the possessor DP. (98b) has a similar logic: as we see in (98b-i), *John* can take the *-'s*. Therefore, if *-'s* were a suffix, we would expect it to appear on *John* in (98b-ii) as well.

The conclusion is that *-'s* is a clitic element that appears in a head. Since it appears only in the Saxon genitive, i.e. only in structures with a [+POSS] feature, we can say

that it is a reflex of the D/Poss head.[71] With *-'s* in D/Poss, we obtain the structure in (99):

(99) a. the man's car

b.



This structure is very much like the structure I developed for Arabic: the possessor noun is generated inside the lexical projection, and D/Poss needs to value its set of φ-features. It does so by establishing an agree relation with the DP complement of N. In this process, the possessor DP *the man* is assigned genitive case, which is null in English, just like nominative and accusative.

One difference with Arabic is that the possessor DP moves and merges with D/Poss. We must assume that D/Poss in English has an EPP feature which it does not have in Arabic. In this respect, D/Poss is similar to T in the clause, which can also have an EPP feature that forces the movement of the subject to spec,TP.

Looking at the tree in (99), we see that the ordering of S and H in the English noun phrase must be $H > S$, because the complement of D/Poss is linearised after *-'s*. This means that the English noun phrase has the same ordering of the two principles as the Arabic noun phrase. The difference in position of the possessor noun follows from the EPP feature, which the Arabic D/Poss lacks.

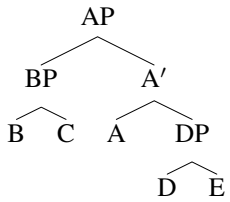
This point raises an important question. The proposal of RLin is based on the idea that word order variation should not be explained with movement. For example, the variation between VO and OV structures is the result of a different ordering of H and S, not of movement of the object over the verb. For this reason, we cannot argue that *the man* has moved solely on the basis of the fact that it appears first in the DP, especially considering the fact that we cannot give any real motivation for this movement: we say it takes place because Poss has an EPP feature, but this is of course a stipulation.[72]

Yet, there is good reason to say that *the man* has moved. As discussed in chapter 1, RLin searches a tree branch by branch. That is, if we have a tree such as in (100), the branch A′ will be searched in its entirety either before or after the branch BP:

[71] This analysis is very similar to one of two possible analyses for *-'s* that Abney (1987) discusses. Abney does not reach a definitive conclusion on the question which of the analyses is the correct one. He argues that both are in fact able to account for the facts and then says that he prefers the alternative analysis, in which *-'s* is a case marker. For the reasons discussed above, I do not follow him in this.

[72] Ideally, the EPP can be derived from more fundamental principles. See for example Haeberli (2000).

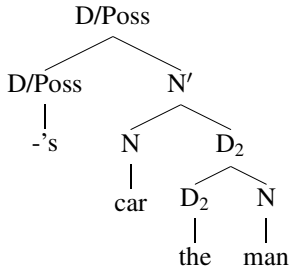
(100)



Suppose that BP is selected. Then (100) will yield as surface order either [BP A DP] or [BP DP A], depending on the relative ordering of H and S. Note that in both orders A and its complement DP are adjacent. This follows from the fact that RLin searches a tree branch by branch: RLin will always linearise A and DP one after the other, either in the order A-DP or in the order DP-A. It is not possible for RLin to linearise A and DP with other material, e.g. BP, intervening.

However, when we examine the linear string in (99a), we see that the noun *car* and the possessor DP *the man* are *not* adjacent: the head *-'s* intervenes. Suppose now that the tree structure were (101):

(101)



In order for RLin to derive the correct surface order from this tree, it would have to go down the branch N′, skip N and go directly to D_2, spell out everything there, then go back up, leaving the branch N′ *unfinished* (since N has not been spelled out yet), spell out *-'s*, and then go back into N′ again, in order to spell out *car*. This, however, is not possible, as we have just seen.

For this reason, we must conclude that *the man* has moved. Given that *the man* is in an agree relation with Poss, and that Agree can be followed by Merge, it follows that *the man* has moved to spec,PossP.

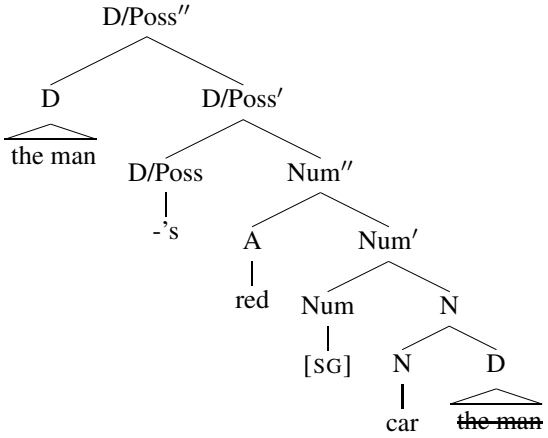
Let us see what happens when there is an adjective in the phrase:

(102) the man's red car

Assuming that the adjective is adjoined to Num, just like Arabic adjectives,[73] we get the following structure:

[73] Abney (1987) argues that in English, prenominal adjectives are actually heads that take the noun as a complement. Since the present discussion aims to present RLin rather than develop an account of the English noun phrase, I do not go into this matter here.

(103) a.

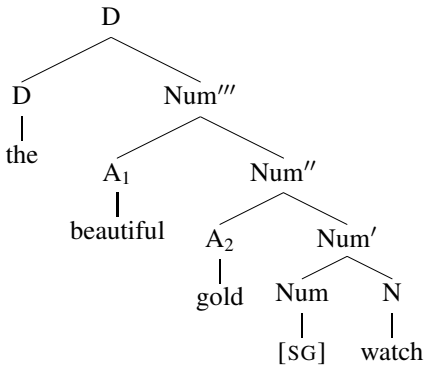


b. the man -'s red SG car ~~the man~~

It is easy to see that the adjunct parameter must be set to adjunct-first if we wish to derive the correct order for English. With this setting, the adjective will be linearised first in the node Num″, resulting in the order given in (103b).

When there are two adjectives in the English noun phrase we get a structure such as in (104):

(104)



Note that the tree structure for this phrase is identical to the structure of its Arabic equivalent in (91). The difference between the two languages lies in the settings of the adjunct parameter. As we just concluded, the English noun phrase is adjunct-first. At the highest level, that of D, there is no difference with Arabic: D, being the projecting node and a head, is linearised first. Then, Num‴ is linearised. Because English is adjunct-first, the adjective will be linearised first, before the rest of the tree in Num″ is linearised. At the level of Num″, the same thing happens: A_2 is linearised before Num′ and everything contained in it, giving the order in (105):

(105) the beautiful gold watch

These examples clearly show the advantage of the system of recursive linearisation. The tree structures of the Arabic and English noun phrases are very similar. The only difference is the movement of the complement of N to spec,Poss, which takes place in English but not in Arabic, and the setting of the adjunct parameter. Ideally, noun-phrase structure will be very similar cross-linguistically, with all variation captured by a few parameters. The present discussion shows that we can indeed do this with RLin when we compare English and Arabic.

The important thing is that we do not need a movement account for the variation in adjective placement. This is an attractive result, because we can now distinguish between the effects of N-movement and the effects of linearisation on the ordering of adjectives. Cinque (1996) notes there are three adjective orderings that are common across languages:

(106) a. $A_1\ A_2\ N$
b. $N\ A_1\ A_2$
c. $N\ A_2\ A_1$

We find (106a) in English. (106b), postnominal adjectives with an English order, is a pattern known from Irish. (106c), postnominal adjectives in mirror-image order is what we see in Arabic.

I have shown that the order in (106c) can be derived without movement of the noun or the NP. The English order can also be derived from the same tree, without any movement.[74] The difference with respect to Arabic is that the linearisation parameters have different values.

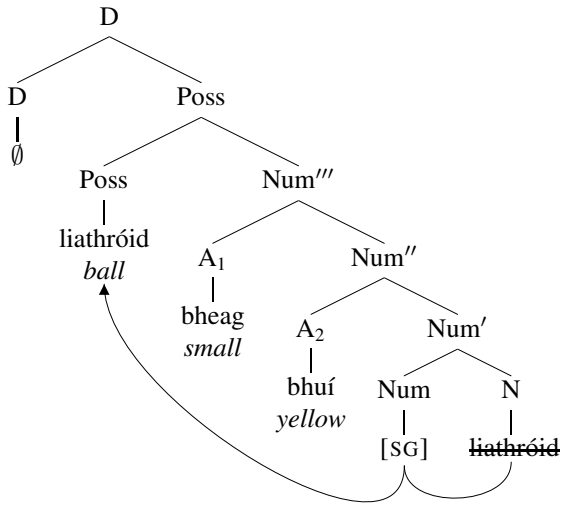
The Irish order in (106b) can be derived through N-movement. Take the following examples (from Sproat & Shih 1991, p. 587):

(107) a. liathróid bheag bhuí
ball small yellow
'a small yellow ball'
b. cupán mór Sasanach
cup big English
'a big English cup'
c. pláta cruinn dearg
plate round red
'a round red plate'

The assumption that Sproat & Shih make for these structures is that N moves up, although they do not specify to which head. Without wanting to make any specific claims on this, let us say for the sake of exposition that the Irish noun moves to Poss. The resulting tree structure is (108):

[74] Of course, there is movement from comp,NP to spec,PossP in English, but this does not affect the adjective ordering.

(108)

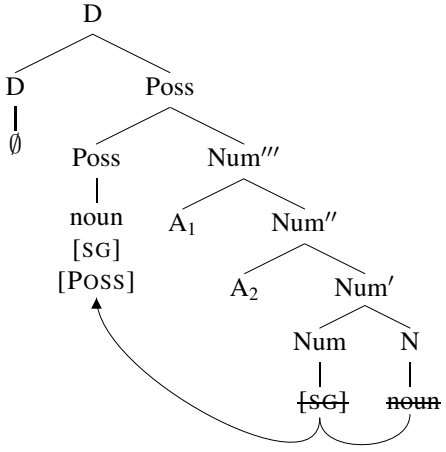


If we linearise this tree with the same settings as in English, i.e. $H > S$ and adjunct-first, we derive the variant ordering. Because N has raised to a head above the adjectives, it is spelled out before them. And because the adjunct parameter has the value adjunct-first, the adjectives are spelled out before their sister nodes, resulting in the English adjective order.

So we see that head movement can create a variant ordering in the case of adjunct-first: English and Irish only differ from each other in this respect. One could therefore ask whether head movement can create a similar variant ordering in the case of Arabic, which has the parameter adjunct-second.

Interestingly, this is not the case. Suppose we have a language X that has the same settings as in Arabic, but it moves the head to a position above Num:

(109)



Because the adjunct parameter is set to adjunct-second, the adjectives are not spelled out first in their nodes. Rather, RLin first spells out the sister nodes of the adjectives, Num″ and Num′. The resulting order is the following:

(110) D noun-SG-POSS ~~noun~~ A_2 A_1

This ordering is indistinguishable from the Arabic ordering, which also has all the heads preceding the adjectives. We see that head movement only creates a variant ordering in the case of adjunct-first.

3.5 Summary

In this chapter, I have presented an analysis of the noun phrase in Arabic. I have argued for the presence of a number of functional projections: D, Poss and Num. Each of these heads projects a feature: DEF, POSS and NUM, respectively. There are two other syntactic features in the Arabic noun phrase, CASE and GEN, but they do not project independent heads. Instead, CASE is present on the noun and is inherited by the other heads, which means that it can be spelled out on any head, and GEN projects a hybrid head with NUM.

In a genitive construction, Poss has the value [+POSS] and additionally a set of unvalued φ-features. In this case, Poss projects a hybrid head with D, resulting in a D/Poss head that has an unvalued DEF feature. This DEF feature is valued in the agreement process that takes place to value the φ-features on Poss. The result of this is that the head noun of a genitive construction "inherits" the definiteness of its complement.

The agreement process between Poss and the genitive complement in the noun phrase also results in genitive case assignment to the complement noun. This means that genitive case is assigned in a mechanism that is identical to the mechanism that assigns nominative and accusative in the clause. In other words, genitive case is a structural case, on a par with nominative and accusative, and not an inherent case.

If the heads Poss and Num contain overt material, this material is always affixal. As a result of this, N moves to these heads in order to pick up these affixes. The same happens when D projects [–DEF], which manifests itself as a suffix *-n* or, if CASE is expressed on D, in variant case endings (the so-called diptotic nouns.) When D projects the feature [+DEF], D does not contain an affix but the element *al-*, which is morphologically independent, although it cliticises onto the noun. That *al-* is an independent element is supported by the fact that it does not show so-called "morphological interference" with other morphological elements, and by the fact that N-to-D movement does not take place in generic usage.

Most modifiers in the Arabic noun phrase can appear both before and after the noun. Closer examination shows that prenominal modifiers are very different from postnominal modifiers, however. They do not modify the noun phrase as an adjunct, rather they head the noun phrase and take the noun that is semantically the nucleus of the phrase as a complement. Postnominal modifiers on the other hand are adjuncts inside the noun phrase, and are adjoined to one of the functional projections, which I

have assumed to be Num. The fact that pre- and postnominal modifiers are so different means that we cannot account for them with a standard antisymmetric approach, which argues that all modifiers are in specifier positions in the noun phrase, and that ordering variation is a result of movement or the lack thereof.

When different modifiers are combined, the postnominal modifiers show a reverse order in comparison to the prenominal modifiers. Similarly, when two adjectives occur in one noun phrase, they have a preferred order, and they also show mirror-image effects when compared to combinations of adjectives in English. The standard assumption in antisymmetric approaches is that these preferred orders are caused by the syntactic ordering of the heads that provide specifier positions for these modifiers. However, such a syntactic account cannot explain the flexibility that adjective orderings allow: adjectives can appear in non-preferred orders if the semantics of the phrase requires this. This shows that the ordering of adjectives (and presumably other modifiers as well) is not a syntactic phenomenon, but a semantic one.

In an antisymmetric approach, the mirror images are accounted for by positing a number of unmotivated movements to specifier positions of a series of undesignated heads. One is forced to assume this if one adopts Kayne's (1994) proposal that UG specifies a fixed word order. I do not accept this idea, and instead adopt the proposal that linear order is derived at PF. Deriving linear order is done by a procedure which I call RLin, which searches the tree for terminal elements to spell out. In searching the tree, RLin will have to decide at each compound node which of the two subnodes to search first. Two parameters are relevant for this decision: the adjunct parameter, which determines whether adjuncts (non-selected specifiers) are to be linearised first or second, and a parameter which orders two principles: principle S states that a selected element is to be linearised first, and principle H states that a head is to be linearised first. For selected specifiers, only S is relevant, which means that selected specifiers are always linearised first. For nodes composed of a head and its complement, both principles are relevant, which means that the relative ordering of S and H determines whether it is the head or the complement that is linearised first.

With these two parameters, we can describe the linear ordering of the Arabic noun phrase. In the Arabic noun phrase, H is ordered before S, and the adjunct parameter has the value adjunct-second. With these settings, we derive the postnominal ordering of adjectives, and we also derive the mirror-image effects. Furthermore, we can describe the order of the English noun phrase with the same variables. The English noun phrase also has $H > S$, but unlike Arabic, it is adjunct-first. One other difference between English and Arabic is that Poss in English has an EPP feature: the complement noun that is assigned genitive moves to spec,PossP, which means that it appears phrase-initially. This movement is familiar to us because it is very similar to the movement of the subject of a clause to spec,TP. Although it is not clear what the rationale for this movement is, we must assume that it takes place, because if it did not, a tree structure would result that violates the common assumption that two elements that are in the same branch of the tree cannot be separated in the linear structure by an element that is in a different branch in the tree.

Order variation can also come about by head movement, as shown by Irish, which has postnominal adjectives in the English order. RLin thus allows us to describe order

variation across languages with two parameters combined with a small number of already familiar movement operations. This is an important advantage over approaches that must resort to ill-defined heads and movement operations.

4

Adjectives

4.1 Introduction

In the discussion of adjectives, the main question is usually where adjectives are located in the noun-phrase structure. I have taken a position on this in the previous chapter, but I believe there are other, more important questions to be answered. Probably the most important one is what the internal structure of the adjective phrase is like. The initial assumption, in light of Abney's (1987) analysis, is of course that the adjective phrase has a clause-like structure, with equivalents for C and T, just as the noun phrase has.

Data in Arabic shows that this is indeed the case. I will start the discussion by looking at agreement between an adjective and the noun it modifies, sometimes called 'concord'. There are some peculiar agreement facts in Arabic that suggest that the structure of the adjective phrase is more complex than sometimes assumed. This discussion will give us an initial framework for the analysis with which we can then look at other aspects. These include genitival complements, the Deg head, and the D head that shows up on adjectives in Arabic.

Adjectival agreement in Arabic shows the pattern familiar from Romance languages: there is agreement in gender (1a,b) and number (1c,d):[1]

(1) a. rağul-un ṭawīl-un
man-NOM tall.M-NOM
'a tall man'

b. imra'at-un ṭawīl-at-un
woman-NOM tall-F-NOM
'a tall woman'

[1] The case markers in the examples in (1) also contain an indefiniteness marker.

c. riğāl-un ṭiwāl-un
men-NOM tall.M.PL-NOM
'tall men'
d. nisā'-un ṭawīl-āt-un
women-NOM tall-F.PL-NOM
'tall women'

Adjectives also agree with the noun in case:

(2) a. ra'aytu imra'at-an tawīl-at-an
I.saw woman-ACC tall-F-ACC
'I saw a tall woman'
b. nağlisu ḥawla al-ṭāwilat-i al-mustadīrat-i
we.sit around the-table-GEN the-round-GEN
'we sit down around the round table'
(SASG p. 153)

There is, however, another phenomenon, which distinguishes the Arabic concord pattern from that of Romance languages: there is also agreement in definiteness. The adjective takes the same definiteness marker as the noun it modifies:

(3) a. rağul-u-n ṭawīl-u-n
man-NOM-INDEF tall-NOM-INDEF
'a tall man'
b. al-rağul-u al-ṭawīl-u
the-man-NOM the-tall-NOM
'the tall man'
c. fī 'amrīkā -l-lātīniyyat-i
in America.GEN Latin-GEN
'in Latin America'
(SASG p. 153)
d. 'ağlisu ᶜalā maqᶜad-in fāḫir-in ğildiyy-in
I.sit on chair-GEN luxurious-GEN leather-GEN
'I sit down in a luxurious leather chair'
(SASG p. 153)

(3a) and (3b) show the contrast between an indefinite noun and a definite one: an adjective has the same definiteness marker as the noun, either *-n* or *al-*. (3c) shows that this is not merely a copying of the determiner: the proper noun *'amrīkā* does not have a determiner but is inherently definite. The adjective accompanying the noun takes the determiner in agreement with this. (3d) is provided as an extra example, and can be contrasted with (2b).[2]

[2]This phenomenon of definiteness agreement seems very similar to the phenomenon of Determiner Spreading found in Greek, (see, for example, Androutsopoulou 1995 and Alexiadou & Wilder 1998), but there are some differences. DS in Greek is not obligatory. Adjectives that allow it, do not have to undergo it. Alexiadou & Wilder even claim that there is a limited possibility to have partial DS; i.e. a determiner on one adjective but not on another in the same noun phrase. In Arabic, however, definiteness agreement is

4.2 The internal structure of the adjective phrase

4.2.1 Agreement

The common assumption is that concord consists of a direct agreement relation between the head noun and its modifiers. This position is taken by Carstens (2000), for example. However, data from Arabic suggests that there is more going on than that. Take the following phrase:

(4) li -l-ğazā'ir-i -l-mutaqaddim-i ḏikr-u-hā
to the-islands.F-GEN the-preceding.M-GEN mentioning.M-NOM-their
'to the aforementioned islands'

The construction in (4) has no equivalent in English. The head of the phrase is the noun *al-ğazā'ir* 'the islands'. It takes genitive case because of the preposition *li*. The noun is modified by an adjectival participle, *al-mutaqaddim* 'preceding'. However, although it is modified by the participle *al-mutaqaddim*, the noun *ğazā'ir* 'islands' is *not* the subject of the participle. The subject of the participle is *ḏikr-u-hā* 'their mentioning'. This is a gerund-like deverbal noun, modified by a pronominal suffix *-hā*. This resumptive pronoun expresses the object of the action expressed by the deverbal

obligatory:

(i) a. *al-raǧul-u ṭawīl-un
the-man-NOM tall-NOM.INDEF
'the tall man'
b. *al-ṭāwilat-u mustadīrat-un
the-table-NOM round-NOM
'the round table'

The examples in (i) cannot have the indicated meanings. (They are in fact grammatical with a sentential reading: *the man is tall* and *the table is round*.)

Another difference is that in Greek, DS is only allowed with so-called predicative adjectives:

(ii) a. o ipotithemenos (*o) dolofonos
the alleged (*the) murderer
b. *o dolofonos itan ipotithemenos
the murderer was alleged
(Alexiadou & Wilder 1998)

This is notably different in Arabic. All adjectives are required to agree in definiteness with the noun they modify, no matter whether they are predicative or not:

(iii) a. al-qātil-u (+al)-maz^{c}ūm-u
the-murderer the-alleged
'the alleged murderer'
b. *al-qātil-u maz^{c}ūm-un
the-murderer alleged-INDEF
'the murderer is alleged'

As shown in (iiib), the adjective *maz^{c}ūm* cannot be used as a sentence-level predicate, which indicates it is not a predicative adjective. However, as (iiia) shows, the determiner is still required when the adjective is used attributively.

noun, and it refers back to 'islands'. Note that the noun *ḏikr-u-hā* has nominative case, since it is the subject of the participle.

The combination *mutaqaddim ḏikr-u-hā* means 'their mentioning preceding'. When it is used attributively with the noun *ǧazā'ir*, the whole has the meaning 'the islands whose mentioning preceded', which is best translated in English as indicated.

The agreement facts in (4) are particularly interesting. The head noun *al-ǧazā'ir* is feminine plural, definite and has genitive case. The subject of the participle, *ḏikr-u-hā*, is masculine singular, definite and has nominative case. Somewhat surprisingly, the participle *al-mutaqaddim* shows a mixed set of features. It is masculine singular, definite and has genitive case. That is, its φ-features are assigned by its subject, *ḏikr-u-hā*, whereas its case and definiteness features are assigned by the noun it modifies, here *ǧazā'ir*.

The following examples show the versatility of this construction:

(5) a. ra'aytu -mra'-at-an ǧamīl-an waǧh-u-hā
I.saw woman-F-ACC.INDEF beautiful.M-ACC.INDEF face.M-NOM-her
lit. 'I saw a woman beautiful her face'
'I saw a woman with a beautiful face'

b. ǧā'at min balad-in maᶜrūf-at-in
it.came from country.M-GEN.INDEF famous-F-GEN.INDEF
šidd-at-u ḥarārat-i-hi
strength-F-GEN heat-GEN-its
lit. 'it came from a country famous the strength of its heat'
'it (the heat) came from a country famous for (the strength of) its heat'
(SASG p. 187)

c. 'ilā silsilatin ǧadīdatin min al-ḥurūb-i_i -l-ṣaᶜb-i
to chain new of the-wars.F-GEN the-difficult.M-GEN
-l-taḥakkum-u bi natā'iǧ-i-hā$_i$
the-containing-NOM with results-GEN-their
lit. 'to a new chain of wars their effects difficult to contain'
'(this tension could lead) to a new chain of wars whose effects will be difficult to contain'
(SASG p. 187)

First of all, the examples show that the construction is not limited to participles, but also occurs with adjectives. They also provide extra illustration of the two agreement processes. In (5a), the modified noun, *imra'a* 'woman', is feminine, indefinite, and takes accusative case. The modifying adjective, *ǧamīl* 'beautiful', is masculine, agreeing with *waǧh* 'face', but the adjective is at the same time indefinite, agreeing with *imra'a* rather than with *waǧh-u-hā*, which is definite. Note that the adjective also has accusative case, like the head noun.

Both (5b) and (5c) show a difference in gender between the head noun and the modifying adjective. In (5b), the head noun is *balad* 'country', which is masculine, whereas the modifying adjective is *maᶜrūfa* 'famous', which has a feminine form. The subject of this adjective is feminine also: *šiddat al-ḥarāra*. This example clearly

shows that the modifying adjective agrees in gender with its DegP-internal subject, not with the head noun. (5c) is similar: the head noun *ḥurūb* 'wars' is feminine. This noun is also an inanimate plural, which means it will trigger feminine singular agreement. The modifying adjective *ṣaᶜb* 'difficult', however, is masculine. The subject of the adjective, *taḥakkum* 'containing', is also a masculine noun. This shows that the adjective agrees in number with its own subject, not with the head noun.

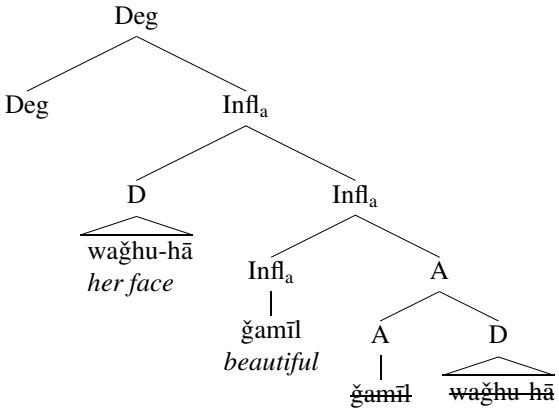
The last example, (5c), furthermore shows that the resumptive pronoun does not have to occur on the subject of the adjective. Here, the subject is a deverbal noun, *al-taḥakkum* 'the containing', and the resumptive pronoun occurs on the object of that infinitive *natā'iğ-i-hā* 'their results'.

What all these examples clearly show is that there is not one but there are two agreement processes involved in the adjective concord in (4) and (5). Agreement in φ-features is distinguished from agreement in case and definiteness. In other words, the way in which concord is established is more complex than usually assumed.

Let us look at this structure to see how we can analyse it. I will only look at the adjectival phrase for the moment. Taking (5a) as an example, this phrase contains two elements: the A head *ğamīl* and the subject *wağhu-hā*. I will follow proposals by Abney (1987) and Zwarts (1992) that the adjective phrase is a DegP.

The evidence shows that there is an agree relation between the adjective and its subject. The initial assumption that I will make is that the subject is generated as a sister to the adjective and moves to the specifier position of some agreement position, which I will call Infl_a:

(6)



The adjective moves to Infl_a in order to pick up the agreement features. In this tree I have positioned the subject of the adjective in spec,InflP_a, having moved from comp,AP.[3]

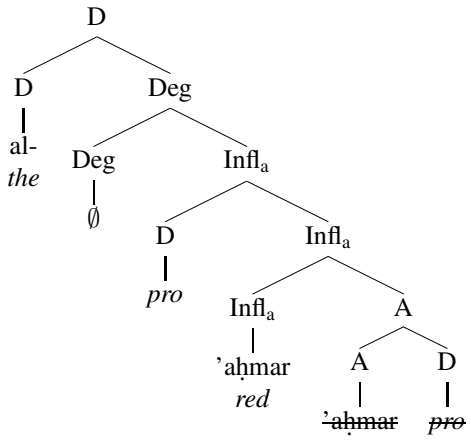
[3]Note that this is just a preliminary structure. Further evidence will show that the adjective is higher in the tree, and I will also argue that the subject is generated outside the AP. (Although there are probably cases where the subject *is* generated inside the AP, e.g. ergative adjectives. I will not go into this matter, however.)

So far we have only seen examples where a DP-internal adjective in Arabic has an overt subject. Most DP-internal adjectives in Arabic do not have such an overt subject. Usually, they just modify the noun, as in (7):

(7) a. al-baytu -l-'aḥmaru
the-house-NOM the-red-NOM
'the red house'
b. 'abniyat-u landan al-qadīmat-u
buildings-NOM London the-old-NOM
'the old buildings of London'
(SASG p. 187)

We have two options open to us. We can either say that the structure of the examples in the previous section is exceptional, and assume that examples such as (7) have a much simpler structure. The alternative is to say that (7) has a structure very similar to that of the earlier examples. Because a unified analysis of adjectives is preferable, I will assume that the latter is in fact the case. This means we must posit the presence of an empty element in the DegP-internal subject position:

(8)



Here the argument of the adjective is syntactically realised as a *pro* element. The structure is essentially the same as the one for (4): the adjective has its own subject with which it agrees. The only difference is that this subject is now a covert element: *pro*. This *pro* is the resumptive pronoun that we also saw in the structure of (4).

4.2.2 Genitive complements of adjectives

In the previous section, I have tentatively assumed that the DegP-internal subject is generated as a sister of the adjective, that is, inside the AP. I already indicated that I would not maintain that analysis. The DegP-internal subject is not the only argument that adjectives in Arabic can take. They can also take genitival complements. Such a genitival complement is an internal argument of the adjective. The subject is an external argument, and is therefore generated outside the AP (Zwarts 1992).

The construction that I will discuss in this section has already been described for Hebrew by Siloni (1998), Hazout (2000) and Kim (2000), and the Arabic construction resembles the Hebrew one closely, although there are some differences.

The basic construction is (9):[4]

(9) imra'-at-un ğamīl-at-u -l-waǧh-i
woman-F-NOM.INDEF beautiful-F-NOM the-face-GEN
'a woman with a beautiful face' (lit. 'a woman beautiful of face')

The adjective is in construct state and is followed by a noun with genitive case. The genitive noun must have a definite article.[5] The adjective agrees with the head noun (*imra'a* 'woman') in number, gender and case, and also in definiteness. This last fact may be surprising, since the adjective is in construct state. But an adjective in this construction can take an additional definite determiner:[6]

(10) al-mar'-at-u -l-ğamīl-at-u -l-waǧh-i
the-woman-F-NOM the-beautiful-F-NOM the-face-GEN
'the woman with a beautiful face'

In spite of the fact that the adjective is modified by a genitive noun, it still takes a definite article. Note that the adjective is indeed in construct state. We can see this when we consider the plural ending *-ūna*. As we saw in the previous chapter, this ending takes the form *-ū* when it appears on a noun or adjective in construct state. (11) contains an adjective that is obviously in construct state, as shown by the special form of the plural ending, but it still takes a definite determiner:

(11) (al-šabāb-u) -l-ḥadīṯ-ū -l-taḫarruǧ-i
(the-youths-NOM) the-new-PL.NOM the-graduation-GEN
lit. '(the youths) new of graduation'
'the new graduates'
(SASG p. 179)

[4]Note that we already encountered one example of this construction in example (42) of chapter 3.

[5]Unlike the Hebrew construction, where the genitive noun can be indefinite.

[6]This, too, is markedly different from the Hebrew construction. In Hebrew, it is the genitive noun that takes the definite article to signal agreement in definiteness:

(i) a. na'ara sxorat se'ar
girl black hair
'a girl with black hair'
b. ha-na'ara sxorat ha-se'ar
the-girl black the-hair
'the girl with black hair'
c. *ha-na'ara ha-sxorat se'ar
the-girl the-black hair
'the girl with black hair'
(Hazout 2000)

When the noun is definite, as in (ib) and (ic), the adjectival construct takes a definite article, which is placed on the noun as in (ib), not on the adjective as in (ic).

Siloni (1998) and Hazout (2000) argue that the head noun must be an inalienable possessor of the genitive noun. This is certainly not the case in Arabic, as the following examples show:

(12) a. bayt-un kaṯīr-u -l-'abwāb-i
house-NOM many-NOM the-doors-GEN
'a house with many doors' (lit. 'a house many of doors')
(SASG p. 176)

b. al-raǧul-u -l-ᶜaẓīm-u -l-ḥaẓẓ-i
the-man-NOM the-great-NOM the-luck-GEN
'the man who is very lucky' (lit. 'the man great of luck')
(SASG p. 176)

c. 'āṯār-un bāliġat-u -l-ḫuṭūrat-i
effects-NOM extreme-NOM the-danger-GEN
'extremely dangerous effects' (lit. 'effects extreme of danger')
(SASG p. 177)

d. al-šarikāt-u -l-mutaᶜaddidat-u -l-ǧinsiyyāt-i
the-companies-NOM the-multiple-NOM the-nationalities-GEN
'the multi-national companies' (lit. the companies multiple of nationalities')
(SASG p177)

None of the genitive nouns in the examples in (12) (*'abwāb* 'doors', *ḥaẓẓ* 'luck', *ḫuṭūra* 'danger' and *ǧinsiyya* 'nationality', respectively) are inalienably possessed by the head nouns (*bayt* 'house', *raǧul* 'man', *'āṯār* 'effects' and *šarikāt* 'companies').

In the analysis of Siloni (1998) and Hazout (2000), the genitive noun is actually the argument that the adjective is predicated of. In other words, the genitive noun fills the external argument position of the adjective, that is, its subject position. This analysis seems reasonable because in a phrase such as *the girl black of hair* it is obviously the hair that is black, not the girl. However, if the genitive noun were the external argument, one would expect that the adjective agrees with it, which is not the case.

The idea that the adjective is predicated of the genitive noun seems to be supported by the fact that the two phrases in (13) are very similar in meaning:

(13) a. al-mar'-at-u -l-ǧamīl-at-u -l-waǧh-i
the-woman-F-NOM the-beautiful-F-NOM face.M-GEN
lit. 'the woman beautiful of face'
'the woman with beautiful face'

b. al-mar'-at-u -l-ǧamīl-u waǧh-u-hā
the-woman-F.NOM the-beautiful.M-NOM face.M-NOM-her
lit. 'the woman beautiful her face'
'the woman with the beautiful face'

(13a) is the structure under consideration, (13b) is the structure discussed in the previous section, in which the adjective has a DegP-internal subject. This subject is the external argument of the adjective, as described above. Because (13a) is very

similar in meaning to (13b), it stands to reason to assume that the noun *waǧh* 'face' is the subject of the adjective in (13a) as well.

However, the adjective+genitive construction under consideration can be used as a sentence-level predicate:

(14) a. hāḏihi -l-mar'atu ǧamīl-at-u -l-waǧh-i
this.F the-woman-NOM beautiful-F-NOM the-face-GEN
'this woman has a beautiful face' (lit. 'this woman is beautiful of face')

b. hāḏihi -l-ṣaḥrā'-u ᶜadīmat-u -l-ḥayāt-i
this the-desert-NOM empty-NOM the-life-GEN
'this desert is void of life'

Because the structure can be used as a sentence-level predicate, we must conclude that it has an open argument position. A structure that has no open argument position cannot be used as a predicate, for the simple reason that it has no argument position available for the element it is to be predicated of. In normal adjectives, the open argument position is the external argument, that is, the subject of the adjective. This means that the genitive noun in (14) cannot fill the external argument position. If it did, it would not be possible to predicate the adjective-genitive combination of something else, as happens here.

Therefore, we must conclude that the genitive noun in these structures fills the position of an internal argument, not the position of the external argument. What I will argue is that the noun actually fills the position that Higginbotham (1985) and others call the *attribute*. Higginbotham says: "When an adjective combines with an N to form a complex N′, as in *tall man, big butterfly*, or *good violinist*, then it is taken as grading with respect to the attribute given in the N." (Higginbotham 1985, p. 563) In other words, in a phrase such as *a big butterfly*, the noun *butterfly* is not only the element of which the adjective is predicated, it also provides the attribute with respect to which the adjective is graded.[7] That is, the noun does not fill one but two positions in the adjective's thetamatic grid.[8]

Normally, it is the noun itself that gives the value of the attribute. What I will say is that in the construction under consideration it is not the noun itself but an aspect of the noun that provides the value for the attribute, and that it is this aspect that is expressed by the genitive noun. This accounts for the observation that the head noun must be a possessor of the genitive noun: the attribute is an aspect of the noun. We also see why the adjective-genitive structure is so similar in meaning to the construction in (13b): a phrase such as *a girl black of hair* can be paraphrased as *a girl that is black with respect to her hair*, in which it is obviously the hair that is black.

The idea that the genitive noun is an internal argument is supported by the fact that the structure under consideration also occurs with participles, in which case the genitive noun expresses the object, i.e. an internal argument:

[7] To clarify: *a big butterfly* is something that is a butterfly and that is big *for a butterfly*. The phrase *for a butterfly* expresses the attribute. See Higginbotham (1985) for discussion.

[8] More precisely, Higginbotham says that one argument position of the adjective is *theta-identified* with the noun, whereas the other is *autonymously theta-marked* by the noun. I will not go into the distinction between these notions here.

(15) a. zayd-un al-ḍārib-u ra's-i -l-ǧānī
Zayd-NOM the-hitter-NOM head-GEN the-perpetrator.GEN
'Zayd, who hits the head of the perpetrator'
b. (al-riǧāl-u) al-muqīm-ū -l-ṣalāt-i
(the-men) the-performing-PL the-prayer-GEN
'(the men) performing the prayer'
(Wright 1981, vol ii, p. 222)

In (15a), the participle *ḍārib* 'hitting' is followed by a genitive, *ra's al-ǧānī* 'the head of the perpetrator', which is the object of the action described by the participle. The participle modifies a proper name, (which is definite despite its indefinite form) and therefore takes the definite article. We have seen the same phenomenon with the adjectives taking a genitive attribute: in spite of the construct state of the adjective, the definite article is still allowed.

(15b) shows that the participle is indeed in construct state. As explained in the previous chapter, the masculine plural ending *-ūna* drops *-n*V when in construct state. It is this form that is used on the participle *muqīmū* 'performing', of which the absolute (non-construct) state form is *muqīmūna*.

As we can see, the construction in (15) has the same properties as the construction used for expressing the attribute argument: the adjective is modified by a genitive noun, and it fully agrees with the head noun. This observation confirms the point made in the previous chapter: one of the tasks of the functional complex is to create positions for arguments of the projecting lexical item. The Poss head of the adjective phrase licenses an internal argument of the adjective, whether it be an attribute or a complement.[9]

Let us now look at the structural analysis of these adjective constructions. Since they can license a genitive noun, there must be a Poss head present in these structures. And since the adjective shows agreement with the head noun, an Infl_a head must also be available. The question then becomes in which order the two occur. The answer to this question is straightforward, because the Poss head licenses an internal argument,

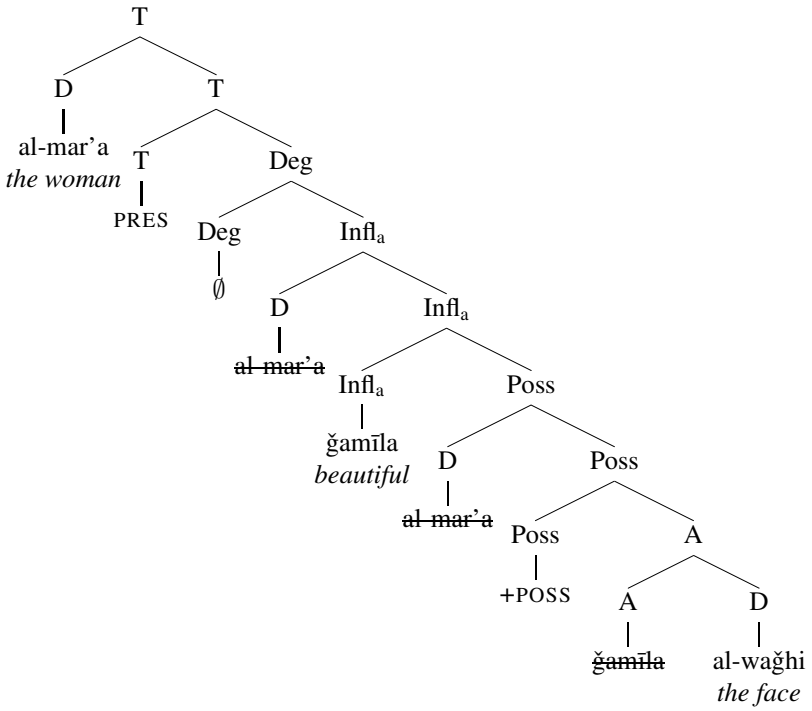
[9]Given that the attribute of the adjective and the object of the participle are internal arguments and that they are licensed with the genitive, one may wonder why other types of internal arguments of adjectives cannot be licensed with the genitive, but instead require inherent case or a preposition:

(i) a. raǧul-un faḫ̱ūr-un bi -bn-i-hi
man-NOM proud-NOM with son-GEN-his
'a man proud of his son'
b. *raǧul-un faḫ̱ūr-u -bn-i-hi
man-NOM proud-NOM son-GEN-his
'a man proud of his son'

(ia), with the object of *faḫ̱ūr* 'proud' expressed with the preposition *bi* 'with' is grammatical, but if the preposition is left out, the phrase becomes ungrammatical. This is unexpected if we assume that the prepositional object is an internal argument, like the attribute and the object. I have no explanation why thematic objects of adjectives have to be licensed with a preposition, but it seems that this is quite a consistent phenomenon across languages. Adjectives usually require prepositions or inherent case to license thematic arguments.

whereas the Infl_a head is responsible for the agreement between the adjective and its external argument. We must conclude, then, that Poss is dominated by Infl_a:

(16)



As I argued in the previous chapter, Poss assigns genitive case but it does not attract the element to which it assigns it, here *waǧh* 'face'. This element, the attribute of the adjective, is generated inside the lexical projection of the adjective, since it is an internal argument. I have used spec,Poss as the position where the subject is base-generated. The subject of an adjective, as argued by Zwarts (1992), is the adjective's external position. This effectively means that the subject must be generated in a position outside the lexical projection of the adjective. Spec,Poss is the most obvious candidate for this position.

So far, we have established that the adjective phrase contains at least two functional projections: an inflectional head Infl_a and a genitive-assigning head Poss. In this respect, it is much like the clause if we compare Infl_a to T and Poss to v. However, it is commonly assumed that Poss is to be equated with T (e.g. Szabolcsi 1994), something that is supported by the analysis in chapter 3. This is an important discrepancy. The nominal equivalent of T seems to be Poss, but the adjectival equivalent of T is obviously Infl, with Poss below it.

The analysis suggests, then, that we cannot say that the functional shells of the clause and the adjective are exact parallels of each other. Rather, we must conclude that the parallelism is more flexible: individual heads can be equivalent, but it is not the case that the functional structure as a whole is identical.

4.2.3 The Deg head

Abney (1987) argues that there is another functional head in the noun phrase. (See also Zwarts 1992). Abney designates it Deg, for degree, which he sees as the adjectival equivalent of C and D. Abney notes that adjectives can usually be preceded by an element that expresses a measure or degree, and he argues that this feature is present on the Deg head. Typical examples are the following, where the Deg element is slanted:

(17) a. this is *too* heavy to carry
b. he is not *that* rich
c. *how* beautiful it is!
d. this is not *as* bad as we expected[10]

The Deg head is also the locus of the comparative and the superlative. In English, there are two ways to form the comparative and superlative: by adding a suffix *-er* and *-est*, or by using *more/most*. Abney argues that both are expressed in Deg and that in the case of the suffix, A moves to Deg.

The interesting thing to note is that Arabic does not have any of the elements listed in (17). In fact, there is usually no simple and straightforward way in Arabic to express what they express. Instead, one needs to use descriptive phrases. (18), for example, would be a way to express the notion of *too*:

(18) fa huwa ’akbar-u min-mā yaḥullu-hu ᶜaql-un wāḥid-un
and it bigger-NOM than-what solves-it mind-NOM one-NOM
‘and it is too big for one mind to solve’
(lit. ‘it is bigger than what one mind solves’)

The translation of (18) shows that the phrase is the equivalent of *too big to solve*, but it is constructed as *bigger than what (one mind) can solve*. There is simply no way to express *too* directly.

Another example is (19):

(19) kuntu ’aᶜğabu limāḏā yabdū ṣaġīr-an hākaḏā
I.was I.wonder why it.seems small like-that
‘I was wondering why it seemed so small’
(SASG p. 175)

(19) may seem at first sight to be an example of a Deg head: *hākaḏā* ‘so, like that, in that manner’ might be considered a Deg head just like English *so*. There are some facts that argue against this, however. First of all, heads in Arabic generally precede their complements. This *hākaḏā* would be the only head that follows its complement. Furthermore, *hākaḏā* is generally used as an adverb with the meaning *like that, in that manner*. In this use, it is certainly not a Deg head, but rather an Adv head. The position that *hākaḏā* occupies in (19b) is typically a position of adverbs, as is shown by (20):

[10]The element *as* requires a phrase starting with *as* to follow. For more discussion, see Abney (1987).

(20) al-malābis-u -l-šatawiyyat-u -l-ṯaqīlat-u naw^{c}an mā
the-clothes-NOM the-winter-NOM the-thick-NOM sort-ACC some
'the somewhat thick winter clothes'
(SASG p. 175)

The phrase *naw^{c}-an mā* 'somewhat' is formed of the noun *naw^{c}* 'sort, type' in the accusative indefinite plus the modifier *mā* which emphasises the indefiniteness of the preceding word. The accusative on the noun shows that it is an adverbial modifier. *Hākaḏā* in (20b) is in the same position as *naw^{c}-an mā*. The fact that *hākaḏā* is an adverb in other contexts and the fact that the position it occupies in (19) is a position that adverbials can appear in suggest that *hākaḏā* is an adverb in (19), and not a Deg head.

There is one exception to the observation that Arabic does not have Deg heads. A subset of Arabic adjectives have a form called the elative, which can be used to express the notion of comparative and superlative:

(21) a. huwa 'aṭwal-u min-nī
he taller-NOM than-me
'he is taller than me'
b. hāḏihi 'aqdam-u -l-mudun-i fī -l-^{c}ālam-i
this.F oldest-NOM the-cities-GEN in the-world-GEN
'this is the oldest city in the world'
(lit. 'this is the oldest of the cities in the world')

In (21a) *'aṭwal*, the elative form of *ṭawīl* 'tall', is used. It is indefinite and functions as the equivalent to the comparative in English. In (21b), the elative form is *'aqdam*, from the adjective *qadīm* 'old'. It is definite here, and has the value of a superlative.[11]

[11]There are other constructions in which the elative is involved. For example, it can be followed by a singular indefinite noun in the genitive:

(i) hāḏihi 'aqdam-u madīnat-in fī -l-^{c}ālam-i
this.F oldest-NOM city-GEN.INDEF in the-world-GEN
'this is the oldest city in the world'

In this use, the elative also expresses a superlative. (i) has the same meaning as (21b).
The elative can also be used as a modifier of an indefinite noun:

(ii) yaḥtāǧu 'ilā qamīṣ-in 'akbar-a min ḏālika
he.needs to shirt-GEN.INDEF bigger-GEN than that
'he needs a bigger shirt (than that)'

As the translation shows, the elative in this construction also has the value of a comparative.
The elative can also have the meaning of the positive grade, but with an intensive meaning. This use is generally found in fixed expressions:

(iii) a. al-qurūn al-wusṭā
the-centuries the-middle
'the Middle Ages'
b. barīṭāniyā -l-^{c}uẓmā
Britain the-great
'Great Britain'

I will not discuss the peculiarities of the elative here. For the present discussion, it is sufficient to note that the form exists and that the morphology that expresses it is generated in Deg. This leads to the conclusion that the adjective must move to Deg, for it must pick up the elative feature in order to receive its specific morphological form.

This in itself does not explain why Arabic does not have any elements comparable to English *too, as, how, that*, etc. After all, Abney (1987) argues that A-to-Deg movement takes place in English as well, but only when the comparative/superlative is expressed with the suffix *-er* or *-est*. Then why would it not be possible for the Arabic adjective to move only when it is an elative?

The difference between Arabic and English in this respect is that in English, the synthetic comparative and superlative are expressed with suffixes. Deg contains *-er* or *-est* as lexical items which cannot stand on their own. As such, they have the ability to attract A. In other cases, the Deg head is not filled with any morphological material, or it is filled with an element that can stand on its own. Therefore, no A-to-Deg movement is necessary.

In Arabic, however, Deg *always* contains morphological material that needs to combine with a lexical stem. A root in Arabic consists of (usually) three consonants, to which a vowel pattern is applied to form a word that can be pronounced. We have already seen examples of this process in the singular/plural alternation in words such as *raǧul – riǧāl* 'man, men', in which the consonants *r-ǧ-l* form the root of the word and the vowel patterns *-a-u-* and *-i-ā-* express number.

The elative form of the Arabic adjective is also formed with a specific vowel pattern: *'a--a-*.[12] So for example the adjective *sāhir* 'sleepless; vigilant' has an elative *'ashar*. The consonantal root is *s-h-r*, the vowel pattern *-ā-i-* expresses the positive grade, the pattern *'a--a-* expresses the elative. Similarly, the adjective *kabīr* 'large' has an elative form *'akbar*. Here, the vowel pattern of the elative is the same, the pattern of the positive grade is different: *-a-ī-*. Both patterns *-ā-i-* and *-a-ī-* are very common among adjectives.

The facts indicate that in the adjective it is the vowel pattern that expresses the degree. The patterns *-ā-i-* and *-a-ī-* express a positive grade, the pattern *'a--a-* expresses an intensive grade, called the elative. These patterns are generated in the Deg head, not in the A head. Because Deg is filled with this pattern, there is no room for elements like the English *too, that, how* or *as*.[13]

Both adjectives in (iii) are in the elative form: *wusṭā* is the feminine elative of *wasīṭ* 'middle', *ʿuẓmā* is the feminine elative of *ʿaẓīm* 'great'. Note that the elative only agrees in gender and number when it is used in this way. When it is used as a comparative or superlative, the form is always masculine singular.

[12]Actually, the elative also contains a consonantal affix: the first vowel *a-* is preceded by a glottal stop, which behaves like a consonant in Arabic morphology.

[13]This analysis has as a direct consequence that the structure that Alexiadou & Wilder (1998) propose for attributive adjectives in Greek cannot be applied to Arabic. They argue that an attributive adjective has the following structure:

(i) a. o ipotithemenos (*o) dolofonos
the alleged (*the) murderer

4.3 Definiteness agreement

I have already mentioned the phenomenon of definiteness agreement in chapter 3 and in the first section of this chapter. In this section, I will discuss the phenomenon and see how we can account for it. First, let us look at the relevant example again:

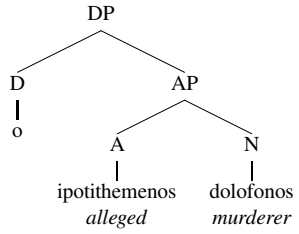
(22) a. rağul-u-n ṭawīl-u-n
man-NOM-INDEF tall-NOM-INDEF
'a tall man'
b. al-rağul-u al-ṭawīl-u
the-man-NOM the-tall-NOM
'the tall man'

As can be seen, the definiteness feature of the adjective manifests itself in the same way as it does on the noun phrase: indefiniteness is marked with a suffix *-n*, whereas definiteness is marked with the determiner *al-*.

In the previous chapter I analysed the determiner *al-* and the indefiniteness marker *-n* as projections of the head D. Apparently, this D head is present in the adjective phrase as well, even though Zwarts (1992) argues that the Deg head is the adjectival equivalent of D and C, which would mean that there can be no extra D head in the adjective phrase.[14]

Not only is the D head visibly present in the adjective phrase, it also has a function. In section 4.2.1 I reached the conclusion that every adjective phrase contains a DegP internal subject argument and a resumptive pronoun that refers back to the modified noun. DP-internal adjectives usually have *pro* as subject, which functions as the resumptive pronoun:

b.



This is basically the same structure that Abney (1987) proposes for all prenominal adjectives in English. Alexiadou & Wilder use it to account for the fact that Determiner Spreading is not possible with these adjectives: there is simply no position for the extra D head to appear in. The structure in (ib) is not possible in Arabic, because the A head only contains the adjectival root. The Deg head is needed to complete the adjective's morphological form, but in (ib) no Deg head is present.

The fact that (ib) is an impossible structure in Arabic is not problematic, since typically attributive adjectives such as *alleged* and *former* require definiteness agreement in Arabic. They behave like any other adjective, so there is no reason to assume that they would have a different structure.

[14]Szabolcsi (1994) argues that the D and the C head should each be separated into two heads. If she is correct, the occurrence of both a D and a Deg head in the Arabic adjective phrase may not be problematic at all.

(23) al-baytu [$_{\text{DegP}}$ -l-'abyaḍu *pro*]
the-house the-white
'the white house'

Following Higginbotham (1985), who argues that all adjectival modification is in fact predication, we can give the following semantic representation of the adjective phrase:

(24) $\iota x(\textbf{house}(x) \wedge \textbf{white}(x))$

What (24) tells us is that the resumptive pronoun is in fact a variable. The variable x, which is bound by the outer ι-operator, also occurs in the adjective phrase. We can plausibly say that the *pro* element in the syntactic structure is the equivalent of this variable. And it is this variable that requires the presence of the adjectival determiner.

Under common assumptions, the (nominal) determiner functions as the binder of the open argument position in the noun phrase. This argument position is the *R* role of the noun, which is generally not syntactically realised. As such, the determiner is the syntactic equivalent of the semantic ι operator.

When we look at the structure of the adjective phrase, we see that the adjectival determiner is also a binder. The variable it binds is the resumptive pronoun present in the adjective phrase. The adjectival determiner functions as a binder for the resumptive pronoun, making sure that the adjective phrase can be used as a DP-internal modifier.

With adjectives that have an overt DegP-internal subject, the analysis is the same:

(25) a. ra'aytu -mra'-at-an ğamīl-an
I.saw woman-F-ACC.INDEF beautiful.M-ACC.INDEF
waǧh-u-hā
face.M-NOM-her
lit. 'I saw a woman beautiful her face'
'I saw a woman with a beautiful face'

b. $\iota x(\textbf{woman}(x) \wedge \iota y(\textbf{face}(y) \wedge \textbf{of}(x)(y) \wedge \textbf{beautiful}(y)))$

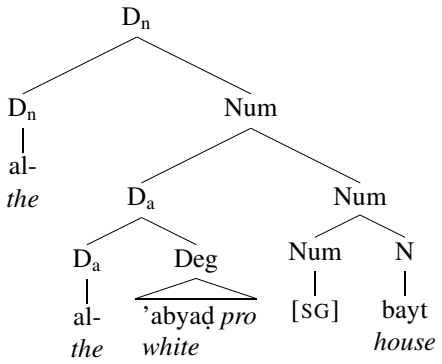
For convenience, I have used a predicate **of** to indicate possession.[15] Again we see that the adjective phrase contains a variable that refers back to (the *R* role of) the head noun. This variable in the syntactic structure of (25a) is the resumptive pronoun *-hā* 'her' which is the possessor of *waǧh* 'face'.

As we see, the adjectival D head functions as a binder for the resumptive pronoun present in the adjective phrase. However, when we look at the semantic structure, we see that there is only one operator that binds both occurrences of the variable x. In the syntactic structure, there are two binders: the nominal D and the adjectival D. This raises the question why the syntactic structure needs two binders.

This question becomes even more compelling when we examine the proposed tree structure for (25):

[15]Note that the variable y is the *R* argument of the adjective's subject 'face'. It is irrelevant to the point at hand.

(26)



The resumptive pronoun in the DegP is in the c-command domain of the nominal D_n, which would mean D_n should be able to bind it.

The answer to this question can be found in Chomsky's (1999) assumption that derivations are built up phase by phase. As explained in chapter 1, Chomsky defines phases on the basis of propositional content. Since the adjective phrase contains all the elements that make up a proposition (i.e. a predicate, the predicate's arguments and a subject) we can conclude that the adjective phrase is a phase. In other words, the adjective phrase is built separately, and only when it is finished is it included in the noun phrase.

It is reasonable to assume that the resumptive pronoun needs to be licensed locally, inside the phase it is contained in, i.e. adjective phrase. For this reason, a D head is added to the adjective phrase, which makes sure that the variable is bound, and in this way licenses it.

The D head that is inserted must of course be identified itself. Because it is at the edge of the phase, we can argue that this does not need to take place locally. There are basically two ways in which the D head can be licensed. First, the adjective phrase can be used independently, as in (27):

(27) al-ṭawīl-u
the-tall-NOM
'the tall one'

Here, the D head is identified in the same way that the D head of any noun phrase is identified.[16] If the adjective phrase is merged inside a noun phrase, modifying the head noun, it will be bound by the noun's D head. In this case, the features of the nominal D are transferred to the adjectival D. These features include DEF and CASE, but also the φ-features. The φ-features are then transferred to the resumptive pronoun, which is bound by the adjectival D.[17]

[16] Which is presumably some interpretational process beyond the scope of syntax.

[17] The exact nature of the binding that takes place between the nominal and the adjectival D heads needs further explanation, because it is not the typical operator binding. I will leave this matter to future research.

4.4 Relative clauses

We now have developed an analysis of the DP-internal adjective phrase in Arabic. It turns out that an adjective phrase has a clause-like structure, and that the adjective agrees with a subject internal to the DegP. Furthermore, we have seen that the adjective phrase contains a resumptive pronoun, and that the definiteness marker on the adjective plays a role in identifying this resumptive pronoun. In this section, I take a quick look at relative clauses, which appear to have a very similar structure.

As already discussed in chapter 3, a relative clause in Arabic is a clause with normal word order that follows the noun it modifies. The relative clause contains a resumptive pronoun and there is no *wh*-element.[18] The relative clause is introduced with a relative clause marker:

(28) al-raǧul$_i$ allaḍī ra'aytu-hu$_i$
the-man REL I.saw-him
'the man that I saw'

The relative clause marker agrees with the antecedent in gender and number. In (28), *allaḍī* is marked for masculine singular. When the antecedent is feminine and/or plural, it takes a different form:

(29) a. al-mar'a$_i$ allatī ra'aytu-hā$_i$
the-woman REL.SG.F I.saw-her
'the woman that I saw'
b. al-riǧāl$_i$ allaḍīna ra'aytu-hum$_i$
the-men REL.PL.M I.saw-them
'the men that I saw'
c. al-nisā'$_i$ allātī ra'aytu-hunna$_i$
the-women REL.PL.F I.saw-them
'the women that I saw'

The relative marker also has dual forms. These forms have an additional property: they agree with the head noun in case:[19]

(30) lam 'aǧidi -l-raǧulayni -llaḍayni baḥaṯā
not I.found the-men.DUAL.ACC REL.M.DUAL.ACC they(DU).searched
ᶜan-nī
for-me
'I did not find the two men that were looking for me'

Note that the accusative case of the relative marker *allaḍayni* is the same as the case of the antecedent *al-raǧulayni* 'the two men', but different from the nominative case of the resumptive pronoun in the relative clause (which in (30) is a *pro* subject).

[18] Substantive relative clauses, i.e. relative clauses without an antecedent, are formed with *wh*-elements, but I will not discuss those here.

[19] The reason that the singular and plural forms of the relative marker do not agree in case is probably due to the fact that they are frozen oblique case forms (Wright 1981).

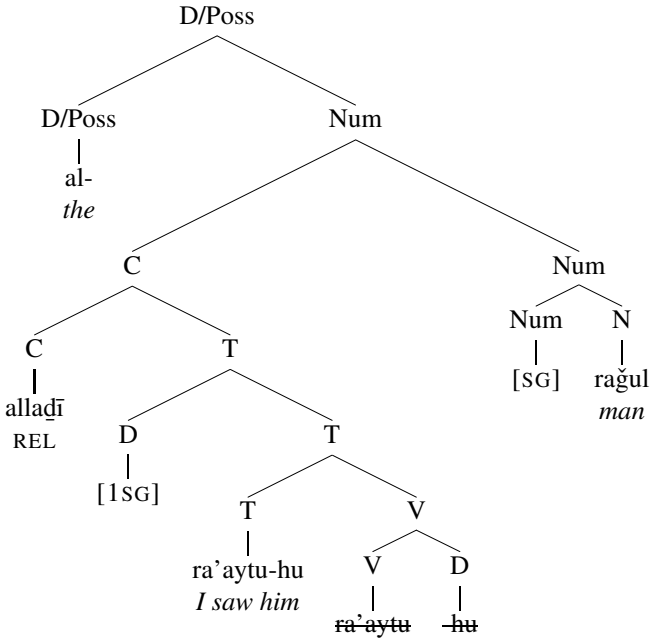
So we see that the relative marker agrees with the antecedent in gender, number and case. Interestingly enough, it also agrees in definiteness. When the antecedent noun is indefinite, the relative marker is dropped:

(31) ğā'a bi kitāb-in ∅ lam yaqra'-hu ba^{c}du
he.came with book-GEN.INDEF (REL) not he.read-it yet
'he brought a book that he had not read yet'

In (31), the antecedent *kitāb* is immediately followed by the relative clause. Like the previous cases, the relative clause contains a resumptive pronoun, but now there is no relative marker. These facts indicate that a relative clause in Arabic has a C head that contains the relative marker *allaḏī* and that agrees with the head noun. If we assume that relative clauses are adjoined to Num, just like adjectives, the structure of a phrase as in (32a) will be (32b):

(32) a. al-raǧul$_i$ allaḏī ra'aytu-hu$_i$
the-man REL I.saw-him
'the man that I saw'

b.



The structure of (32) is very similar to the structure of the adjective phrase in (26) above. In the adjective phrase, the D head is bound by the matrix D. In (32), the C head is also bound by the nominal D head and receives its features in this way. Furthermore, the relative marker C binds the resumptive pronoun in the clause in the same way that the adjectival D binds the resumptive pronoun in the adjective phrase.

It turns out, then, that we do not need any extra assumptions to explain the agreement in relative clauses. Relative clauses use the same mechanisms that adjectives use.

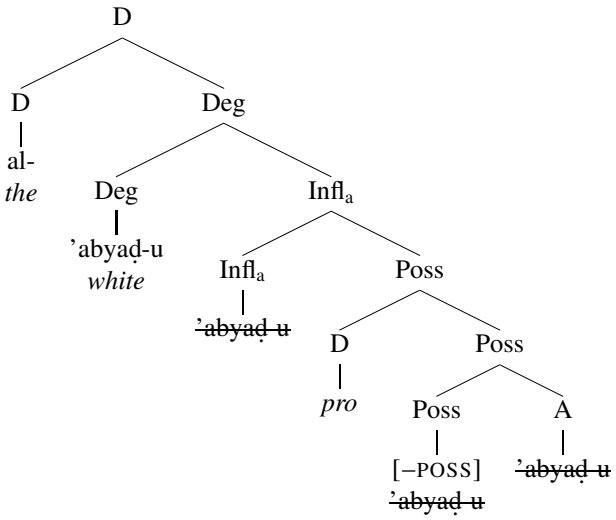
4.5 Linearisation

I have identified a number of elements in the adjective phrase, and I have given them a position in the hierarchical structure, but I have not yet discussed how we can derive actual word orders from this structure.

The heads in the adjective phrase that have an independent morphological form are D (when it is filled with *al-*) and A itself. The other heads are either affixal or non-overt. We have already established that D has $H > S$, (the adjunct setting is undefined for D because there are no known elements that adjoin to D) which means that we only need to determine the settings for A. The first hypothesis to test is obviously that it has the same settings as N. Let us see what this would give us. (33) contains a straightforward case of an adjective modifying a noun. Note that I only give the tree of the adjectival DP, not of the entire DP:

(33) a. al-bayt-u -l-'abyaḍ-u
the-house-NOM the-white-NOM
'the white house'

b.

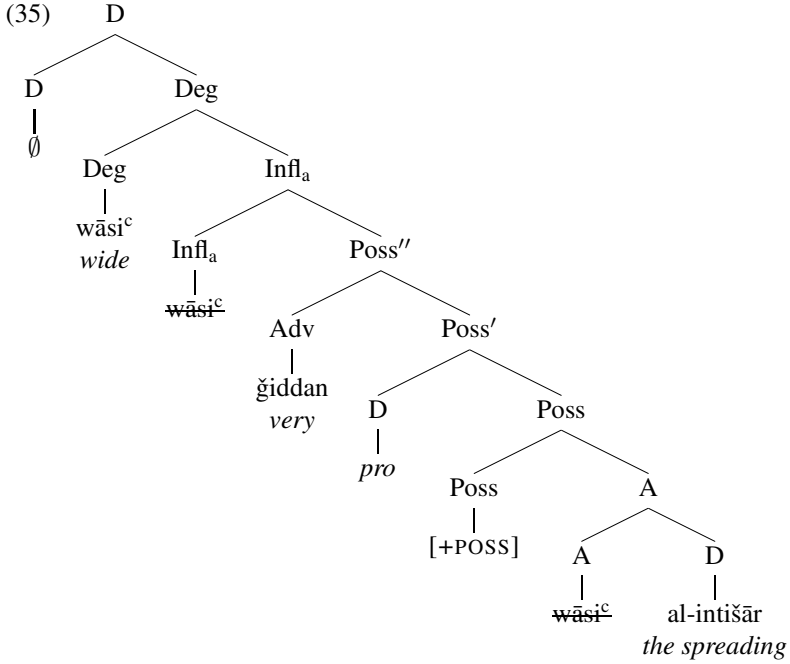


The ordering $H > S$ derives the correct linear order from (33b), but note that this example does not really tell us very much: the A head does not have a complement, so we can only use it to confirm the ordering of S and H for the category D. The example in (34) gives us more information:

(34) a. hāḏihi maǧall-at-un wāsi^{c}-at-u -l-intišār-i ǧiddan
this magazine-F-NOM wide-F-NOM the-spreading-GEN very
lit. ‘this is a magazine very wide of spreading’
‘this is a magazine with a very wide circulation’

(34) includes an adverb, *ǧiddan* ‘very’, which has not been discussed so far. Given the assumption that adjectives are adjoined to a head in the noun phrase, we can plausibly say that adverbs in adjectival phrases have a similar status. I will consider them specifiers of the lowest functional projection in the adjective phrase, which is Poss.[20] The tree structure that results is (35):

(35)



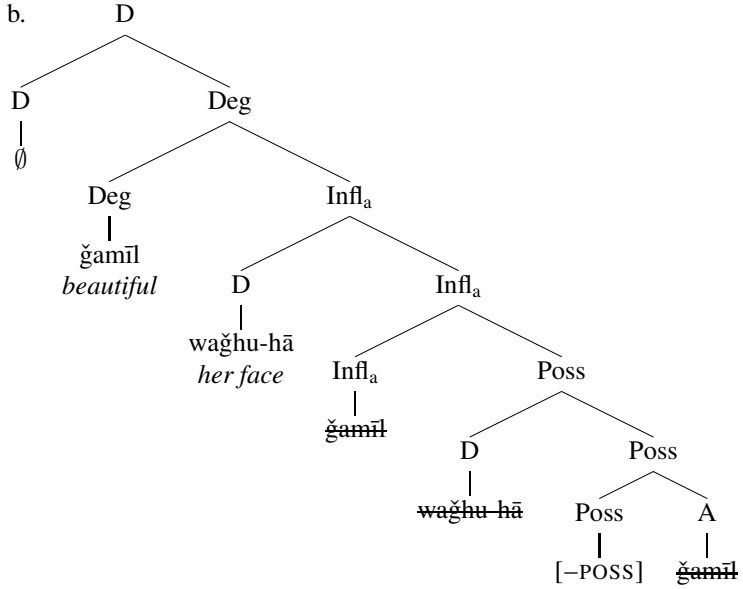
With an ordering of $H > S$ and adjunct-second, which is identical to the settings of parameters in the noun phrase, we derive the correct order: the heads are all linearised before their complements, which yields an order in which the complement of the adjective immediately follows the heads (of which only Deg contains overt material). The adverb *ǧiddan* ‘very’ is a non-selected specifier of Poss, and is linearised second in the node Poss″, which means it will follow all overt material in the branch Poss′, which contains the complement *al-intišār* ‘the spreading’.

With the type of construction discussed in section 4.3, in which the adjective has a DegP-internal subject, the correct order is also derived:

[20] In the noun phrase, the lowest functional projection is of course Num. There is something to be said for claiming that there is a Num head in adjectives as well, because like nouns, they are inflected for number. Like Poss, Num would adopt the linearisation parameters of A, which means it does not influence the ordering.

(36) a. imra'atun ğamīl-un waǧhu-hā
woman.NOM beautiful.M.NOM face-NOM-her
'a woman with a beautiful face'

b.



Note, by the way, that this tree shows the effect of the movement of the adjective to Deg: because of this movement, the adjective appears before the DegP-internal subject in the linear string.

4.6 Summary

Adjectives in Arabic agree with the noun they modify in gender, number, case and definiteness. However, when examined more closely, we see that there are actually two agreement processes: the adjective agrees with a DegP-internal subject in φ-features, and it agrees in case and definiteness with the noun it modifies. The DegP-internal subject can be expressed overtly in Arabic, in which case the adjective can be seen to agree with this subject in φ-features and with the head noun in case and definiteness.

The conclusion is that adjective phrases have a clause-like structure, with a subject and a head Infl$_a$ that establishes the agreement between the subject and the adjective. In cases where the adjective phrase does not have an overt DegP-internal subject, the subject is *pro*. This *pro* is a resumptive pronoun that refers back to the head noun. If the DegP-internal subject is overt, it will contain an overt resumptive pronoun.

At the semantic level, this resumptive pronoun is a variable. It will need to be bound, and for this reason, a D head is projected in the adjective phrase. Usually, another binder would be available once the adjective phrase has been inserted into a larger phrase, but there are configurations in which no such binder is present. The adjectival D head will make sure the variable has a binder no matter where the adjective

phrase is later inserted.

This analysis of the adjective phrase carries over *mutatis mutandis* to relative clauses. Relative clauses are projections of a V head rather than an A head, which means they will have a C head at the highest level, rather than a D head. This C head behaves exactly like the adjectival D head, however: it agrees with the head noun in case, definiteness, gender and number. Relative clauses also contain a resumptive pronoun, which is identified in the same way that the resumptive pronoun in the adjective phrase is. In other words, relative clauses are very similar to adjective phrases.

Adjectives have an external argument, the subject of which they are predicated, but they can also have an internal argument. Adjectival participles can project an object, but this is not the only internal argument that appears in adjectives. It is also possible that the attribute of the adjective is overtly realised. Normally, the attribute argument of the adjective is filled by the noun that is modified, but it is possible that the attribute is filled by a noun referring to an aspect of the noun, in which case the attribute will be expressed as a genitive argument of the adjective.

The linearisation of adjective phrases does not pose many problems. The heads that contain independent morphological material are D and A. The linearisation parameters of D have already been established in the previous chapter, and the linearisation parameters of A turn out to be the same as the parameters of N.

5

Deverbal nouns

Arabic verbal morphology includes a form that Arab grammarians have termed *maṣdar*. Western Arabists often employ the term 'infinitive' to refer to this verb form, but, unlike infinitives in Western languages, which are usually considered to be verbal forms, the Arabic infinitive is a nominal form. To emphasise this fact, I will avoid the term 'infinitive' and use the term *masdar* instead.

Two different uses of masdars can be distinguished. On the one hand, they can have a more verbal use, which closely resembles gerunds in English and other languages, and on the other hand, they have a more nominal use, which resembles that of simplex event and result nominals as formulated by Grimshaw (1990).[1]

5.1 Properties of masdars

I discussed the properties of Arabic nouns in chapter 3. In this first section, I will show that masdars are indeed nominal forms, since they have the properties of nouns. The most typical properties of nouns in Arabic is their ability to take the definite article, take case endings and form plurals, both broken and sound. The examples in (1) show that masdars are no exception:[2]

[1] In fact, in traditional Arabic grammar, only the verbal use of these forms is called the *masdar*, which means that the term has a meaning very similar to *gerund* in English. I will not follow this convention, however. Instead, I will use the term *masdar* to refer to the word form, comprising both the verbal and the nominal use.

[2] I will use the English gerund to gloss the Arabic masdar. Note, however, that I do not always use the gerund in the translation.

(1) a. waṣf-u-n; al-waṣf-u
describing-NOM-INDEF the-describing-NOM
'a description; the description'
b. 'awṣāf-u-n; al-'awṣāf-u
describing.PL-NOM-INDEF the-describing.PL-NOM
'descriptions; the descriptions'

(2) a. taṭawwur-u-n; al-taṭawwur-u
developing-NOM-INDEF the-developing-NOM
'a development; the development'
b. taṭawwur-āt-u-n al-taṭawwur-āt-u
developing-PL-NOM-INDEF the-developing-PL-NOM
'developments; the developments'

In (1), the masdar *waṣf*, from the verb *waṣafa* 'to describe', takes a broken plural *'awṣāf*. In (2), *taṭawwur*, from *taṭawwara* 'to develop', takes the sound plural ending *-āt*. Furthermore, the various forms show the definite article, the indefinite marker and case endings on the masdars.

Masdars also have the distribution of nouns. They appear in all positions in which non-event nominals appear: subject, object, complements of prepositions and adjuncts:

(3) a. tuqliqu hāḏihi -l-taṭawwur-āt-u -l-ḫubarā'-a
worry these the-developing-PL-NOM the-experts-ACC
'these developments worry experts'
b. qara'tu waṣf-a -l-ḥādiṯ-i
I.read describing-ACC the-accident-GEN
'I read the description of the accident'
c. lā 'aḫāfu min fi^{c}l-i hāḏā
not I.fear from doing-GEN this
'I do not fear doing it'
d. fa sakattu iḥtirām-an li law^{c}at-i-hi
and I.was.silent respecting-ACC.INDEF to grief-GEN-his
'I was silent out of respect for his grief'

In (3a), the (plural) masdar *taṭawwurāt* is subject of the verb *tuqliqu* 'worry'. In (3b), *waṣf* is object of *qara'tu* 'I read'. In (3c), the verb *'aḫāfu* 'I fear' takes a prepositional object. The complement of this preposition *min* 'from' is the masdar *fi^{c}l* 'doing'. Lastly, (3d) shows an example in which the masdar *iḥtirām* 'respecting' is used as an adjunct.

There is, however, one major difference in distribution between non-event nominals and masdars. The latter can appear as objects of verbs that semantically require a verbal complement. One such case is already shown in (3c). Other examples are given in (4):

(4) a. wa lākinna-ka taf^{c}alu mā lā ’astaṭī^{c}u fi^{c}l-a-hu
(and) but-you you.do what not I.am.able doing-it
‘but you do what I am unable to do’
(Cant. II p. 402)
b. ḥāwaltu ’an ’aṣifa -l-ḥādiṯ-a
I.tried that I.describe.SUBJ the-accident-ACC
‘I tried to describe the accident’
b′. ḥāwaltu waṣf-a -l-ḥādiṯ-i
I.tried describing-ACC the-accident-GEN
‘I tried to describe the accident’

In (4a), the masdar *fi^{c}l* ‘doing’ functions as object of the modal verb *’astaṭī^{c}u* ‘I am able to’. In (4b), the main verb is *ḥāwaltu* ‘I tried’. This verb requires a verbal complement, and is usually construed with a subjunctive complement clause, as in (4b). Alternatively, however, it can be construed with a masdar, as in (4b′). In this verbal use, the masdar does not allow pluralisation: (5) does not make any sense.

(5) *ḥāwaltu ’awṣāf-a -l-ḥādiṯ-i
I.tried describing.PL-ACC the-accident-GEN

What we see, then, is that masdars have two uses, which we can characterise as ‘nominal’ in (3), and as ‘verbal’ in (4). This distinction between a ‘nominal’ and a ‘verbal’ use of masdars is strongly reminiscent of ambiguities found in deverbal nouns in English. I will discuss this matter in the following sections.

Another property that masdars share with nouns is the ability to assign genitive case. Take (4b), for example, repeated as (6):

(6) ḥāwaltu waṣf-a -l-ḥādiṯ-i
I.tried describing-ACC the-accident-GEN
‘I tried to describe the accident’

The masdar *waṣf* has one argument: the object *al-ḥādiṯ*. It is expressed in a genitive construction: the masdar is in the construct state, having neither a definite article nor an indefinite marker. The argument takes genitive case and follows the masdar immediately.

In (6), the object of the verbal predicate expressed by the masdar is assigned genitive case. The genitive is not specific to the object, however. If the only argument of a masdar is the subject, this subject also takes genitive case:

(7) a. tazāyud-u -l-muᶜāraḍat-i -l-rasmiyyat-i -l-ᶜarabiyyat-i
increasing-NOM the-resistance-GEN the-official-GEN -the-Arabic-GEN
li ḍarb-i -l-ᶜirāq-i
to striking the-Iraq-GEN
'increasing (official) Arabic opposition to a strike against Iraq'
(lit. 'increase of the official Arabic resistance to a strike against Iraq')

b. yuḍḥiku-nī raqṣ-u-hu
makes.laugh-me dancing-NOM-his
'his dancing makes me laugh'

The masdar in (7a), *tazāyud* 'increase' has a genitive complement *al-muᶜāraḍa* 'resistance, opposition'. This complement serves as the subject of the verbal predicate of the masdar.[3] The fact that this verb happens to be unaccusative is of no consequence here: in (7b), the subject of the masdar *raqṣ* 'dancing', which is an accusative verb, also takes genitive case: the pronominal suffix *-hu* 'his' is the genitive form of the pronoun.

A noun can only license one genitive complement. This is no different for masdars. So when both subject and object are expressed, only one can take genitive case. The argument that takes genitive case is always the subject. The object must be licensed in another way, and there are two ways to do this. One way is to introduce the object with the preposition *li* 'for, to':

(8) a. šāhadtu tadmīr-a -l-qayṣar-i li -l-madīnat-i
I.witnessed destroying-ACC the-Caesar-GEN to the-city-GEN
'I witnessed Caesar's destroying of the city'

b. fa-qad kāna yaᶜrifu (...) min hā'ulā'i -l-šabāb-i
and-PART was he.knows (...) of these the-young-GEN
ḥubb-a-hum li -l-ᶜilm-i
loving-ACC-their to the-knowing
'he knew how much these young people loved learning'
lit. 'he knew of these young people their love of learning'
(Cant. II p. 404)

In (8a), the subject *al-qayṣar* 'Caesar' is the genitive modifying the masdar *tadmīr* 'destroying'. The object, *al-madīna* 'the city', is introduced with the preposition *li*. In (8b), the relevant masdar is *ḥubb* 'loving, love'. Its subject here is the pronoun suffix *-hum* 'them', its object is *al-ᶜilm*, 'knowing, knowledge', which is introduced with *li*.

[3]The phrase in (7a) shows how ubiquitous the masdar is in Arabic. Not only *tazāyud* 'increasing' is a masdar, but *muᶜāraḍa* 'opposing, resisting' and *ḍarb* 'striking, hitting' are, too.

The alternative method for expressing the object is to give it accusative case:

(9) a. šāhadtu tadmīr-a -l-qayṣar-i -l-madīnat-a
I.witnessed destroying-ACC the-Caesar-GEN the-city-ACC
'I witnessed Caesar destroying the city'

b. la-mā naqaṣa ḥubb-u-hā 'iyyā-ya ḏarrat-an wāḥidat-an
EMPH-not diminished love-NOM-her OBJ-me grain-ACC one-ACC
'her love for me did not diminish one bit'
(Cant. II p. 403)

c. (...) ṭalab-ī taḥwīl-a -l-ḥurriyyat-i li -l-nisā'-i
(...) demanding-my granting-ACC the-freedom-GEN to the-women-GEN
'(...) my demand to grant freedom to women'
(Cant. II p. 403)

(9a) shows the same phrase as (8a), but now with the object in accusative case. In (9b), the object is a pronominal suffix introduced with the particle *'iyyā*. This particle is used when a pronominal suffix is required as object, but the verb or masdar already has a suffix. This occurs for example when a double-object verb has two pronominal objects, or when a masdar has an overt subject and a pronominal object, as in (9b).[4] Lastly, in (9c), the masdar *ṭalab* 'demanding' has a suffix pronoun expressing the subject and an accusative object *taḥwīl* 'granting' expressing the object.

5.2 Comparison to English event nominals

The distinction between verbal and nominal uses of masdars is strongly reminiscent of the distinction between complex event nominals and simplex event/result nominal made by Grimshaw (1990). The basic idea of her theory is that some deverbal nouns retain the event and argument structure of the underlying verb. These nouns, which she calls complex event nominals, have certain verb-like properties that distinguish them from other nouns.

Other deverbal nouns do not retain the event and argument structure of the underlying verb. These are called simplex event nominals, if they refer to an event, or result nominals, if they refer to the result of an action or event. These nominals behave like non-event nominals in every way, and they do not share the typical properties of complex event nominals.

[4] Classical Arabic sometimes allows the use of two suffixes, as in (i):

(i) faqd-ī-hi
losing-my-it
'my losing it'

In Modern standard Arabic, (i) would normally be phrased as in (ii):

(ii) faqd-ī 'iyyā-hu
losing-my OBJ-it
'my losing it'

At first sight, masdars in their verbal use seem a typical instance of complex event nominals. In this section, I take a closer look at the properties that Grimshaw ascribes to complex event nominals, and see if they apply to verbal masdars as well. This comparison brings to light some of the most striking differences between verbal and nominal uses of masdars. In the next section, I will show how these differences can be accounted for.

Because most deverbal nouns in English are ambiguous between a complex event and a simplex event/result reading, Grimshaw introduces several means of disambiguation, such as the addition of certain aspectual modifiers and the expression of a subject. It is possible to use the methods she describes in Arabic, but there is another, clearer method. As I already remarked, masdars are often used in a position where a verb or a subclause would be expected. This is a very useful method of disambiguation, because the deverbal noun must retain the verb's event and argument structure in order to be able to take the position of a verb. One typical case is the light verb *tamma* 'to come to pass, happen'. This verb is often used to paraphrase the passive,[5] in which case it takes a complex event masdar as subject:

(10) a. tamma -ᶜtirāf-u-hu bi -l-ḏanb-i
happened confessing-NOM-his with the-crime-GEN
'he came to confess his crime'
(Fassi Fehri 1993, p. 236)

b. tamma -ktišāf-u ᶜilāğ-in ğadīd-in
happened discovering-NOM cure-GEN.INDEF new-GEN.INDEF
'a new cure has been discovered'
(Holes 1995, p. 258)

Another example is the use of a masdar in the position of a subclause. Such a masdar must have a complex event reading in order for it to be an accurate paraphrase of the subclause. One instance of this has already been given earlier in (4), repeated here as (11a):

(11) a. ḥāwaltu waṣf-a -l-ḥādiṯ-i
I.tried describing-ACC the-accident-GEN
'I tried to describe the accident'

b. 'adhaša-nī -ᶜtirāf-u-hu bi -l-ḏanb-i
astonished-me confessing-NOM-his with the-crime-GEN
'it astonished me that he confessed to the crime'
(lit. 'his confessing to the crime astonished me')

In (11a), the masdar is the complement of a control verb. In this case, the subject of the masdar is controlled by the subject of the matrix verb, *ḥāwala* 'to try'. In (11b), the masdar *iᶜtirāf* 'confessing' is in subject position and has the value of a subclause.

[5] Although, as Holes (1995) notes, most commonly with punctual and intentional actions, and with a perfective aspect. Note also that it is not a true passive, in that the subject can be expressed in the normal way, i.e., with a genitive. It does not need some sort of *by*-phrase.

Grimshaw argues that complex event nominals retain the event and argument structure of the verb from which they are derived. Therefore, one would expect that complex event nominals obligatorily take arguments, just like verbs. Grimshaw argues that this is true: all of the arguments of a complex event nominal must be expressed in syntax. The following examples illustrate this:[6]

(12) a. the felling (+of trees)
b. the destroying (+of the city)

Both examples in (12) are only grammatical if the object is expressed. Note, however, that in neither case the subject is expressed. Grimshaw argues that the subject is an exception to the requirement that all arguments must be expressed. She claims that deverbal nouns are like passive verbs in that the subject is reduced.

There is evidence, however, that indicates that this is not the case. The most important evidence for this comes from control verbs. (11a) shows that masdars can be used as the complement of a control verb. The matrix verb in this example, *ḥāwaltu* 'I tried', controls the subject of the masdar, which must therefore be present. Under usual assumptions, this covert subject is PRO. In other words, (11a) shows that complex event masdars have a PRO subject if no overt subject is present.

This claim is also supported by evidence from binding facts. When a masdar has a reflexive argument, the data show that the masdar must be an interpretational domain with its own subject. Take, for example, (13):

(13) šadād-u yuwāṣilu -rtidā'-a-hu li malābis-i-hi wa
Shadad continued putting.on-ACC-his of clothes-GEN-his and
tahyi'at-a nafs-i-hi
preparing-ACC SELF-GEN-his
'Shadad continued putting on his clothes and preparing himself'

The finite verb in (13), *yuwāṣilu*, 'he continued' has two (conjoined) masdars as complement. Both these masdars are modified by a genitive pronominal element. the first is modified by a pronoun suffix (which is the equivalent of a genitive phrase), the second is modified by a genitive reflexive. Both the pronominal modifier of the first masdar and the reflexive modifier of the second refer to the subject of the finite verb, *šadād*. With the first masdar, *irtidā'* 'putting on (of clothes)', it expresses the subject, with the second, *tahyi'a* 'preparing', it expresses the object.

These facts indicate that the masdar is the binding domain for the anaphor. If it were not, it would be difficult to explain why in one case a pronoun is required to refer to *šadād* and in the other case a reflexive. After all, the pronoun and the reflexive are in the same syntactic position (the genitive complement of a masdar) and have the same referent (the subject of the finite verb.) If the masdar is the binding domain and if there is a PRO subject present in the masdar if no other subject is expressed, the facts are easily explained: the pronoun suffix *-hu*, being the subject of the masdar, is free in

[6] As Grimshaw argues, English gerunds only have a complex event reading, which means no specific method of disambiguation is needed in (12).

its domain, whereas the reflexive, being the object, is bound by the PRO subject.[7]

We see that Grimshaw's claim that complex event nominals do not have to express their subject is not correct for Arabic. Verbal masdars need to express the subject, and if the subject is not explicit, PRO is present. Fassi Fehri (1993), however, argues that masdars can easily be used without any arguments at all, even when they have an event reading. He gives the following example:

(14) lā 'urīdu -l-intiqād-a
not I.want the-criticising-ACC
'I do not want to criticise'

The verb *intaqada* 'to criticise' is usually construed with an object. In (14), however, its masdar *intiqād* is used without an object, even though it does have a complex event reading, being the complement of the control verb *'arāda* 'to want'. Because the matrix verb is a control verb, a PRO subject must be present.

Fassi Fehri argues that (14) shows that it is easier for masdars to leave the object unexpressed than it is for finite verbs. However, it is not problematic to use a finite verb instead of the masdar in (14):

(15) lā 'urīdu 'an 'antaqid-a
not I.want that I.criticise-SUBJ
'I do not want to criticise'

In other words, there is no evidence to suggest that (verbal) masdars behave very differently from finite verbs with respect to expressing their arguments. Although under certain pragmatic conditions it is possible for masdars to drop for example the object, this is also possible for finite verbs.

Another property of complex event nominals that Grimshaw argues for is that they only allow *the* as their determiner. Other determiners, such as *a, one, that/this* only occur with result nominals:

(16) a. they studied the/an/one/that assignment
b. they observed the/*an/*one/*that assignment of the problem
c. the/*that assignment of that problem too early in the course always causes problems

Although complex event nominals clearly do have a more restricted choice of determiners, it is not the case, at least in English, that only *the* is allowed. Given the appropriate context, a demonstrative can also occur:

(17) this constant harassing of passers-by must stop

In (17), *harassing* is a gerund, which makes it a complex event nominal. Yet, it takes the demonstrative *that*.

[7] The Reflexivity framework of Reinhart & Reuland (1993) accounts for these facts in a similar way. See Kremers (1997) for details.

Arabic verbal masdars show a similar restriction in determiners. Fassi Fehri claims that Arabic does not have the ability to use a demonstrative:

(18) ḥāwala zayd-un (*hāḏā) -l-iᶜtirāf-a
tried Zeid-NOM (*this) the-confessing-ACC
'Zeid tried to confess'
(Fassi Fehri 1993, p. 236)

However, just as in English, a demonstrative may be marginally correct in certain cases:

(19) waqaᶜa tadmīr-u -l-madīnat-i hāḏā
happened destroying-NOM the-city.F this.M
'this destroying of the city took place'

Apparently, the restriction against demonstratives is not absolute, neither in English nor in Arabic.

Similarly, Fassi Fehri claims that a verbal masdar cannot be indefinite:

(20) *yurīdu taqdīm-an li -ᶜtirāf-in
he.wants offering-ACC.INDEF of confession-GEN.INDEF
'he wants a confessing of a crime'

However, as we will see below in (24), it is possible for a verbal masdar to have an indefinite subject (or object) genitive, in which case the masdar itself is grammatically indefinite, because it inherits the definiteness of its genitive complement.

Obviously, the claim that Grimshaw and Fassi Fehri make that complex event nominals can only take the determiner *the/al-* is too strong, both for English and for Arabic. However, it still seems to be the case that complex event nominals do not allow the full range of determiners that other nouns allow. English does not allow the indefinite determiner *a*, and Arabic only allows an indefinite masdar if it is indefinite by inheritance from a genitive complement. In other words, Arabic verbal masdars behave like complex event nominals in this respect.

Another property of complex event nominals is that they do not allow pluralisation:

(21) a. the assignments were long
b. *the assignments of the problems took a long time

This is a very robust property, and it holds for Arabic as well:

(22) *tammat -iᶜtirāf-āt-u-hu bi -l-ḏanb-i
happened confessing-PL-NOM-his with the-crime-GEN
'his confessings of the crime took place'
(Fassi Fehri 1993, p. 236)

This inability to form plurals also follows from the fact that complex event nominals have the event structure of verbs. An event, the referent of a verb, is not compatible with number.[8]

Grimshaw furthermore argues that complex event nominals allow a subject possessor, but not if it is indefinite. This follows from the fact that a complex event nominal cannot be indefinite (as seen in (21) above) and that the definiteness of the possessor determines the definiteness of the entire phrase:

(23) a. ??a teacher's assignment of the problem
b. the assignment of the problem by a teacher

In Arabic, a construction like this is not unacceptable:

(24) ḍarb-u rağul-in mar'at-an ẖata'-un
hitting-NOM man-NOM.INDEF woman-ACC.INDEF mistake-NOM.INDEF
lit. 'a man's hitting a woman is a mistake'
'it is wrong for a man to hit a woman'

Note that the masdar in (24), *ḍarb* is indeed a complex event nominal: it assigns accusative to its object.

Grimshaw notes that the example in (23a) is not entirely ungrammatical, although it is rather degraded. The Arabic example is also not very acceptable. These facts indicate that indefinite subjects of complex event nominals are not disallowed, (in the same way that indefinite subjects of verbs are not disallowed), but there is some restriction against it, which I assume is pragmatic in nature.

Another property that Grimshaw notes for complex event nominals in English is that they do not occur as predicates or with equational *be*, while result nominals do:

(25) a. that was the/an assignment
b. *that was the/an assignment of the problem

This seems to be a property typical of English. Sentences such as (26) are grammatical in Arabic:[9]

(26) a. kāna hāḏā -ftitāḥ-a -l-ra'īs-i -l-mu'tamar-a
was this opening-ACC the-president-GEN the-conference-ACC
lit: 'this was the president's opening the conference'
b. kāna hāḏā qawl-a-hu -l-ḥaqq-a
was this saying-ACC-his the-truth-ACC
'here, he said the truth'
lit. 'this was his saying the truth'

Note that the objects of the masdars in (26), *al-mu'tamar* 'the conference' and

[8]There are languages that have some sort of "plural" marking on the verb, as discussed for example in Collins (2001), but they can probably be analysed as a specific kind of aspect markers, which would support the assumption made by some that Asp and Num are equivalent heads.

[9]Note that the verb *kāna* 'to be' assigns accusative case to its predicate. Hence the accusative case on *iftitāḥ* 'opening' and *qawl* 'saying'.

al-ḥaqq ‘the truth’ take accusative case, which means the masdars are indeed verbal masdars. Nonetheless, they are used across a copula, as the predicate of an equational sentence.

Grimshaw surmises that the reason that complex event nominals cannot be used as predicates in English lies in the fact that they cannot be indefinite. Predicates, she claims, must be indefinite. In general, that is a correct statement, but there is an exception. In equational sentences, the predicate[10] can easily be definite:

(27) a. this is my brother
b. that man is the man I saw yesterday robbing the bank

There is no reason why complex event nominals should not occur in this position. And indeed, as shown, they can, in Arabic.[11] Apparently, the fact that complex event nominals in English cannot, is typical for English. Since there is nothing that would predict that masdars *cannot* occur across a copula, Arabic does not show any unpredicted behaviour. I will leave the matter of why English does not allow these structures to future research.

Another property of complex event nominals is that they allow control into an infinitival purpose clause. As Grimshaw argues, the controller in such cases is presumably the event denoted by the nominal, rather than an implicit argument:

(28) a. the translation of the book (in order) to make it available to a wider readership
b. the examination of the patient in order to determine whether...

Unambiguous result nominals never allow control:

(29) a. *the translations of the book (in order) to make it available to a wider readership
b. *the exam in order to determine whether...

This is also a property that follows quite directly from the fact that complex event nominals retain the verb’s event and argument structure. We would therefore expect that verbal masdars also allow control into a purpose clause. The following example shows that this is indeed the case:

[10]Note that *predicate* may not be the correct term here. The structures are equational in nature, expressing identity, rather than a true predicative relation.

[11]They can in Dutch, too:

(i) de eerste stap is het installeren van Linux
the first step is the installing of Linux
‘the first step is to install Linux’

In (i), the noun *installeren* is formally an infinitive that takes the definite article *het*. Such nouns in Dutch have some of the properties of gerunds, specifically that they can be modified by an adverb. It is safe to assume, then, that they are complex event nominals. Yet, they can occur across a copula.

(30) qirā'at-u-ka -l-ğarīdat-a li fahm-i -l-siyāsat-i
reading-NOM-your the-paper-ACC for understanding-GEN the-politics-GEN
-l-dawliyyat-i
the-international-GEN
'your reading the paper in order to understand international politics'

In (30), the masdar *qirā'a* is modified by a purpose clause introduced by *li* '(in order) to'. The subject of this purpose clause,[12] is controlled by the subject of the matrix masdar *qirā'atu-ka* 'your reading'.

Verbal masdars and English complex event nominals show very similar behaviour. This leads to the conclusion that verbal masdars are complex event nominals in the sense of Grimshaw (1990). However, as explained above, masdars are also able to assign accusative case. This is a property that most deverbal nouns in English do not have. Only one category of deverbal nouns does: gerunds.

5.3 Abney's analysis of gerunds

Grimshaw (1990) discusses deverbal nouns, but she focuses on nouns such as *translation, assignment, destruction*, etc. These are nouns with an obvious verbal root, but they are lexicalised: their formation is not productive. Grimshaw does not devote much attention to another obvious class of deverbal nouns, the gerunds.

Gerunds have some properties that distinguish them from the deverbal nouns that Grimshaw discusses. First of all, their formation is productive. In principle, a gerund can be formed of any verb. Furthermore, they sometimes occur with typical verbal modifiers such as adverbs, which other deverbal nouns do not, and gerunds can assign accusative case to their objects, which is also something that other deverbal nouns cannot do.

Abney (1987) discusses three main types of gerund constructions, generally known as *Acc-ing, Poss-ing* and *Ing-of*. The following examples show these constructions:

(31) *Acc-ing*: John singing the Marseillaise
Poss-ing: John's singing the Marseillaise
Ing-of: John's singing of the Marseillaise

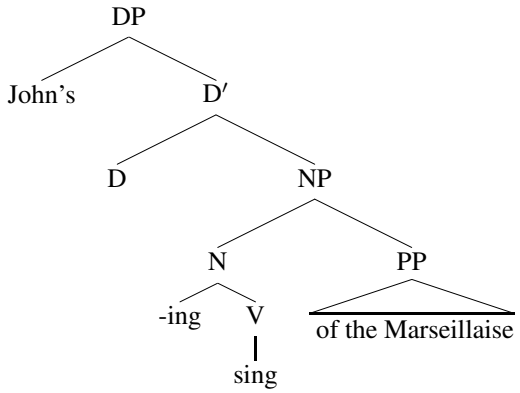
In Abney's analysis, there is a single affix *-ing* that can attach to a verbal element at different levels in the morphosyntactic structure. The affix *-ing* has the feature [+N],[13] making the verbal element a nominal one. So, for example, in the *Ing-of* construction, the affix attaches at the head level:[14]

[12]This subject is PRO, because the purpose clause itself is also a masdar: *fahm* 'understanding'.

[13]In fact, what *-ing* must do is change the feature matrix [+V,-N] to [-V,+N].

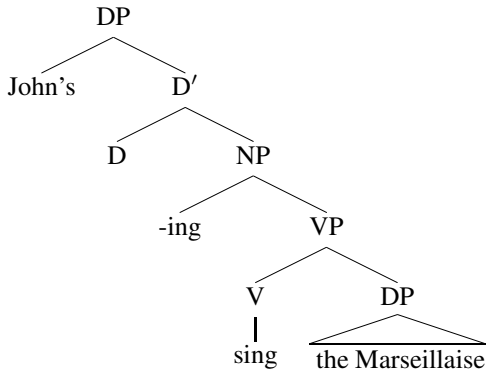
[14]This basically means that the affixation takes place in the lexicon rather than in syntax. However, Abney proposes a system in which the lexical tree and the syntactic tree are subject to the same principles, making the distinction arbitrary.

(32)



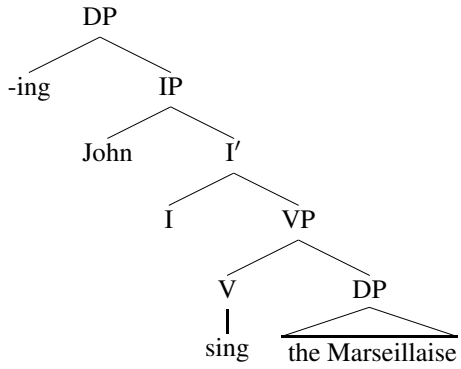
In the *Poss-ing* construction, the affix attaches at the level of VP, turning it into an NP:

(33)



In the *Acc-ing* construction, the verb projects to the level of IP. Only then is the *-ing* affix added, to produce a DP:

(34)



I will not go into the question of what it means for an affix to attach to a maximal projection, since I will propose a somewhat different formalisation later on.

The analysis proposed by Abney accounts for several of the properties of gerunds. For example, in some constructions gerunds can occur with adverbials:

(35) a. John's constantly criticising the customers is annoying
b. *John's constant criticising the customers is annoying

The gerund *criticising* in (35a) is modified by the adverbial *constantly*. It is not possible to use an adjective here, as can be seen from (35b). When the object is licensed with *of*, however, the facts are reversed:

(36) a. John's constant criticising of the customers is annoying
b. *John's constantly criticising of the customers is annoying

Abney's analysis offers a rather straightforward explanation for this. In the case of (36), the *-ing* affix is attached at the level of V, which means that there is no assigner of accusative case. It also means that adjectives must be used for modification, as it is a projection of N that is modified. In the case of (35), however, *-ing* is attached at the level of VP, which means that there is room for adverbial modification.

Verbal masdars in Arabic have several typical properties in common with English gerunds. First of all, the formation of (verbal) masdars is indeed a productive process in Arabic. For the majority of verbs, the formation is regular, for a smaller class, the masdar is irregular. With any verb, a masdar can be used with a verbal meaning, e.g. with the value of a subclause. Nominal masdars, on the other hand, are not productive. Their meanings are not predictable, and not every verbal masdar has a nominative use.

Secondly, masdars are able to assign accusative case:

(37) intiqād-u -l-raǧul-i -l-mašrūᶜ-a
criticising-NOM the-man-GEN the-project-ACC
'the man's criticising the project'

In (37), the masdar *intiqād* 'criticising' assigns accusative case to the object *al-mašrūᶜ* 'the project'. This fact argues for a gerund-like status of (verbal) masdars.

The third property that sets apart English gerunds is their ability to be modified by adverbs. Masdars have this property as well:

(38) intiqād-u -l-raǧul-i -l-mašrūᶜ-a bi šiddat-in
criticising-NOM the-man-GEN the-project-ACC with sharpness
'the man's sharply criticising the project'

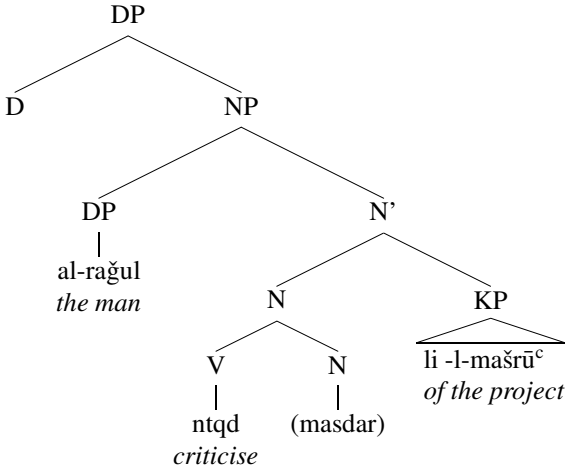
(38) is identical to (37), except that the adverbial phrase *bi šiddatin* 'sharply' is added. So it seems that Arabic masdars are on a par with English gerunds in this respect as well.

5.4 A minimalist view

Fassi Fehri (1993) analyses the structure of masdars along the lines of Abney (1987). So, for example, a masdar that uses *li* to license its object is analysed in the following way:

(39) a. intiqād-u -l-rağul-i li -l-mašrūᶜ-i
criticising-NOM the-man-GEN to the-project-GEN
'the man's criticising of the project'

b.

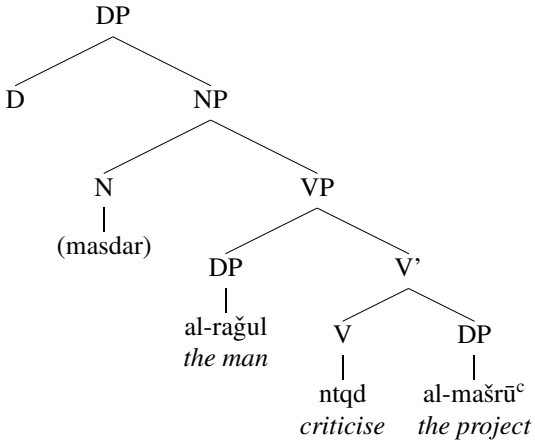


In (39b), the nominalisation affix (which I have indicated simply with *masdar*), is adjoined to V. As a result, the construction has an N-head, which projects.

Masdars with accusative objects have the nominalisation affix added to them at VP-level, forming an NP. Because the head of such a masdar is still V, it can assign accusative case to its object:

(40) a. intiqād-u -l-raǧul-i -l-mašrūᶜ-a
criticising-NOM the-man-GEN the-project-ACC
'the man's criticising the project'

b.



In (40b), the nominalisation affix is adjoined to VP. That means that V has already projected, assigning accusative to the object *al-mašrūᶜ* 'the project'.

It is not clear in the current theory what the status is of an affix that can attach to the structure at different levels, such as the *-ing* affix that Abney proposes, or the Arabic equivalent posited by Fassi Fehri. For this reason, I will use a somewhat different formalisation of the same intuition.

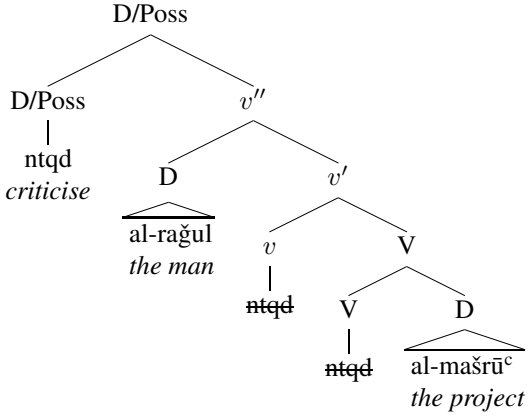
The formalisation that I use is based on the idea that the functional heads in the clause have their equivalents in the noun phrase. During the projection of V to CP, it is possible that a nominal lexical head is projected, rather than its verbal equivalent. So for example in the phrase *John's singing the Marseillaise*, rather than projecting T, it is Poss, the nominal equivalent of T, that takes v as complement.[15]

In the same way, the Arabic phrase in (39) is formed by projecting D/Poss where T would normally be projected:

[15]Note that in Arabic, as well as in Germanic and Romance languages, there is no regular, productive way to derive verbs from nouns, which suggests that a switch from a nominal to a verbal projection is not possible, and that language only allows a switch from a verbal to a nominal projection. However, there are languages, such as Nahuatl, that have quite elaborate systems of denominal (and de-adjectival) verbs. Whether or not such systems can be analysed as projections switching from a nominal to a verbal category is a question that must be answered by future research.

(41) a. intiqād-u -l-rağul-i -l-mašrū^{c}-a
criticising-NOM the-man-GEN the-project-ACC
'the man's criticising the project'

b.



The tree in (41b) shows that the root of the lexical entry, which is generated in V, has moved to D/Poss. In chapter 3, I argued that N does not move to D if D is filled with the definite determiner *al-*,[16] but I left open the question of whether N moves to D/Poss in a genitive construction, because this question could not be answered based on the available material. The present example shows that the lexical head does indeed move to D/Poss in this structure. This movement accounts for the observed word order of noun-subject-object.

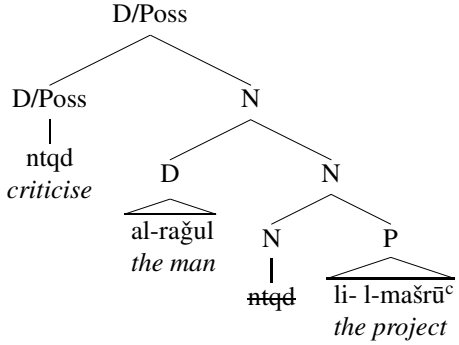
There are several elements of the masdar construction that need some attention. First, the availability of accusative case is easily explained. There is a small v present in the structure, which assigns accusative to the object.[17] We can similarly explain the absence of accusative in the masdar+*li* (equivalent to *ing-of*) by saying that v is absent in those structures:

[16] If D is indefinite, N does move to D.

[17] Note that it cannot be determined whether the accusative object moves to spec,v. I follow Mulders's (1997) analysis of mirrored specifiers. This explains why a moved object does not cause defective intervention effects when D/Poss probes for a match, but it also makes it impossible to determine whether the object actually moved, because there is no overt material between comp,VP and the inner spec,vP.

(42) a. intiqād-u -l-raǧul-i li -l-mašrūᶜ-i
criticising-NOM the-man-GEN of the-project-GEN
'the man's criticising of the project'

b.



The tree in (42) indicates that the switch of V to N takes place in the lexicon. This stands to reason, because the construction in (42) can be used with simplex event and result nominals, which, as we have seen in English, are typically formed in the lexicon rather than in syntax.[18]

The tree also indicates that the subject of the masdar is generated inside the NP, rather than as the specifier of some functional projection. This is in accordance with what I said in chapter 3: the external argument of a noun is the referential argument *R*, which means that the lexical arguments, including the subject, are internal, even though the subject is the external argument of a verb.

The analysis so far suggests that there is a strict parallel between English gerunds and Arabic verbal masdars. There are, however, several striking differences between them, which must be accounted for. First of all, in English, a gerund that assigns accusative can be modified by an adverbial, whereas a gerund that uses *of* cannot. This is different in Arabic, however. Adverbials are allowed with masdars that use *li*:

(43) 'aqlaqa-nī -ntiqād-u -l-raǧul-i bi -stimrār-in li
worried-me criticising-NOM the-man-GEN with duration-GEN of
l-mašrūᶜ-i
the-project-GEN
'the man's constantly criticising the project worried me'

The masdar in (43), *intiqād* 'criticism' has both a subject and an object, which means it allows a complex event reading. Even though the masdar does not assign accusative case, it still allows modification with an adverbial.

One might think that Arabic masdars can be modified by an adverbial no matter whether they are nominal or verbal, but this is not the case. A masdar that is used as a

[18]Note that complex event nominals can also occur in this structure. The difference between simplex event or result nominals and complex event nominal, as Grimshaw (1990) argues, is that the latter retain the event and argument structure of the verb because they do not have the normal *R* referential argument of verbs. Instead, they have an *Ev* argument, which resembles the *E* referential argument of verbs. Simplex event and result nominals, on the other hand, do have the *R* argument of non-event nominals.

simplex event or result nominal cannot be modified by an adverbial:

(44) qara'tu -ntiqād-a-hu (*bi šiddat-in)
I.read criticising-ACC-his with sharpness-GEN
'I read his (*sharply) criticism'

As expected, an adjective is possible in this structure:

(45) qara'tu -ntiqād-a-hu -l-šadīd-a
I.read criticism-ACC-his -the-sharp-ACC
'I read his sharp criticism'

In (44) and (45), the masdar *intiqād* 'criticism' has a result reading: the verb *qara'a* 'to read' cannot take a verbal complement. As a result noun, *intiqād* does not have event structure, and consequently cannot be modified by an adverbial.

The fact that the *masdar+li* construction can be modified by an adverbial if the masdar is verbal contrasts with the English facts. In English, an adverbial can only be used with a gerund that also assigns accusative case. The explanation for this difference may be found in the fact that Arabic adverbials are usually not directly derived from an adjective, but are formed with prepositional phrases. For example, in (43), the adverbial is *bi -stimrārin*, which literally means 'with duration'.

The semantic licensing of an adverbial is dependent on the event and argument structure of the lexical item. As Grimshaw (1990) shows, complex event nominals retain the event and argument structure of the verb, even if they are formed in the lexicon. As such, complex event nominals have the right semantics to allow an adverbial. This explains why Arabic prepositional adverbials are allowed with *masdar+li* constructions.

Yet, English adverbs are not allowed with *ing-of* nor with non-gerundive complex event nominals. We can account for this if we say that adverbs not only have to be semantically licensed, but have to be syntactically licensed as well. Although I will not go into the exact nature of this licensing process, it stands to reason that it can only take place in a verbal domain. Therefore an *ing-of*, which lacks such a verbal domain entirely, cannot license English adverbs. Prepositional adverbials, however, do not require syntactic licensing, which explains why they can occur in nominal contexts.

In this view, the contrast between English-type adverbs and prepositional adverbials would be similar to that between accusative arguments versus prepositional arguments. Both types of arguments require semantic licensing (i.e. they must be assigned a theta role), but only the former also require syntactic licensing, by being assigned structural case.

If this analysis is in the right direction, we predict that English will also allow adverbial prepositional phrases with *ing-of* constructions and not with non-gerundive nominals, which seems to be borne out by the facts:

(46) a. it was Paul's denying of the facts with such persistence that made Lucy so angry.
b. it was Paul's denial (*with such persistence) that made Lucy so angry

Furthermore, we also make the prediction that Arabic will not allow English-type adverbs with the *masdar+li* construction. Although Arabic does not have a specific adverb-forming suffix such as English *-ly*, it can occasionally form adverbs by using an adjective in the indefinite accusative:

(47) kāna ġālib-an yaqūdu šaḫṣiyyan hağamāt-in ğarī'at-an
he.was general-ACC he.lead personally attacks-ACC bold-ACC
'he would generally lead bold attacks personally'

The adjective *ġālib* 'predominant, general' is used in the accusative indefinite here with the meaning of the adverb *generally*. If the licensing of true adverbs requires a verbal complex, it is expected that the *masdars+li* construction does not allow such adverbs, which is indeed the case:

(48) qiyādat-u-hu (*ġālib-an) li hağamāt-in ğarī'at-in
leading-NOM-his (*generally) of attacks-GEN bold-GEN
'his generally leading of bold attacks'

Another question that presents itself if we analyse Arabic masdars along the lines of English gerunds, is why Arabic does not have a structure similar to *Acc-ing*:

(49) a. John singing the Marseillaise
b. *al-intiqād-u -l-raǧul-u/a -l-mašrū^{c}-a
the-criticising-NOM the-man-NOM/ACC the-project-ACC

Abney (1987) analyses the *Acc-ing* construction as involving adjunction of the nominalising affix *-ing* at the IP level, turning the IP into a DP. In other words, *Acc-ing* is on a par with *Poss-ing* and *ing-of* in being a DP.

One must wonder, however, if the structure really is a DP. Note, for example, that if the subject is PRO, *Acc-ing* cannot take a definite determiner, unlike *Ing-of*:

(50) a. (*the) PRO singing the Marseillaise
b. the singing of the Marseillaise

Furthermore, the *Acc-ing* construction can appear in positions where a normal DP cannot:

(51) Elaine's winking at Roddy was fruitless, *he being a confirmed bachelor*
(Reuland 1983, 101)

In (51), the *Acc-ing* construction appears as a sort of causal adjunct, which is not a position a DP can appear in. It is, however, a position in which a CP can appear.[19]

Reuland (1983) argues for an analysis of *Acc-ing* constructions as CPs containing an IP headed by *-ing*. In his analysis, *-ing* is a nominal element. One possibility seems to be that it is in fact a participial ending, and that the *Acc-ing* structure is a participial construction rather than a gerund construction.

[19]The case of the subject is obviously not accusative here, but there seems to be some variation, which I will not go into.

I will not go into the analysis of *Acc-ing* here. Suffice it to note that the answer to the question why Arabic lacks an equivalent of this structure must probably be sought not in the analysis of gerund or masdar constructions, but in the use and distribution of participles.

5.5 Participles

Another class of deverbal nominals is formed by participles. Although the topic is too broad to discuss it in-depth here, I will say a few things about them. There are two types of participles in Arabic: active participles and passive participles. In other words, Arabic makes a distinction that differs from Germanic and Romance languages, which usually make a distinction between present and past participles: *walk-ing* and *walked*, respectively.[20] The two participles of the verb *kataba* 'to write' are the following:

(52) a. kātib-un
writing ; having written
b. maktūb-un
(being) written

If we look for an analysis for participles, a good place to start is the impersonal passive participle, because it has a striking similarity to the adjective structure discussed in section 4.2.1 of chapter 4. Arabic allows impersonal passives with verbs that do not assign accusative to their objects but instead take a prepositional object:

(53) a. yuḥkamu ᶜalay-him bi -l-'iᶜdām-i
is.sentenced.3SG on-them with the-execution-GEN
'they have been sentenced to die'
lit. 'it is sentenced on them with execution'
b. al-riǧāl-u -l-maḥkūm-u ᶜalay-him bi
the-men-NOM the-sentenced.SG-NOM on-them with
-l-'iᶜdām-i
the-execution-GEN
'the men sentenced to die, the condemned'

[20] Past participles in Germanic or Romance often show a tendency towards taking a passive meaning. For example, in Dutch, substantivised past participles are interpreted as passives:

(i) de verdacht-e ; de overwonnen-e-n
the suspected-SUBST ; the conquered-SUBST-PL
'the suspect ; the conquered'

The past participles *verdacht* 'suspected' and *overwonnen* 'conquered' in (i) are substantivised with the ending *-e*. In their meaning, the participles are passive: *the one who is suspected* and *those who have been conquered*. Because of this tendency, a speaker of Germanic or Romance may be inclined to interpret the Arabic passive participle as past tense. This is not correct, however: both active and passive participles can in principle have a present and past meaning. They are not tied to a specific tense.

The verb *ḥakama* ‘to sentence’ takes two prepositional objects: the preposition *ᶜalā* ‘on’[21] expresses the person being sentenced, the preposition *bi* ‘with’ introduces the sentence being passed. In (53a), this verb is used in its passive form. Because the active voice does not have an accusative object to promote to subject, the verb takes a default third person masculine singular form and the two prepositional objects remain as they are. (53b) shows the participle of this verb. It modifies the noun *al-riğāl* ‘the men’. Note that the participle takes masculine singular form, even though the noun that is modified is plural. Impersonal passive participles always take this form, just like impersonal passive finite verbs.

If we compare this structure to the adjectives that have a DegP-internal subject, we see several striking similarities. To repeat an example from the previous chapter:

(54) ra’aytu -mra’-at-an ğamīl-an wağh-u-hā
I.saw woman-F-ACC.INDEF beautiful.M-ACC.INDEF face.M-NOM-her
lit. ‘I saw a woman beautiful her face’
‘I saw a woman with a beautiful face’

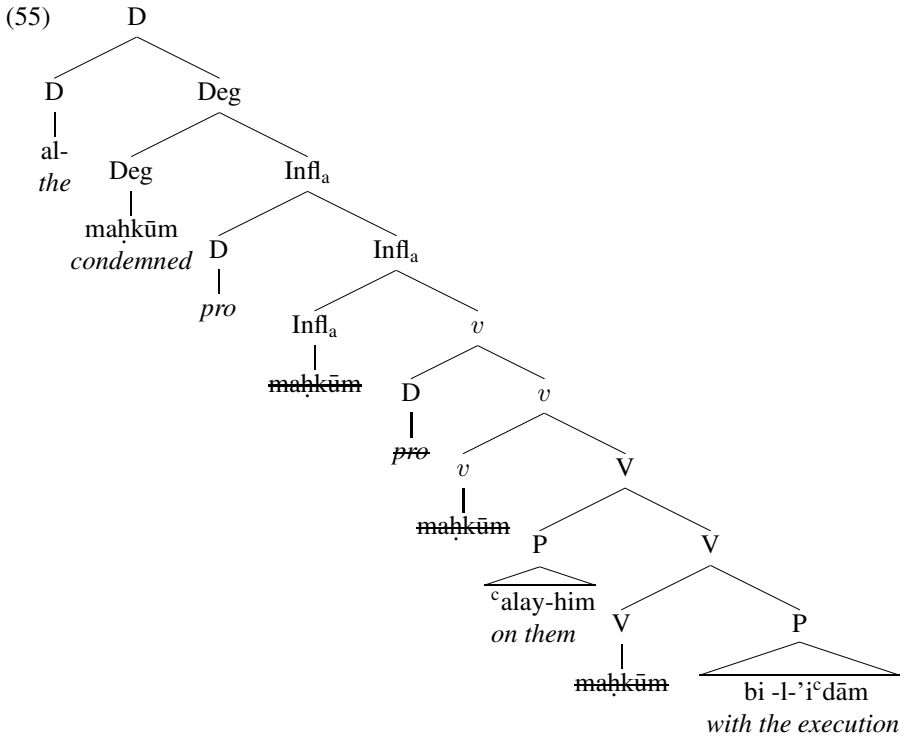
Recall that the adjective in these structures has a DegP-internal subject, and that it agrees with this subject in gender and number. It agrees with the head noun of the entire DP in case and definiteness. So in (54), the adjective *ğamīl* ‘beautiful’ is masculine singular, agreeing with the DegP-internal subject *wağhu-hā* ‘her face’, but it is also accusative and indefinite, agreeing with the head noun *imra’a* ‘woman’. Furthermore, the DegP contains a resumptive pronoun that refers to the head noun. In (54), this pronoun is the possessive suffix pronoun *-hā* ‘her’ on *wağh* ‘face’.

The impersonal passive participles have the same structure. The subject of the participle is an impersonal *pro*, just like the subject of impersonal passive verbs. The participle agrees with this subject in gender and number, which means it always takes masculine singular form. In case and definiteness, on the other hand, the participle agrees with the head noun. Furthermore, there is a resumptive pronoun in the participial construction, which is the complement of the preposition.

Taking the structure of the adjective phrase as a starting point, we can construe the following tree for the participial phrase in (53b):

[21] Like *’ilā* ‘to’, *ᶜalā* ‘on’ changes its form to *ᶜalay-* if it takes a suffix pronoun.

(55)



Because the participial phrase is a noun-phrase modifier, it must be a DP, with a DegP as its complement. The structure will also have an Infl$_a$, because a position is required to license the subject, even if the subject is only an impersonal *pro*. Furthermore, because the phrase has such obvious verbal properties, both VP and vP are present.[22]

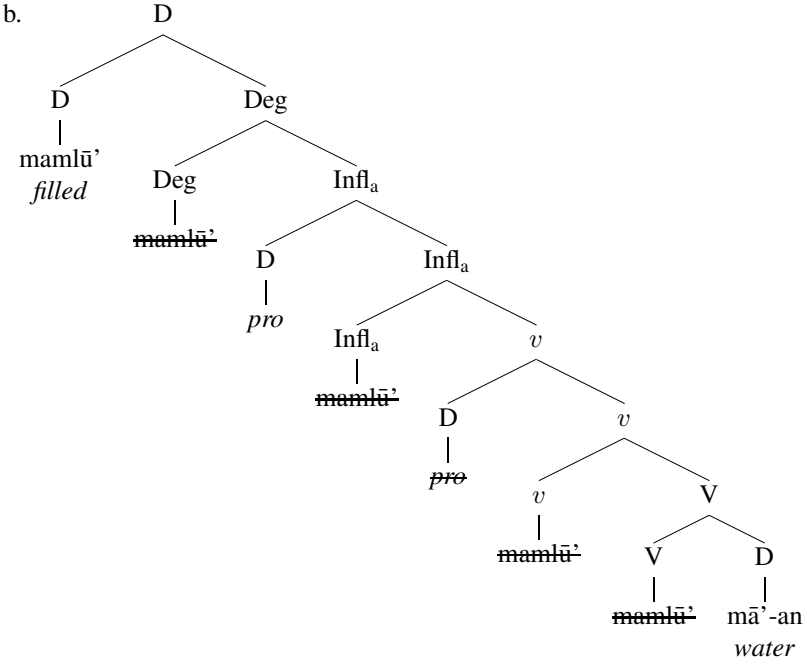
Already with this phrase a pattern suggests itself. I argued that deverbal nouns are built on a verbal projection, and that somewhere along the way a nominal functional head is projected. (55) suggests that we can analyse participles in much the same way: they also contain a verbal core, but at some point an adjectival head is projected instead of a verbal functional head.

Participles can in certain contexts assign accusative case. For example, the passive form of a double accusative verb can only promote one of its accusative objects to the subject position. The other object remains in object position and receives accusative case. The passive participle of such a verb behaves in the same way. Take the verb *mala'a* 'to fill', for example. This verb can take two accusatives, one expressing that which is filled, the other expressing that which it is filled with. In (56a), the passive participle of this verb is used to modify the noun *ḥawḍ* 'basin', which is the passivised subject:

[22]It is not clear where the passive morphology in this structure is located. If we follow Reinhart (2002), we can say it is on V, because passivisation involves changing the argument structure, which Reinhart argues to be a lexicon operation.

(56) a. (ra'ā) ḥawḍ-an mamlū'-an mā'-an
(he.saw) basin-ACC filled-ACC water-ACC
'(he saw) a basin filled with water'
(Cant. II p. 416)

b.

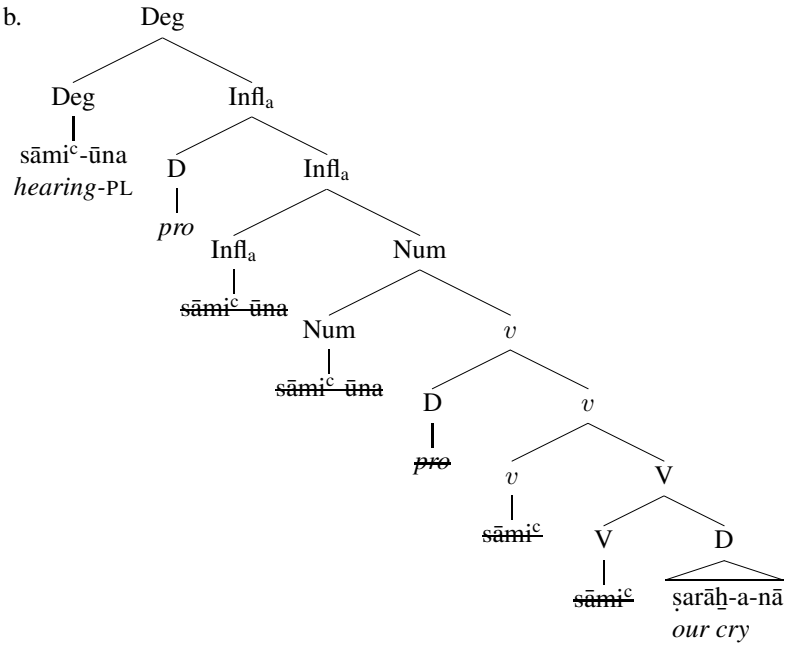


Note that the tree in (56b) only gives the participial DP, not the entire DP. The participial DP contains a *pro* subject, just like an adjective phrase. This subject refers back to the head noun *ḥawḍ* 'basin'. Because the internal object *mā'* 'water' receives accusative case, a small v head must be present.

Active participles can also assign accusative case. Such participles are sometimes used in Modern Standard Arabic as the main verb of a clause, replacing a finite verb:

(57) a. hal ’antum sāmi^{c}-ūna ṣarāḫ-a-nā?
Q you.PL hearing-PL.NOM cry-ACC-our?
‘do you hear our cry?’
(Cant. II, p. 412)

b.



The DegP is obviously embedded in a clausal functional structure, but in (57b) I only give the participial tree.[23] Just like in the tree of (56), there is a v present in the structure to assign accusative case. We also have a Num projection, which is the position where the plural suffix *-ūna* is base-generated.[24]

Masdars, as I explained above, have a verbal and a nominal use. The difference is basically a semantic one, but it is reflected in the syntactic properties of the two uses. A similar distinction can be made for participles. Verbal participles often have the value of a clause, either a main clause, a subclause or a relative clause. Such participles have verbal properties, e.g., in being able to assign accusative case, as we have seen. Nominal participles have the value of a noun or adjective, and like nominal masdars, they are not able to assign accusative case. Instead, they will use the genitive or the preposition *li* ‘to, for’ to license their objects:

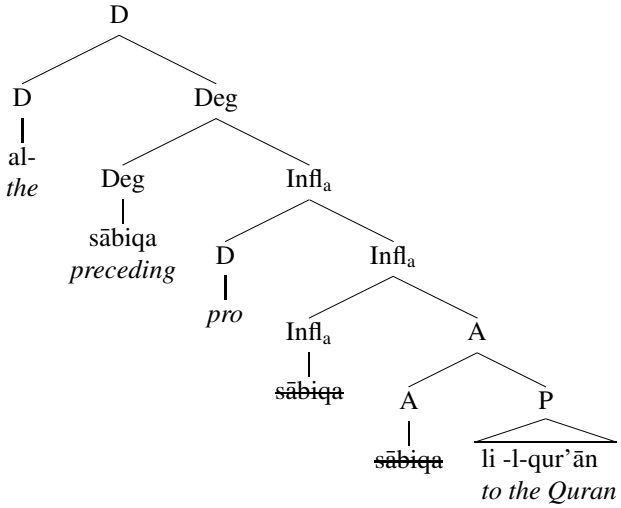
(58) al-kutub-u -l-muqaddasat-u -l-sābiqat-u li -l-qur’ān-i
the-books-NOM the-holy-NOM the-preceding-NOM to the-Quran-GEN
‘the holy books that preceded the Quran’

[23] Note that I have put a *pro* as the DegP-internal subject. It is also possible that the clausal subject *’antum* ‘you.PL’ is generated in the DegP and then moves on to spec,TP. I will not discuss this option here.

[24] In fact, a Num projection must also be present in the structures discussed above, because every participle is marked for number.

In (58), the participle *sābiqa* 'preceding' uses the preposition *li* 'to, for' to license its object *al-qur'ān* 'the Quran'. Taking the structure of masdars as an example, we can say that the structure of the participial phrase is derived by making the switch from a verbal to an adjectival category in the lexicon, much like what happens in the case of the nominal masdar in (42):

(59)



This quick look at participles suggests that they can be analysed along the same lines as masdars: both masdars and participles start out as verbs, but at some point in the derivation, possibly already in the lexicon, a head is projected that is of a different category, nominal or adjectival, respectively. Once this switch has taken place, it is not possible to switch back when some higher head is projected.

This analysis suggests that there are indeed equivalent heads in the three domains: verbal functional categories have their equivalents in nominal and adjectival functional categories. The category switch can only take place between equivalent heads. That is, it is possible to project Poss instead of T, or N instead of V, but it is not possible to project N instead of T, or D instead of v, for example.

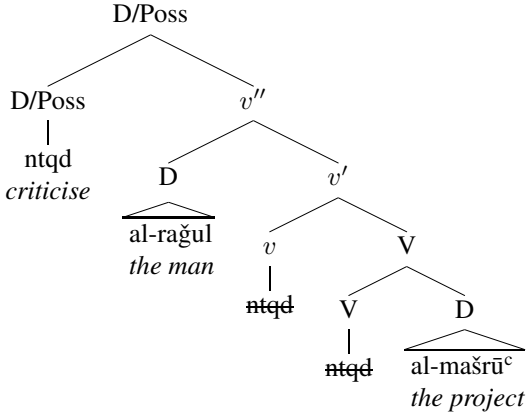
5.6 Linearisation

To a large extent, linearisation of masdars is rather straightforward. However, we run into a problem when we try to linearise prepositional objects. When we examine these structures, it turns out that there is one other operation that can take place during linearisation: postponement.

Before we discuss prepositional objects, let us look at two straightforward cases. First, a masdar that is nominalised at the level of D/Poss:

(60) a. intiqād-u -l-raǧul-i -l-mašrū^{c}-a
criticising-NOM the-man-GEN the-project-ACC
'the man's criticising the project'

b.

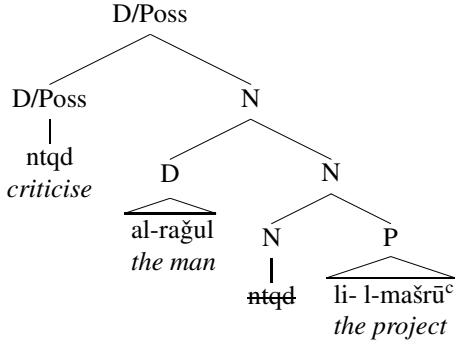


The example in (60) shows how RLin deals with the subject argument of the verb, *al-raǧul* 'the man'. This is a selected argument, as it is s-selected by v. Because of this, principle S will force it to be linearised first in the node v''. This means that if there are two s-selected arguments in a lexical projection, the higher argument will always be spelled out before the lower argument. This is a desirable result, because arguments that are hierarchically higher in the structure generally appear first in the linear string.

A masdar that is nominalised at the level of N is dealt with in the same way:

(61) a. intiqād-u -l-raǧul-i li -l-mašrū^{c}-i
criticising-NOM the-man-GEN of the-project-GEN
'the man's criticising of the project'

b.



We again see the effect of the linearisation principle S, which forces a selected specifier to be linearised first. In (61b), the subject argument *al-raǧul* 'the man' will be linearised before its sister node, which contains the object. In this way, principle S is responsible for the order subject-object that we observe in masdars.[25]

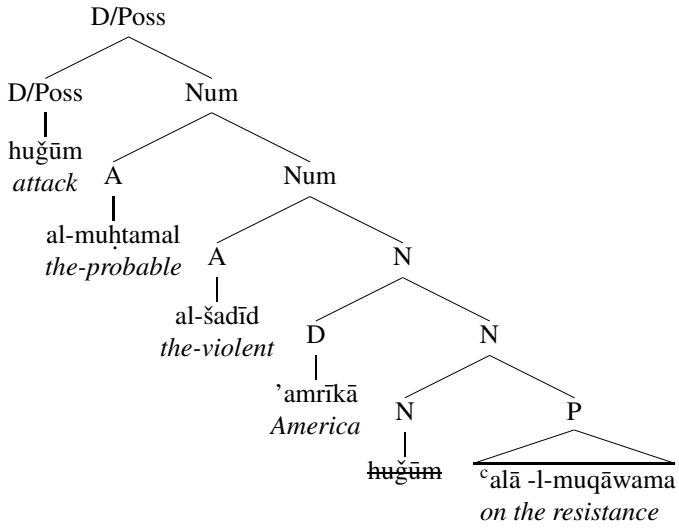
[25] In a way, a structure such as in (61) is very similar to a double object construction: a lexical head having

Now let us turn to the problem that occurs with prepositional complements. The procedure outlined in chapter 3 cannot linearise a prepositional complement properly if the noun phrase also contains an adjective. In order to see this, let us return to an example that was already discussed in chapter 3:

(62) huğūm-u ʾamrīkā -l-šadīd-u -l-muḥtamal-u ᶜalā
attack-NOM America.GEN the-violent-NOM the-probable-NOM on
-l-muqāwamat-i
the-resistance-GEN
'the probable violent attack on the resistance by the US'

Given the analysis presented in this chapter, the tree structure of this phrase would be the following:

(63)



With the parameter settings of $H > S$ and adjunct-first, we would derive the following order:

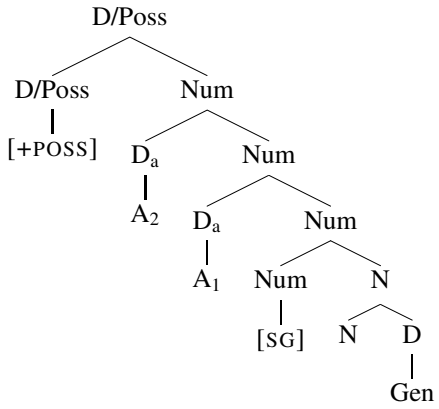
(64) huğūm ʾamrīkā ᶜalā -l-muqāwama al-šadīd al-muḥtamal
attack America on the-resistance the-violent the-probable

This is not the correct linearisation, however: the adjectives follow the prepositional complement, whereas they should in fact precede it, as demonstrated in (62).

To begin with, let me demonstrate the source of the problem. RLin can account for the mirror-image orders we see in adjective placement because of its recursive nature. Take, for example, the following tree:

two internal complements. In general, a double object construction has the property that both objects are linearised in the same way with respect to the verb (i.e., they both appear either before or after the verb, not one before and the other after) and of the two objects, the one that is hierarchically the highest appears first in the linear string. (See, for example, Marantz 1993). Principle S is what is responsible for this.

(65)



With the linearisation parameters set as mentioned, the order that is obtained is the following:

(66) D/Poss - N - Gen - A_1 - A_2

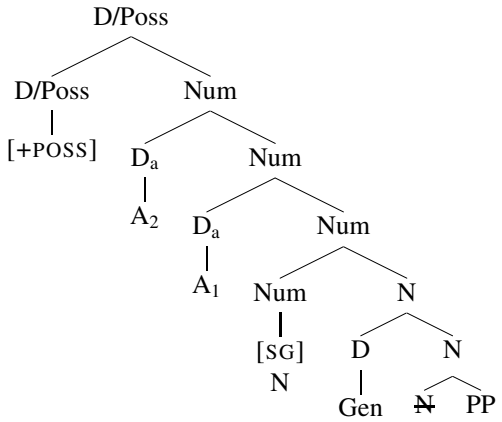
The tree in (65) contains two adjectives. They are ordered in a mirror image order because RLin processes the tree recursively, and because with the current parameter settings it deals with non-selected specifiers last.

If we visualise the procedure in the tree, we can say that RLin goes down from the top of the tree to its lowest branches, spelling out various components according to the settings of the parameters, and then goes back up again to spell out those components that were skipped on the way down. The adjectives are linearised in a mirror image order because they are linearised "on the way up."[26] On the way up, RLin first encounters A_1 and then A_2.

However, following the procedure in this way entails that *everything* inside the NP is linearised before the adjectives. After all, the NP is the lowest element in the tree, and as such it will be spelled out in its entirety when RLin has reached it. In (65), this is not problematic, because everything in the NP *needs* to be spelled out before the adjectives. If, however, the NP contains a prepositional complement, things are different. Take the following tree:

[26]The phrases *on the way down* and *on the way up* should not be taken too literally. They just describe the visual effect when one traces the linearisation in a tree.

(67)



(67) will be spelled out with the prepositional phrase before the adjectives:

(68) D/Poss - N - Gen - P - A_1 - A_2

This order ensues because everything inside the NP, that is, the genitive noun, the trace of N *and* the prepositional phrase are spelled out before RLin gets to the adjectives. However, the observed order is the one with the prepositional phrase *after* the adjectives:

(69) D/Poss - N - Gen - A_1 - A_2 - P

This order is not always the observed order, however. There are indeed cases in which the order predicted by (68) occurs:

(70) a. wa lam yumkin-hu ḏālika ’illā baᶜda riyāḍat-in li
and not enabled-him that except after training-GEN to
-l-nafs-i ṭawīlat-in
the-mind-GEN long-GEN
‘he could only do it after long mental training’

b. hiya tağribat-u ’aṣdiqā’-a l-ī kaṯīr-īna
it experience-NOM friends-GEN.INDEF of-me many-PL
‘it is an experience of many of my friends’
(SASG p. 401)

In (70a), the adjective *ṭawīl* ‘long’ modifies the noun *riyāḍa* ‘training’. We know that it does not modify *nafs* ‘soul’ because this noun is definite, whereas the adjective is indefinite. Yet the adjective is separated from the noun it modifies by the prepositional phrase *li -l-nafs* ‘of the mind’. The same thing occurs in (70b), where the adjective *kaṯīrīna* ‘many’ is separated by the PP *l-ī* ‘of me’ from the noun it modifies, *’aṣdiqā’* ‘friends’.

It is not only the preposition *li* ‘of, to’ that can occur between a noun and its modifying adjective. Other prepositions occur there as well:

(71) a. ḥiqbat-un min al-zaman-i ṭawīlat-un
period-NOM from the-time-GEN long-GEN
'a long stretch of time'
b. ǧundiyy-un min silāḥ-i -l-mašāt-i basīṭ-un
soldier-NOM from service-GEN the-walkers-GEN simple'
'a simple foot soldier' ('a simple soldier from the infantry')
c. kānat tuḥissu bi taṣallub-in fī šuᶜūr-i-hā
she.was she.feels with hardening-GEN in feelings-GEN-her
ġarīb-in
strange-GEN
'she was feeling a strange hardness in her emotions'
(SASG p. 402)

Note that this ordering of PP-Adj is not possible with definite nouns:

(72) a. *al-ǧundiyy-u min silāḥ-i -l-mašāt-i -l-basīṭ-u
the-soldier-NOM from service-GEN the-walkers-GEN the-simple-NOM
b. al-ǧundiyy-u -l-basīṭ-u min silāḥ-i -l-mašāt-i
the-soldier-NOM the-simple-NOM from service-GEN the-walkers-GEN
'the simple foot soldier'

Furthermore, indefinite nouns do not require the PP to precede the adjective:

(73) ǧundiyy-un basīṭ-un min silāḥ-i -l-mašāt-i
soldier-NOM simple-NOM from service-GEN the-walkers-GEN
'a simple foot soldier'

In other words, the placement of the PP after the adjectives is the preferred option, and in the case of definite nouns, it is the only option. Nonetheless, I will argue that this option is obtained through an additional operation.

The way to derive the correct linear ordering from the tree in (67) is to say that the spell-out of the PP is postponed. That is, RLin searches the tree in the normal way, but when it encounters the PP, it holds off on spelling it out. Instead, it continues processing the rest of the tree. This causes the adjectives to be spelled out. When the entire phase has been completed, that is, when RLin has completed the spell out of both subnodes of the top D/Poss node, it will take the PP again and spell it out. As a result, the PP will appear after the adjectives, which have already been spelled out at this point.

One might say that this postponement of spell-out is a rather *ad hoc* solution, but in fact it is not. We need such a mechanism in order to explain the phenomenon of *heavy-NP shift*:

(74) they sent *t* to Mary that book that only got good reviews in the New York Times

An object noun phrase that contains a large amount of phonological material can be displaced to the right edge of the sentence. In (74), the noun phrase *that book that only*

got good reviews in the New York Times should appear in the position of the trace *t*, but instead, it has been displaced across the indirect object PP *to Mary*. No definitive analysis of this phenomenon has been offered yet, but it is generally believed that it is a peripheral operation that is not part of the core grammar but operates at PF.

In the current proposal, in which linearisation is described as a process that operates at PF, we immediately see how such heavy-NP shift must operate: RLin processes the tree in the normal manner, but when it encounters the heavy NP, its spell-out is postponed until the containing phase, in this case the clause,[27] has been spelled out.

What I will argue is that in the case of the PP complement in Arabic noun phrases, such postponement is strongly preferred over direct spell-out. The reason why it takes place is not phonological heaviness, because the spell-out of any PP can be postponed. What the exact reason is will have to be determined, but it probably has to do with the fact that if the prepositional phrase would be spelled out directly, the adjectives would then follow the complement of the PP, which is a noun phrase itself. It would therefore not be immediately clear that the adjectives belong not to the prepositional complement but to the head noun of the noun phrase. Perhaps we can say that spell-out of the PP is postponed in order to avoid this confusion.

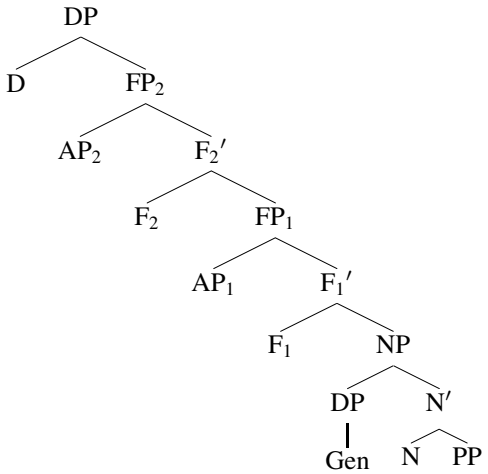
In fact, this reason is very similar to the reason why spell-out of the heavy NP in (74) is postponed: if it were not, it would be difficult for the listener to attach the indirect object PP *to Mary* to the verb, because the distance between the two has become too large. In other words, we can conclude that under certain conditions, it is possible to postpone the spell-out of a constituent.[28] It is important to note that such postponement can only lead to (apparent) rightward movement to the edge of the phrase: the constituent that is postponed will be spelled out when the rest of the containing phase has been completed.

It should be noted that an antisymmetric approach to the Arabic noun phrase also runs into the problem of the PP ending up in the wrong place. For example, in Cinque's (1994) approach, the structure of a phrase with two adjectives, a genitive complement and a PP complement would be (75):

[27] In fact the vP.

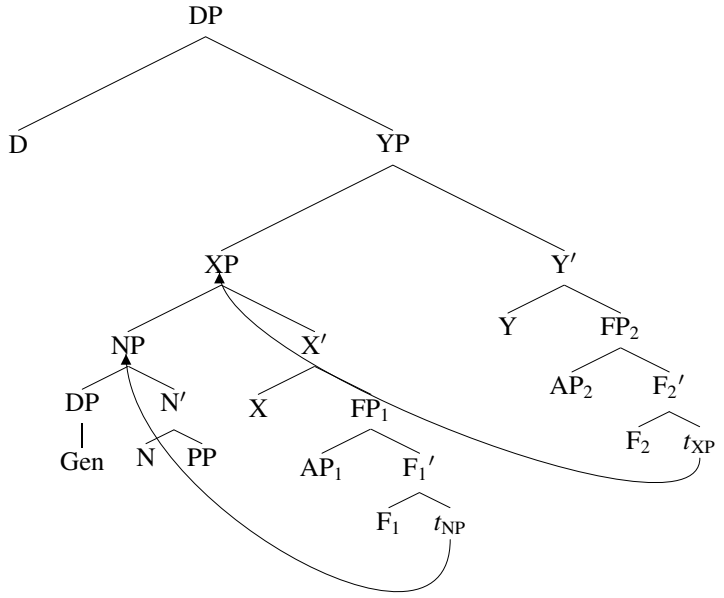
[28] The exact conditions under which this process can take place will have to be specified, but I will leave that question to future research.

(75)



In Cinque's account, the mirror image order is derived by moving the NP to a position above FP_1, say to XP, and then moving XP to a position above FP_2, as shown in (76):

(76)

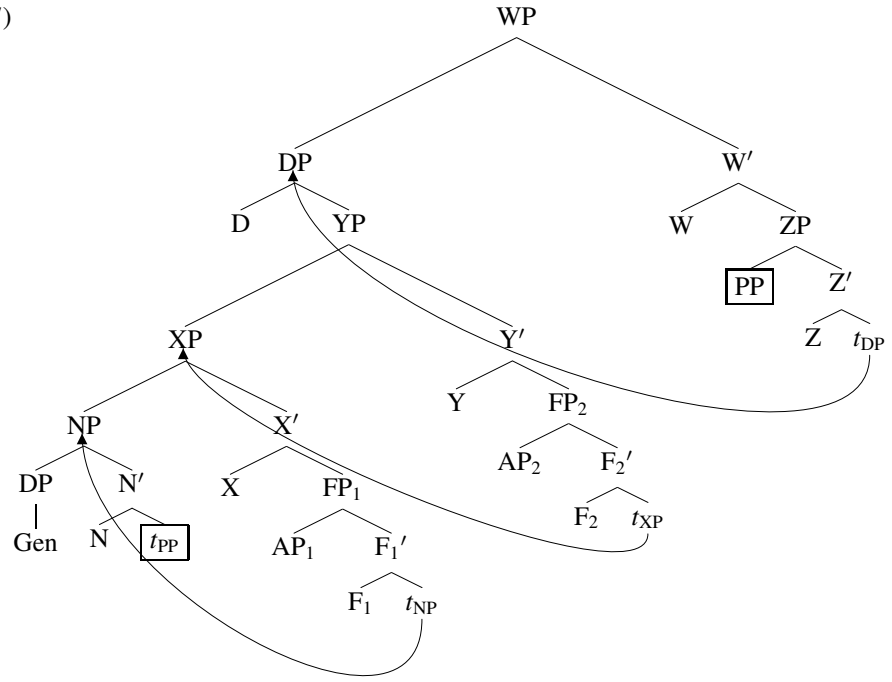


As one can see, this derivation creates the same problem: with the movement of NP to spec,XP, the prepositional phrase moves as well. Therefore, the PP will end up before the adjectives.[29] More recent antisymmetric approaches would offer a solution in which the PP moves out of the NP first, to a position above DP. After that, the

[29]Kayne's (1994) approach to adjective placement as discussed in chapter 4 has the same problem. It also derives the mirror-image order by moving the NP, containing the PP.

movements can take place in the way that either Cinque or Kayne describes. Lastly, the whole DP would move to a position above the PP. With Cinque's approach, the resulting structure would be (77):

(77)



The analysis is problematic for obvious reasons: neither the movement of the PP nor the final movement of the DP are properly motivated: it is not clear what triggers them, it is not clear what sort of projections these landing sites WP and ZP are. The movements described here are very similar to the sort of movements that would be required to account for heavy-NP shift in an antisymmetric approach, but unlike RLin, such an approach does not offer any insight into the process. What, for example, is the restriction that makes sure that the movement of the PP to spec,ZP is always accompanied by later movement of the DP to spec,WP? Why would it not be possible to only move the PP to spec,ZP, creating an order of PP-D-N-A_1-A_2? In RLin, it is obvious that we need the notion of postponement of spell-out as a way to describe the effects, and we see clearly what the effect of this operation is: the postponed constituent will be spelled out at the right edge of the phrase.[30]

[30]Furthermore, the way RLin deals with rightward movement suggests strongly that it can only apply to phases. That is, it is not possible to postpone the spell-out of a NumP if linearisation has already started on the DP that contains it.

5.7 Summary

Deverbal nouns in Arabic, so-called *masdars*, have two different uses: they can have a verbal use and a nominal use. The verbal use of masdars is very similar to the structure of English gerunds: the masdars are complex event nominals, showing many of the properties that Grimshaw (1990) illustrates for this class of nouns. Furthermore, verbal masdars can assign accusative case and to a certain extent allow adverbials.

Verbal masdars are generated on a verbal base: the lexical projection is a verb that is productively transformed into a noun: at a certain point in the derivation, a nominal functional head is projected instead of its verbal counterpart. For example, instead of projecting T, which would yield a clause, a Poss is projected, yielding a DP.

A preliminary investigation of participles shows that they behave much like masdars. They too show a distinction between a verbal and a nominal/adjectival use: participles can be used with the value of a clause (either main clause, subclause or relative clause) but they can also have the value of an adjective or a noun. This similarity suggests that we must analyse them as verbal projections where at some point a nominal or adjectival functional head is projected.

In the linearisation of masdars, it turns out that prepositional complements cause a problem. Because they are located inside the NP, it would be expected that they are linearised before any adjectives, which are linearised last. Although there are structures in which this order occurs, the predominant order has the adjectives preceding the prepositional phrase. We can account for this if we assume that RLin has the ability to postpone spell-out of a phase until the containing phase has been spelled out. This possibility of postponement is needed independently to account for the phenomenon of *heavy NP-shift*. Postponement of a phase results in apparent rightward movement of that phase to the edge of the containing phase.

6

Conclusions

The main purpose of this thesis has been to develop an analysis of the Arabic noun phrase in minimalist terms. It turns out that such an analysis is indeed feasible. The Arabic noun phrase projects the heads D, Poss and Num. It also contains the features CASE and GENDER, but they do not project an independent head. Instead, CASE is present on N and is inherited by the other heads, allowing it to be expressed morphologically on any of them. GENDER only projects a syncretic head with Num, or it is present as a feature on N that is not expressed morphologically.

When the noun has a complement, the head Poss will have the feature [+POSS] and a set of unvalued φ-features. These unvalued φ-features will be valued during the derivation in an Agree relation with the complement of the head noun. In the same Agree relation, the complement is assigned genitive case. This means that genitive is a structural case, just like nominative and accusative.

When Poss has the feature [+POSS], it projects a hybrid head which also has a DEF feature. Because there is already a DEF feature present in the projection, no D head is projected, resulting in a noun that does not carry a morpheme marking definiteness. Furthermore, the DEF feature of this D/Poss head is unvalued, which means it has to be valued during the derivation. As just remarked, the unvalued φ-features on Poss trigger an Agree relation with the complement of the head noun. This complement is a DP, and as a consequence has a valued DEF feature. So the DEF feature of D/Poss is valued in the same Agree relation that also values the φ-features of D/Poss and the CASE feature of the complement noun.

The adjective in Arabic shows certain phenomena that make clear that the adjective does not agree directly in φ-features with the noun it modifies. Instead, it agrees with a subject that is internal to the adjective phrase. In most cases, this subject is *pro*, but Arabic allows a construction in which this subject is overt, containing a resumptive pronoun that refers back to the head noun. This means that the adjective phrase has a clause-like structure: there is a head that mediates in subject-adjective agreement, which in Arabic can also assign nominative case.

The adjective can also license an internal argument, by assigning it genitive case. This internal argument can be a complement in the case of participles, but also the argument that Higginbotham (1985) calls the *attribute*.

Adjectives in Arabic show definiteness agreement with the head noun: if the noun is marked for definiteness, so is the adjective. This adjectival D head is added to guarantee that the resumptive pronoun, which functions as a variable at the semantic level, has a binder. Because derivations are built up phase by phase, the variable must be bound locally. For this reason, the D head is added to the adjective phrase.

This analysis of adjective phrases is very similar to the analysis of relative clauses. They too contain a resumptive pronoun that refers to the head noun, and they too are headed by an element that serves as a binder for the resumptive pronoun. This binder is not a D but a C head, which shows agreement in φ-features, definiteness and case with the head noun. The binder, C in the relative clause and D in the adjective phrase, adopts the φ-features from the D head of the head noun, and passes these on to the resumptive pronoun. If the resumptive pronoun in the adjective phrase is *pro*, (i.e. if it is in subject position) the adjective will adopt the features of the head noun through agreement with this *pro*.

Deverbal nouns in Arabic (*masdars*) have two different uses: they can have a verbal use, in which case they resemble English gerunds, and they can have a nominal use, in which case they behave like other, non-event nouns. In their verbal use, masdars can be used in positions where a verbal clause is required. As such, verbal masdars are typical complex event nominals in the sense of Grimshaw (1990). Furthermore, masdars can assign accusative case and they can license adverbials, which means that they have a verbal base. This shows that verbal masdars can have a V head. Because all masdars have the distribution and many of the properties of nouns (i.e., case marking, definite determiner), we must conclude that at some point in the derivation, a nominal functional head is projected where in a clause a verbal functional head would be. For example, instead of projecting a T head that takes v as complement, a Poss head is projected. Such a switch in category can only happen between equivalent heads. Poss can take the place of T, but not the place of v.

Participles in Arabic show behaviour that is very similar to the behaviour of masdars. They too have a distinct verbal use, in which they can assign accusative case and replace verbs, and a nominal use, in which they behave as nouns or adjectives. Just like verbal masdars have nominal properties, verbal participles have adjectival properties: they have the same distribution, and share many of the properties of adjectives (i.e., case marking, definiteness agreement). Therefore, we can analyse participles in the same way as masdars: the verbal participles have a V head, and at some point in the projection an adjectival functional head is projected instead of a verbal one.

As explained in chapter 2, the bare phrase structure approach to tree structures developed by Chomsky (1995) does not account for linearisation. There is no explanation of the way in which the tree structure is mapped onto a linear structure. Nonetheless, this is an essential part of the PF derivation. Because syntactic structures can be modelled as binary trees, linearising a tree amounts to searching the tree for terminal elements which can be spelled out. The most straightforward way of searching a

binary tree, and the only way that is compatible with common syntactic assumptions, is a so-called depth-first search. This search method works recursively: it takes the tree that is to be linearised; if this tree is simplex (i.e., if it only consists of a single terminal element) this element is spelled out; if the tree is compound, it is split up into its two subnodes, and the search procedure applies itself to both subnodes. Because of its recursive nature, I have dubbed this procedure *recursive linearisation*, or RLin.

If the (sub)tree that is being linearised is compound, RLin has to decide to which of the two subnodes it will apply itself first. This decision is made solely on the basis of the information that is available to RLin, that is, on the basis of the features of the two subnodes. By comparing the categorial features of the node and its subnodes, RLin determines which of the two subnodes projects. Then, two principles apply: the Selected principle, or S, which states that a selected head must be linearised first, and the Head principle, or H, which states that a head (i.e. a non-compound node) must be linearised first.

When the node that is being linearised consists of a selected specifier and its X′ sister, only S applies. As a result, selected specifiers are always linearised first, and consequently will appear before their heads in the linear string. When, on the other hand, the node being linearised consists of a head and its complement, both S and H apply: H requires that the head be linearised first, S requires that the complement be linearised first. The two principles therefore clash. In order to resolve this clash, the two principles are ordered. This ordering is parameterized: it varies from language to language. If the order is $H > S$, the head will be linearised first, resulting a Head-Complement order, and if the order is $S > H$, the complement will be linearised first, resulting in a Complement-Head order.

When the node being linearised consists of a non-selected specifier (i.e., an adjunct) and its X′ sister, neither of the principles S or H apply. This means that adjuncts can be linearised either first or second in their node. The choice of first or second linearisation is parameterized with the adjunct parameter.

When the adjunct parameter is set to adjunct-first, adjuncts are linearised in hierarchical order: the highest adjuncts will be spelled out first, and consequently will appear in the linear string first. When the adjunct parameter is set to adjunct-second, however, adjuncts are linearised after all other elements in the phase have been linearised. More importantly, they are linearised in reverse order: the hierarchically lowest adjunct will appear first in the linear string. That is, the setting of adjunct-second yields mirror-image orders in adjuncts.

RLin also allows us to describe *heavy-NP shift* purely in PF terms. The linearisation of a phase can be postponed until the linearisation of the containing phase has been completed. As a result, the postponed phase will appear to have been dislocated to the right edge of the containing phase. This type of dislocation has always been regarded as a PF phenomenon that does not take place in core syntax. With RLin, we see how this works. Postponement does not apply to heavy NPs alone. It can also apply to other phases, such as prepositional phrases in the Arabic noun phrase.

RLin obviates the need for movement to account for word order. We do not need to posit unmotivated movement operations in order to account for something like a Complement-Head order. Note, however, that it is not the case that movement does

not have an effect on word order. Movement certainly exists, and when it takes place it puts elements in a different structural position, which *may* have the effect that they end up in a different place in the linear string. The important thing to realise is that not all word order effects are caused by movement.

Summarising, RLin provides us with a very simple and straightforward account of linearisation. We do not need to resort to movement operations to account for word order. There are two parameters, one ordering parameter for the principles S and H, and one adjunct parameter. The values of both of these parameters are directly observable in the primary linguistic data, making them the ideal sort of parameter: their settings are directly accessible to the language-learning child.

Appendices

A

Arabic noun-phrase syntax and morphology

This appendix describes the syntax and morphology of the Arabic noun phrase in traditional terms. It is added as a reference. Although I try to provide enough information in the text for the reader to examine the data himself, some readers may prefer additional information, which is given here.

A.1 The simple noun

A.1.1 Noun morphology

Arabic is a Semitic language, and it shares with other languages of this family its striking morphological structure. The root of most lexical items consists of three consonants. In order to derive words from such a root, the consonants are supplemented with vowels and consonantal affixes. (See, e.g. Holes 1995). The inserted vowels and consonants are together called a 'pattern'. Many of these patterns are regular and give a specific meaning. For example, the vowel pattern -ā-i- (where the hyphens stand for root consonants) indicates the active participle. Thus, from the root KTB, which carries the meaning of 'to write', one derives *kātib* 'writing', or 'writer'.[1]

Most nouns in Arabic are derived from such a three-consonantal root. There are many patterns for deriving nouns, some of which have a regular meaning, but many

[1] Usually, patterns are indicated by putting them in a consonantal root. Most often, the root F^{c}L is used for this, or less frequently, KTB. I will use KTB for convenience. Thus, one speaks of the pattern KāTiB rather than of -ā-i-. Using a consonantal root that has itself a meaning to name patterns may lead to some confusion. It should be noted that the patterns do not represent words themselves. KiTB is just a pattern, there is no such word as *kitb*. It is possible, of course, that a certain pattern is also found with the root KTB, but this is not a necessity. When I write a pattern, I will do this with the capitals KTB. If I write an actual word that has this root, I will give it in lower case and, in the text, in italic.

are just 'nominal patterns' without a specific meaning beyond that. For example, the pattern KaTB is found in such nouns as *bayt* 'house', *nafs* 'soul', *bard* 'cold' and many more. Other nouns have a pattern KiTB, e.g., *fiᶜl* 'action', *kiḏb* 'lie', or maKTaB, such as *maktab* 'desk; office', *maṣnaᶜ* 'factory', *madraǧ* 'course, route'.

Definiteness, case and gender

When a pattern is applied to a root, a word results that can be further supplemented with a number of affixes. For nouns, these affixes indicate definiteness, case, and sometimes number. The first of these is the definite article, the prefix *al-*. Secondly, every noun takes a case suffix. Arabic has three cases: nominative, genitive and accusative.[2] For most nouns, the case endings are as follows: *-u* for nominative, *-i* for genitive and *-a* for accusative. Lastly, most nouns take an indefinite suffix if they do not have a definite article. This suffix takes the form *-n*. Thus, a noun like *bayt* 'house' can take the following forms:

	indefinite	definite
NOM	bayt-u-n	al-bayt-u
GEN	bayt-i-n	al-bayt-i
ACC	bayt-a-n	al-bayt-a

Table A.1: Triptotic nouns (*bayt* 'house')

The nouns in table A.1 are called *triptotic* nouns, because they have three different case endings. There is a group of nouns, called diptotes, that have only two case endings when they are indefinite: they have *-a* for the genitive, as well as for the accusative. Furthermore, they do not take the indefinite suffix. When definite, they have the normal three case endings. For example:

	indefinite	definite
NOM	ṣaḥrā'-u	al-ṣaḥrā'-u
GEN	ṣaḥrā'-a	al-ṣaḥrā'-i
ACC	ṣaḥrā'-a	al-ṣaḥrā'-a

Table A.2: Diptotic nouns (*ṣaḥrā'* 'desert')

There are two genders in Arabic: masculine and feminine. Like in European languages, the gender of many nouns is arbitrary. For nouns referring to persons, however, gender is generally derived from the sex of the referent. Masculine nouns have no specific ending. Feminine nouns often take a special feminine ending *-at*.[3] The case endings and the indefinite suffix follow after this ending. No noun ending in *-at* uses the diptotic endings in table A.2.

[2] These cases are purely syntactic. That is, the case of a noun is determined on syntactic grounds only: the noun's place in the structure. Unlike Latin, for instance, a case ending does not convey some meaning of its own. E.g., in Latin, it is possible to say *Romam eo* 'I go to Rome', where the accusative on *Roma* expresses direction. One cannot use case in Arabic in this manner.

[3] When quoting a word, the *-t* of the feminine ending is usually dropped. I will follow this practice.

Apart from these nouns, which all effectively end in a consonant, there are also nouns ending in *-ā*. These take neither case endings nor indefinite suffix. Most of them, though not all, are feminine.

Plural formation

There are two different ways to form the plural of a noun. Most nouns have a so-called broken plural, in which a different pattern is applied to the root. Note that only the vowels change. If the singular noun has a pattern that comprises also a consonantal affix, this affix is retained in the plural:

(1)	a.	bayt	buyūt
		house	houses
	b.	maṣna^c	maṣāni^c
		factory	factories

In (1a), the plural of the noun *bayt* 'house' is formed by applying the pattern KuTūB to the consonantal root. In (1b), the word *maṣna^c* 'factory' (root ṢN^c), which has the pattern maKTaB, is pluralised by applying the pattern maKāTiB to it. As one can see, the consonantal prefix *m-* is retained. Note that singular and plural patterns often go together: maKāTiB is a typical plural pattern for the singular pattern maK-TaB.

Broken plurals take the same case endings as singular nouns and also take the indefinite marker, as indicated in table A.1. Many plurals, however, are diptotes, and will therefore take the endings exemplified in table A.2.

The other way to form the plural of a noun is to use a plural suffix. There are two such suffixes: *-ūna*, the so-called masculine sound plural, and *-āt*, the feminine sound plural. The masculine sound plural is used mostly with nouns denoting male persons. The feminine sound plural is used for nouns denoting female persons, but also for many inanimate nouns, abstract and concrete. Note that in the latter case, the nouns do not have to be feminine.

The masculine sound plural is marked for case: nominative is *-ūna*, genitive and accusative is *-īna*. Thus, the following forms exist:

	indefinite	definite
NOM	mudarris-ūna	al-mudarris-ūna
GEN	mudarris-īna	al-mudarris-īna
ACC	mudarris-īna	al-mudarris-īna

Table A.3: Masculine sound plural (*mudarris* '(male) teacher')

Note that the masculine plural does not take the indefiniteness ending *-n*.

The feminine sound plural *-āt* takes alternative case endings: *-u* for nominative and *-i* for genitive and accusative, and can also take the indefinite suffix. This gives the following forms:

	indefinite	definite
NOM	mudarris-āt-u-n	al-mudarris-āt-u
GEN	mudarris-āt-i-n	al-mudarris-āt-i
ACC	mudarris-āt-i-n	al-mudarris-āt-i

Table A.4: Feminine sound plural (*mudarrisa* '(female) teacher')

It should be noted that the different plural formations cannot be applied freely. Each noun has a specific plural, either broken or sound.[4]

Dual formation

Apart from a plural, Arabic also has a dual. This is formed with the suffix *-āni*. It applies to all nouns, whether masculine or feminine, animate or inanimate. It is marked for case: *-ayni* is used for genitive and accusative:

	indefinite	definite
NOM	mudarris-āni	al-mudarris-āni
GEN	mudarris-ayni	al-mudarris-ayni
ACC	mudarris-ayni	al-mudarris-ayni

Table A.5: Dual (*mudarris* '(male) teacher')

Like the masculine plural ending, the dual ending does not take the indefiniteness suffix *-n*.

A.1.2 Modification

Arabic nouns can be modified in different ways. An adjective can be added, a demonstrative, etc. In this section, I discuss how these modifications take place.

Demonstratives

Arabic has two demonstratives: one for proximity (*hāḏā)* and one for distance (*ḏālika*). They have singular and plural forms, and the singular furthermore has masculine and feminine forms. The proximity demonstrative *hāḏā* also has dual forms for masculine and feminine. The demonstratives are not declined for case, except for the dual forms of *hāḏā*. The full paradigm is given in the following table:

[4]Sometimes a noun can have more than one plural form. E.g., a noun like *ṭarīqa* 'manner, mode' has two broken plurals: *ṭarā'iq* and *ṭuruq*.

		near		far	
		m	f	m	f
sg		hāḏā	hāḏihi	ḏālika	tilka
du	(nom)	hāḏāni	hātāni		
	(gen/acc)	hāḏayni	hātayni		
pl		hā'ulā'i		'ulā'ika	

Table A.6: Demonstratives

A demonstrative is placed before the noun it modifies. Note that a noun modified by a demonstrative also takes the definite article, as demonstrated in (2a).[5] If, like in (2b), the definite article is omitted, the phrase will be interpreted as a clause in which the demonstrative is subject and the noun predicate:

(2) a. hāḏā -l-bayt-u
this the-house-NOM
'this house'
b. hāḏā bayt-u-n
this house-NOM-INDEF
'this is a house'

Adjectives

Morphologically, adjectives cannot be distinguished from nouns. They take the same case endings as nouns, and also have either the definite article or the indefinite suffix. Most adjectives are triptotic (table A.1), but some types of adjectives are diptotes (table A.2). The only difference with nouns is that adjectives always have a masculine and a feminine form.[6]

Adjectives are strictly postnominal. They agree with the noun in gender, case and number.[7] Apart from that, adjectives also agree with the noun in definiteness. That is, if the modified noun has a definite article, the adjective also takes the article. And in the same way, if the modified noun is indefinite, so is the adjective:

(3) a. bayt-u-n ğamīl-u-n
house-NOM-INDEF beautiful-NOM-INDEF
'a beautiful house'
b. al-bayt-u -l-ğamīl-u
the-house-NOM the-beautiful-NOM
'the beautiful house'

This definiteness agreement may seem strange at first, but similar facts are found

[5]Note that the article *al-* loses its initial *a-* when the preceding word ends in a vowel.

[6]Except for adjectives that only apply to women, such as *ḥāmil* 'pregnant'. They do not take the feminine ending. Note, by the way, that nouns that refer to people usually have both a masculine and a feminine form. In other words, adjectives and nouns also pattern alike in this respect.

[7]There is one quirk: inanimate plurals always trigger feminine singular agreement; not only on adjectives, but also on demonstratives and verbs.

in other languages as well. Kester (1996), for example, argues that the definiteness feature on adjectives plays an important role in the adjectival inflection in Germanic languages.

Note that when the noun is definite and the adjective indefinite, the phrase is interpreted as a sentence:[8]

(4) al-bayt-u ğamīl-u-n
the-house-NOM beautiful-NOM.INDEF
'the house is beautiful'

A.2 The possessive structure

As explained in section A.1.1, Arabic nouns have three cases, one of which is the genitive. The genitive is used after all prepositions and also, as expected, to express possession. However, it is not the case that one can simply add a noun in the genitive to another noun to express possession. In order to do this, a special construction is used.

A.2.1 The *construct state*

A noun can only be modified by one, and no more than one, genitive noun. The two nouns, head noun and modifier, form a rigid structure: their order is fixed, nothing can intervene, no extraction is possible. The order is always *head noun – modifier*. Furthermore, the head noun takes a special form, the so-called *construct state*. In the construct state, the noun is not marked for definiteness: it takes neither definite article nor indefinite suffix. The modifier is a full noun phrase, and must be marked for definiteness. Naturally, the modifier takes genitive case:

(5) a. bayt-u rağul-i-n
house-NOM man-GEN-INDEF
'a man's house'
b. bayt-u -l-rağul-i
house-NOM the-man-GEN
'the man's house'

In the genitive constructions in (5), the head noun is *bayt*. As one can see, it only takes a case ending. It does not have a definite article, nor an indefinite suffix. The genitive modifier, *rağul* 'man', does take a definiteness marker.

Dual and plural nouns can, of course, also be modified by a genitive phrase. The construct state of broken and feminine sound plurals is identical to singular nouns: no definite article and no indefinite suffix. Masculine sound plurals and duals, however, do not have the normal indefinite suffix. In the construct state, they lose *-na* and *-ni*, respectively. Thus, the construct state forms are the following:

[8] Generally, predicative adjectives are indefinite. This means that it is not possible to say 'a house is beautiful'. The structure required for this would be (3a), but this means something else. Note, however, that

	dual	plural
NOM	mudarris-ā	mudarris-ū
GEN	mudarris-ay	mudarris-ī
ACC	mudarris-ay	mudarris-ī

Table A.7: Construct state of dual and plural

For example:

(6) mudarris-ū -l-ğāmiᶜat-i
teacher-M.PL.NOM the-university-GEN
'the teachers of the university'

Pronouns

The possessor that modifies a noun can be a noun itself, as in the previous section, but it can also be a pronoun. In Arabic, this genitive pronoun takes the form of a suffix on the noun. It follows after the case ending. The full paradigm is given in table A.8.

		singular	dual	plural
3	M	bayt-u-hu	bayt-u-humā	bayt-u-hum
	F	bayt-u-hā		bayt-u-hunna
2	M	bayt-u-ka	bayt-u-kumā	bayt-u-kum
	F	bayt-u-ki		bayt-u-kunna
1		bayt-ī		bayt-u-nā

Table A.8: Possessive pronouns (*bayt* 'house')

Note that the vowel *-u-* in the third person forms changes to *-i-* when the vowel preceding the pronominal suffix is *-i-*. Thus *bayt-u-hu* 'his house' in the nominative, but *bayt-i-hi* in the genitive. Similarly, *bayt-u-hum* 'their house' in the nominative and *bayt-i-him* in the genitive. Before the first person singular suffix *-ī*, the case endings disappear.

The possessive suffixes are used after all nouns, whether singular or plural. Note that the noun is in the construct state, which means that for dual and sound masculine plural, the forms in table A.7 are used.[9]

a phrase like 'a house is beautiful' has a generic meaning. Arabic uses a definite noun to express this.

[9]Though some nouns undergo small phonological modifications, which I will not go into here.

A.2.2 Modification

A possessive structure can be modified by demonstratives and adjectives. How this is done is discussed in this section.

Demonstratives

If a demonstrative is added to the genitive modifier, it is placed before this noun, as would be expected:

(7) a. bayt-u hāḏā -l-raǧul-i
house-NOM this.M the-man-GEN
'this man's house'
b. lawn-u hāḏihi -l-tuffāḥ-at-i
colour-NOM this.F the-apple-F-GEN
'the colour of this apple'

The head noun, however, cannot be as easily modified by a demonstrative. If one would put a demonstrative before the noun, the phrase would be interpreted as a sentence:

(8) hāḏā bayt-u -l-raǧul-i
this house-NOM the-man-GEN
'this is the man's house'

In traditional grammar, this is attributed to the fact that the head noun does not have a definite article. This means that "the connection between the demonstrative and the noun cannot be made". In order to obtain the correct effect, the demonstrative must be placed after the modifying noun:

(9) a. bayt-u -l-raǧul-i hāḏā
house-NOM the-man-GEN this.M
'this house of the man's'
b. nāfiḏ-at-u -l-bayt-i hāḏihi
window-F-NOM the-house-GEN this.F
'this window of the house'

Adjectives

Adding an adjective to the genitive modifier of a possessive construction is straightforward: the adjective simply follows the genitive noun and agrees with it in the usual manner:

(10) ṣāḥib-u -l-sayyār-at-i -l-ḥamrā'-i
owner-NOM the-car-F-GEN the-red.F-GEN
'the owner of the red car'

When the head noun of a possessive phrase is to be modified by an adjective, there is a problem. Nothing can intervene between the head noun and the genitive modifier.

Therefore, the adjective is placed after the modifier:

(11) sayyār-at-u -l-raǧul-i -l-ǧadīd-at-u
car-F-NOM the-man-GEN the-new-F-NOM
'the man's new car'

As (11) demonstrates, the adjective agrees with the head noun in number, gender and case. However, the adjective is not in the construct state itself. Instead, it takes a definite article. Actually, the adjective takes its definiteness feature from the genitive noun, even if it modifies the head noun:

(12) bayt-u ṭalabat-i-n kabīr-u-n
house-NOM students-GEN-INDEF large-NOM-INDEF
'a large student house'

In (12), the head noun *bayt* 'house' is modified by the adjective *kabīr* 'large'. Since the head noun is in construct state and therefore not marked for definiteness, the adjective obtains its definiteness feature from the genitive modifier.

B

Breadth-first vs. depth-first

As stated in chapter 2, there are two main strategies for searching (binary) tree structures. They are called breadth-first and depth-first. (See, e.g., Winston 1992 and Bratko 2001.) The basic idea of both methods is the same. While the tree is being searched, two lists must be kept: one search list `S` of the nodes still to be searched, and one result list `R` that contains the elements found. At the beginning of the search, the search list `S` contains only the root node. At the end of the search, a flat list `R` (i.e., a list without hierarchical structure) results that gives the linear ordering of the terminal elements in the tree. The procedure consists of the following steps:

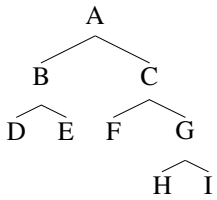
(1) search list of nodes `S`:
   a. take the first element E of `S`
   b. if E is a terminal, add E to result list `R`
   c. if E is non-terminal, add subnodes A and B to `S`
   d. if `S` is not empty, search the new list `S`
   e. if `S` is empty, terminate and spell out `R`

This procedure has three choice points. As stated, the procedure keeps two lists, indicated with `S` and `R`. Each time we add an element to one of these lists, we need to make a choice where to add it: to the front or to the back. This gives us two choice points, one on `S` and one on `R`. Furthermore, when adding subnodes A and B to `S`, we need to choose whether we add them in the order A-B or in the order B-A. This choice is important, because the procedure always takes the first element of the new list `S` to search. Therefore, the order in which the elements are added has a direct effect on the order in which the terminals are found and end up in `R`.

Let us start with `S`. The choice whether to add new elements to the front or back of this list gives us two different search procedures, which are traditionally known as depth-first and breadth-first. In a depth-first search, elements are added to the front of `S`, in a breadth-first search elements are added to the back of `S`.

To see how this works, let us take the following tree:

(2)



The following is a representation of the search procedure in a breadth-first strategy:

(3) S: [A] → [B C] → [C D E] → [D E F G] → [D E F G] → [E F G] → [F G] → [H I] → [I] → []
R: [D E F H I]

S in (3) gives the contents of the search list in each step of the search. At each step, the first element is taken out and examined. If it is found to be a terminal, it is added to R. I have indicated this by underlining the element. If the first element is not a terminal, it is taken out and its subnodes are added to the back of the list. So in the first step, A is taken out and its subnodes B and C are added to the list. In the second step, B is taken out and examined. Since it is not a terminal, its subnodes D and E are added to the back of S, etc.

R in (3) gives the final contents of the result list R, after the entire tree has been searched. One can see why this mechanism is called breadth-first: the tree is searched as it were from left to right.

The breadth-first search strategy can easily be implemented in a computer language such as Prolog:

```
% breadth-first linearisation making use
% of search and result lists

terminal(X) :- atom(X).

order(A,B,A,B).

linearise(Tree) :-
        linearise([Tree], [], ResultList),
        atom_chars(Result, ResultList),
        write(Result).

linearise([],R,R).

linearise([N|S], R, NewNewR) :-
        terminal(N),
        append(R, [N], NewR),
        linearise(S, NewR, NewNewR).
```

```
linearise([[_,A,B]|S], R, NewR) :-
        order(A, B, First, Second),
        append(S, [First,Second], NewS),
        linearise(NewS, R, NewR).
```

We can represent the tree in (2) with the following fact:

(4) `tree([a, [b, d, e], [c, f, [g, h, i]]]).`

The order in which two elements are added to the result list `R` is determined by the `order` predicate. In the listing above, the subnodes A and B will be added in the order A-B. If we use this order, the program will linearise the tree in (4) as `defhi`:

```
?- tree(T), linearise(T).
defhi
```

Breadth-first is obtained when the subnodes of a node are added to the back of the list. As noted earlier, we can also add the subnodes to the front of the search list. If we do this, we get a depth-first search:

(5) `S:` [A] → [B C] → [D E C] → [E C] → [C] → [F G] → [G] → [H I] → [I] → []
`R:` [D E F H I]

This method is called depth-first, because the tree is searched from top to bottom: each node is first searched completely before its sister node is searched. The essential difference between the two search methods is that in a depth-first strategy, each node is dealt with as it is encountered, whereas in breadth-first a node that is encountered is stored and only processed later. This is reflected in the position in which each node is stored in the search list: in breadth-first, it is stored at the end, which means it will only get processed after the rest of the list. In the depth-first strategy, it is stored at the beginning of the search list, which means it will be processed in the immediately following step.

The depth-first search strategy can be implemented with the program above if we make one small change. We have to change the line

```
append(S, [First,Second], NewS)
```

to become

```
append([First,Second], S, NewS)
```

which will add the elements A and B to the front of `S` rather than to the end.

The two methods do not yield a different word order in this case: both give [D E F H I] as the value for `R`. We could also choose to add elements to the back

of `R` rather than to the front,[1] which would yield the reverse order, but this still gives us identical orderings for both search methods. This changes, however, if we choose to add the subnodes of a given node to `S` in reverse order. We can implement this by changing the definition of the `order` predicate. Instead of

```
order(A,B,A,B).
```

we must use

```
order(A,B,B,A).
```

If we do this, breadth-first gives us the following:

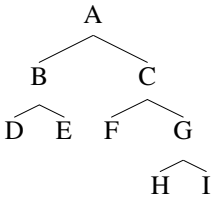
(6) `S:` [A] → [C B] → [B G F] → [G F E D] → [F E D I H] → [E D I H] → [D I H] → [I H] → [H] → []
`R:` [F E D I H]

Depth-first, on the other hand, will result in the following:

(7) `S:` [A] → [C B] → [G F B] → [I H F B] → [H F B] → [F B] → [B] → [E D] → [E D] → [D] → []
`R:` [I H F E D]

When we consider these two outputs, we see that breadth-first searches a tree in a way that can give undesirable results. To see this, look at the search in (6). It produces an `R` list of [F E D I H]. When we compare this to the original tree structure, we see some anomalies:

(8)



In (8), the nodes F and G are sisters. In the `R` produced by the breadth-first search, however, F and the terminal nodes dominated by G are *not* adjacent. The terminal nodes E and D intervene. This effect is caused by the fact that the nodes D and E are at the same level as the nodes F and G. The breadth-first search searches a tree in layers, where each row of nodes constitutes a layer. So in (8), [A] is a layer, [B C] is a layer, [D E F G] is a layer and [H I] is a layer.

Because breadth-first searches in layers, it has the ability to separate sister nodes in the way described. This is highly undesirable, because the standard assumption in syntactic research is that if there is a (sub)segment [X Y Z] in the linear string, this segment has a hierarchical structure of either [[X Y] Z] or of [X [Y Z]]. Breadth-first

[1] This can be done by changing the line `append(R,[N],NewR)` to `append([N],R,NewR).`

search, however, has the ability to create linear strings with a structure of [[X] Y [Z]], where [X ... Y] forms a discontinuous element. The difference between the two search strategies is that with depth-first all the terminals in a single branch will end up adjacent in the linear string, whereas in breadth-first, all the terminals of a single layer will end up adjacent in the linear string. The latter is incompatible with standard syntactic assumptions.

For this reason, breadth-first search is not the appropriate strategy for RLin to use. It must use a depth-first search. The interesting thing about depth-first is that it can be implemented without the search list `S`. Depth-first takes an element of the list `S`, splits it up into its subnodes and puts those elements back to the front of `S`. In the next step, however, it will take the first element of `S` again, which is one of the elements that it just added to `S`.

Therefore, we can implement depth-first search without the search list. Instead of taking an element from `S` and adding its subnodes back to `S`, we can simply apply the search procedure to the tree directly, by splitting up the root of the tree and then searching both subnodes in order. The simplified procedure is the following:

(9) search tree T:
- a. take subnodes A and B of the root of T
- b. if A is a terminal, add A to `R`, otherwise search tree A
- c. if B is a terminal, add B to `R`, otherwise search tree B

After the tree T has been split in its subnodes in step (a), the top node can be discarded, because all the information in it that is relevant is in A and B. We do not need to keep a list of nodes still to be processed, because each node is immediately processed as it is encountered. This differs from the breadth-first strategy, in which a node that is encountered is not immediately processed, and therefore has to be stored.

With the elimination of breadth-first, we have also eliminated the choice point that existed on `S`. We can also eliminate the choice point on the result list `R`, because it does not yield any additional orderings. To see this, consider what happens if we process the subnodes of a node K(A,B) in the order A-B, and add terminal elements to the end of `R`:

(10) `S`: [A] → [B C] → [D E C] → [E C] → [C] → [F G] → [G] → [H I] → [I] → []
`R`: [D E F H I]

The `R` that results is [D E F H I]. If we were to add the terminals to the beginning of `R` instead of to the end, we would get the reverse: [I H F E D]. Now, let us look at a depth-first search that processes a node K(A,B) in the order B-A:

(11) `S`: [A] → [C B] → [G F B] → [I H F B] → [H F B] → [F B] → [B] → [E D] → [E D] → [D] → []
`R`: [I H F E D]

As one can see, the resulting `R` is [I H F E D], which is exactly the reverse of the `R` that resulted from the depth-first search in (10). In other words, we can eliminate

the choice point on R without losing anything.

In Prolog, this simplified depth-first search can be straightforwardly implemented in the following program:[2]

```
% depth-first linearisation without
% making use of search and result lists

terminal(X) :- atom(X).

order(A,B,A,B).

rlin([]).

rlin(Node) :-
        terminal(Node),
        write(node).

rlin([_,A,B]) :-
        order(A, B, First, Second),
        rlin(First),
        rlin(Second).
```

This program is simpler than the previous version because there is no need to keep a search list S. It is also not necessary to keep a result list R. Instead, the terminal elements found are outputted immediately.

The program shows that the algorithm is in essence very simple. The argument given to rlin is the tree that is to be linearised. If this tree is empty, nothing has to be done and rlin terminates. If the tree consists of a single terminal element, the element is spelled out. If the tree is compound, the subnodes A and B are taken, their order is determined and they are linearised by applying rlin to them.

In a full implementation of RLin, we would expand the order predicate, so that it will determine the order in which A and B have to be linearised in the manner described in chapter 3, and we would also have to make provisions for the possibility of postponing linearisation of a phase. But in essence, RLin is described with this simple algorithm.

[2]Neither this program nor the one above is really the most efficient implementation possible of the search procedures. There are ways in which the code can be optimised. But they are the most straightforward implementations, which is why I show them here.

References

Abney, S. (1987). *The English Noun Phrase in its Sentential Aspect*, PhD thesis, MIT. MIT Working Papers in Linguistics.

Ajello, R. & Sani, S. (eds) (1995). *Scritti linguistici e filologici*, Pisa: Pacini Editore.

Alexiadou, A. & Stavrou, M. (1998). (a)symmetries in DPs and clauses: evidence from derived nominals, *The Linguistic Review* **15**: 257–276.

Alexiadou, A. & Wilder, C. (1998). Adjectival modification and multiple determiners, *in* A. Alexiadou & C. Wilder (eds), *Possessors, predicates and movement in the DP*, Amsterdam: John Benjamins, pp. 303–332.

Androutsopoulou, A. (1995). The licensing of adjectival modification. Paper presented at the 14th WCCFL Conference, Workshop on Predication.

Aoun, J. & Sportiche, D. (1983). On the formal theory of government, *The Linguistic Review* **3**: 211–235.

Arends, J., Muysken, P. & Smith, N. (eds) (1995). *Pidgins and creoles: an introduction*, Amsterdam: Benjamins.

Baker, M. (1988). *Incorporation*, Chicago: University of Chicago Press.

Baltin, M. & Collins, C. (eds) (2001). *The Handbook of Contemporary Syntactic Theory*, Malden, Mass./Oxford: Blackwell.

Belletti, A. & Rizzi, L. (eds) (1996). *Parameters and functional heads: essays in comparative syntax*, New York, N.Y.: Oxford University Press.

Bernstein, J. (1993). The syntactic role of word markers in null nominal constructions, *Probus* **5**: 5–38.

Bernstein, J. (2001). The DP hypothesis: identifying clausal properties in the nominal domain, *in* M. Baltin & C. Collins (eds), *The Handbook of Contemporary Syntactic Theory*, Malden, Mass./Oxford: Blackwell Publishers, chapter 17, pp. 536–561.

Bittner, M. & Hale, K. (1996). The structural determination of case and agreement, *Linguistic Inquiry* **27**(1): 1–68.

Bolinger, D. (1967). Adjectives in English: attribution and predication, *Lingua* **18**: 1–34.

Bouchard, D. (1998). The distribution and interpretation of adjectives in French: a consequence of Bare Phrase Structure, *Probus* **10**: 139–183.

Brame, M. (1982). The head-selector theory of lexical specifications and the nonexistence of coarse categories, *Linguistic Analysis* **10**(4): 321–325.

Bratko, I. (2001). *Prolog programming for artificial intelligence, 3rd ed.*, Harlow: Addison-Wesley.

Butt, M. & Geuder, W. (eds) (1998). *The projection of arguments: lexical and compositional factors*, Stanford, CA: CSLI Publications.

Campbell, R. (1998). A null pronominal in the noun phrase, *Linguistic Inquiry* **29**(1): 153–159.

Cantarino, V. (1975). *Syntax of Modern Arabic Prose*, Bloomington/London: Indiana University Press.

Cardinaletti, A. (1997). Subjects and clause structure, *in* L. Haegeman (ed.), *The New Comparative Syntax*, London: Longman, pp. 33–63.

Cardinaletti, A. (2000). Towards a cartography of subject positions. ms. Università di Venezia/SSLMIt Università di Bologna.

Carlson, G. (1999). No lack of determination, *Glot International* **4**(3): 3–8.

Carstens, V. (2000). Concord in minimalist theory, *Linguistic Inquiry* **31**(2): 319–355.

Carstens, V. (2001). Multiple agreement and case deletion: against phi-(in)completeness, *Syntax* **4**(3): 147–163.

Cheng, L. & Sybesma, R. (1998). *Yi-wan Tang, Yi-ge Tang*: classifiers and massifiers, *The Tsing Hua Journal of Chinese Studies* **28**(3): 385–412.

Chomsky, N. (1981). *Lectures on government and binding*, Dordrecht: Foris.

Chomsky, N. (1986a). *Barriers*, Cambridge, Mass: The MIT Press.

Chomsky, N. (1986b). *Knowledge of Language: its Nature, Origin and Use*, New York: Praeger.

Chomsky, N. (1995). *The Minimalist Program*, Cambridge, Mass: The MIT Press.

Chomsky, N. (1998). Minimalist inquiries: the framework. MIT Occasional Papers in Linguistics 15.

Chomsky, N. (1999). Derivation by phase. MIT Occasional Papers in Linguistics 18.

Cinque, C., Koster, J., Pollock, J. & Zanuttini, R. (eds) (1994). *Paths Towards Universal Grammar: Essays in honor of Richard S. Kayne*, Washington, D.C. : Georgetown University Press.

Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP, *in* C. Cinque, J. Koster, J. Pollock & R. Zanuttini (eds), *Paths Towards Universal Grammar: Essays in honor of Richard S. Kayne*, Georgetown University Press, pp. 81–110.

Cinque, G. (1996). The 'antisymmetric' programme: theoretical and typological implications, *Journal of Linguistics* **32**: 447–464.

Cinque, G. (1999). *Adverbs and functional heads: a cross-linguistic perspective*, New York: Oxford University Press.

Collins, C. (1997). *Local Economy*, Cambridge, Mass: The MIT Press.

Collins, C. (2001). Aspects of the Plurality in ǂHoan, *Language* **77**(3): 456–476.

De Schutter, G. (1997). The noun phrase in Dutch, *Leuvense Bijdragen* **86**: 309–356.

De Wit, P. (1997). *Genitive Case and Genitive Constructions*, PhD thesis, Utrecht Insitute of Linguistics OTS.

Delsing, L.-O. (1993). *The Internal Structure of Noun Phrases in the Scandinavian Languages: a Comparative Study*, PhD thesis, University of Lund.

Delsing, L.-O. (1998). Possession in Germanic, *in* A. Alexiadou & C. Wilder (eds), *Possessors, Predicates and Movement in the Determiner Phrase*, Amsterdam/Philadelphia: John Benjamins, pp. 87–108.

Dikken, M. d. (1998). Predicate Inversion in DP, *in* A. Alexiadou & C. Wilder (eds), *Possessors, Predicates and Movement in the Determiner Phrase*, Amsterdam/Philadelphia: John Benjamins, pp. 177–214.

El-Ayoubi, H., Fischer, W. & Langer, M. (2001). *Syntax der Arabischen Schriftsprache der Gegenwart. Teil I, Band 1: Das Nomen und sein Umfeld*, Wiesbaden: Reichert Verlag.

Fassi Fehri, A. (1993). *Issues in the structure of Arabic clauses and words*, Dordrecht: Kluwer Academic Publishers.

Fassi Fehri, A. (1999). Arabic modifying adjectives and DP structures, *Studia Linguistica* **53**(2): 105–154.

Giorgi, A. & Pianesi, F. (1997). *Tense and Aspect: from Semantics to Morphosyntax*, New York/Oxford: Oxford University Press.

Giorgi, A. & Piansi, F. (1996). Verb movement in Italian and syncretic categories, *Probus* **8**: 137–160.

Greenberg, J. H. (1966). Some universals of grammar with particular reference to the order of meaningful elements, *in* J. H. Greenberg (ed.), *Universals of Language*, Cambridge, Mass: MIT Press, pp. 73–113. 2nd ed.

Grimshaw, J. (1990). *Argument Structure*, Cambridge, Mass: The MIT Press.

Grimshaw, J. (1991). Extended projection, *in* P. Coopmans, M. Everaert & J. Grimshaw (eds), *Lexical Specification and Insertion*, The Hague: Holland Academic Graphics, pp. 115–134.

Grimshaw, J. (1998). Locality and extended projection. ms. Rutgers University.

Haeberli, E. (2000). Towards deriving the EPP and Abstract Case, *GG@G (Generative Grammar in Geneva)* **1**: 105–139.

Haegeman, L. (2000). The external possessor construction in West Flemish, *GG@G (Generative Grammar in Geneva)* **1**: 1–19.

Haider, H., Olsen, S. & Vikner, S. (eds) (1995). *Studies in Comparative Germanic Syntax*, Dordrecht: Kluwer Academic Publishers.

Hale, K. & Keyser, S. J. (eds) (1993). *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, Mass: The MIT Press.

Halle, M. & Marantz, A. (1993). Distributed morphology, *in* K. Hale & S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, Mass: The MIT Press, pp. 111–176.

Haspelmath, M. (1999). Explaining article-possessor complementarity: economic motivation in noun phrase syntax, *Language* **75**(2): 227–243.

Hazout, I. (2000). Adjectival genitive constructions in Modern Hebrew: a case study in coanalysis, *The Linguistic Review* **17**: 29–52.

Hazout, I. (2001). Predicate formation: the case of Participial Relatives, *The Linguistic Review* **18**: 97–123.

Higginbotham, J. (1985). On semantics, *Linguistic Inquiry* **16**(4): 547–539.

Hinds, J. (1986). *Japanese*, London: Croom Helm.

Holes, C. (1995). *Modern Arabic: structures, functions and varieties*, London/New York: Longman.

Holm, J. (2000). *An introduction to pidgins and creoles*, Cambridge: Cambridge University Press.

Jelinek, E. (1984). Case and configurationality, *Natural Language & Linguistic Theory* **2**: 39–76.

Kayne, R. (1984). *Connectedness and Binary Branching*, Dordrecht: Foris.

Kayne, R. (1994). *The Antisymmetry of Syntax*, Cambridge, Mass: The MIT Press.

Kenesei, I., Vago, R. M. & Fenyvesi, A. (1998). *Hungarian*, London, New York: Routledge.

Kenstowicz, M. (ed.) (2001). *Ken Hale: a life in language*, Cambridge, Mass: The MIT Press.

Kester, E.-P. (1996). *The nature of adjectival inflection*, PhD thesis, Utrecht Institute of Linguistics OTS.

Kiefer, F. & Kiss, K. (eds) (1994). *The syntactic structure of Hungarian*, San Diego: Academic Press.

Kim, J.-y. (2000). Adjectives in construct. ms. University of Massachusetts, Amherst.

Koopman, H. (1996). The spec head configuration, *in* E. Garret & F. Lee (eds), *Syntax at sunset*, pp. 37–64. UCLA Working Papers in Syntax and Semantics, 1.

Koopman, H. & Szabolcsi, A. (2000). *Verbal Complexes*, Cambridge, Mass: The MIT Press.

Kornfilt, J. (1997). *Turkish*, London: Routledge.

Kremers, J. (1997). *How Arabs speak to each other about themselves: a study of* nafs *and* ba^{c}ḍ *in Modern Standard Arabic*, Master's thesis, University of Nijmegen.

Laenzlinger, C. (2000). French adjective orderings: perspectives on DP-internal movement types, *GG@G (Generative Grammar in Geneva)* **1**: 55–104.

Larson, R. (1988). On the double object construction, *Linguistic Inquiry* **19**(3): 335–391.

Lasnik, H. (1999). *Minimalist analysis*, Malden, Mass: Blackwell Publishers.

Lindauer, T. (1995). *Genitivattribute: eine morphosyntaktische Untersuchung zum deutschen DP/NP-System*, Tübingen: Niemeyer.

Longobardi, G. (1994). Reference and proper names: a theory of N-movement in syntax and Logical Form, *Linguistic Inquiry* **25**(4): 609–665.

Longobardi, G. (1995). A case of construct state in Romance, *in* Ajello & Sani (1995), pp. 293–329.

Mallén, E. (1990). Genitive clitics and nominal Infl, *Canadian Journal of Linguistics/Revue canadienne de linguistique* **35**(3): 237–263.

Marantz, A. (1984). *On the Nature of Grammatical Relations*, Cambridge, Mass: The MIT Press.

Marantz, A. (1993). Implications of asymmetries in double object constructions, *in* S. Mchombo (ed.), *Theoretical Aspects of Bantu Grammar*, Standford, California: CSLI Publications, pp. 113–150.

Menuzzi, S. (1994). Adjectival positions inside DP, *in* R. Bok-Bennema & C. Cremers (eds), *Linguistics in the Netherlands, 1994*, Amsterdam: Benjamins, pp. 127–138.

Mitchell, T. (1990). *Pronouncing Arabic, I-II*, Oxford: Clarendon Press.

Mulders, I. (1997). Mirrored specifiers, *in* J. Coerts & H. de Hoop (eds), *Linguistics in the Netherlands, 1997*, Amsterdam: John Benjamins, pp. 135–146.

Nilsen, Ø. (2003). *Eliminating positions: syntax and semantics of sentence modification*, PhD thesis, Utrecht Institute of Linguistics OTS.

Ogawa, Y. (2001). *A unified theory of verbal and nominal projections*, Oxford: Oxford University Press.

Pesetsky, D. & Torrego, E. (2001). T-to-C movement: causes and consequences, *in* M. Kenstowicz (ed.), *Ken Hale: a Life in Language*, Cambridge, Mass: The MIT Press.

Picallo, M. C. (1991). Nominals and nominalizaton in Catalan, *Probus* **3**: 279–316.

Rabel, L. (1961). *Khasi, a Language of Assam*, Baton Rouge: Louisiana State University Press.

Radford, A. (2000). NP shells. Essex Research Reports in Linguistics, vol. 33, p. 2–22.

Rappaport, G. (2001). The Geometry of the Polish Nominal Phrase: Problems, Progress, and Prospects, *in* P. Banski & A. Przepiórkowski (eds), *Generative Linguistics in Poland: Syntax and Morphosyntax*, Warsaw: Polish Academy of Sciences, pp. 173–189.

Reinhart, T. (1976). *The syntactic domain of anaphora*, PhD thesis, MIT.

Reinhart, T. (2002). *The Theta System: syntactic realization of verbal concepts*, Cambridge, Mass: The MIT Press.

Reinhart, T. & Reuland, E. (1993). Reflexivity, *Linguistic Inquiry* **24**(4): 657–720.

Reinhart, T. & Reuland, E. (1995). Pronouns, anaphors and case, *in* H. Haider, S. Olsen & S. Vikner (eds), *Studies in Comparative Germanic Syntax*, Dordrecht: Kluwer Academic Publishers, pp. 241–268.

Reuland, E. (1983). Governing *-ing*, *Linguistic Inquiry* **14**(1): 101–136.

Reuland, E. (1990). Head movement and the relation between syntax and morphology, *Yearbook of Morphology 3*, Dordrecht: Foris, pp. 129–161.

Rijkhoff, J. (1992). *The noun phrase: a typological study of its form and structure*, PhD thesis, University of Amsterdam.

Ritter, E. (1991). Two functional categories in noun phrases: evidence from Modern Hebrew, *in* S. D. Rothstein (ed.), *Perspectives on Phrase Structure: Heads and Licensing*, San Diego: Academic Press, pp. 37–62.

Ritter, E. (1993). Where's gender?, *Linguistic Inquiry* **24**: 795–803.

Rothstein, S. (ed.) (1991). *Perspectives on Phrase Structure: Heads and Licensing*, San Diego: Academic Press.

Schoorlemmer, M. (1998). Possessors, articles and definiteness, *in* A. Alexiadou & C. Wilder (eds), *Possessors, Predicates and Movement in the Determiner Phrase*, Amsterdam/Philadelphia: John Benjamins, pp. 56–86.

Scott, G.-J. (1998). Stacked adjectival modification and the structure of nominal phrases. SOAS Working Papers in Linguistics and Phonetics, vol. 8, 59–89.

Shlonsky, U. (2000). The form of the Semitic noun phrase: an antisymmetric, non N-movement account. ms. University of Geneva.

Siegel, M. (1976). *Capturing the adjective*, PhD thesis, University of Massachusetts at Amherst.

Siloni, T. (1990). On the parallelism between CP and DP: The case of Hebrew semi-relatives, *Proceedings of LCJL 1*, pp. 135–153.

Siloni, T. (1996). Hebrew noun phrases: Generalized noun raising, *in* A. Belletti & L. Rizzi (eds), *Parameters and functional heads: essays in comparative syntax*, New York, N.Y., [etc.]: Oxford University Press.

Siloni, T. (1997). *Noun Phrases and Nominalizations: the syntax of DP*, Dordrecht: Kluwer Academic Publishers.

Siloni, T. (1998). Adjectival constructs and inalienable constructions. ms. Tel-Aviv University.

Sportiche, D. (1990). Movement, agreement and case. ms. UCLA.

Sproat, R. & Shih, C. (1988). Prenominal adjective ordering in English and Mandarin, *Proceedings of NELS 18*, pp. 465–489.

Sproat, R. & Shih, C. (1991). The cross-linguistic distribution of adjective ordering restrictions, *in* C. Georgopoulos & R. Ishihara (eds), *Interdisciplinary Approaches to Language. Essays in Honor of S. Y. Kuroda*, Dordrecht: Kluwer, pp. 565–593.

Stiebels, B. (1999). Noun-verb symmetries in nahuatl nominalizations, *Natural language & linguistic theory* **17**(4): 783–834.

Szabolcsi, A. (1983). The possessor that ran away from home, *The Linguistic Review* **1**: 89–102.

Szabolcsi, A. (1987). Functional categories in the noun phrase, *in* Kenesei, István (ed.), *Approaches to Hungarian*, Szeged: Jate, pp. 167–189.

Szabolcsi, A. (1994). The noun phrase, *in* F. Kiefer & K. Kiss (eds), *The Syntactic Structure of Hungarian*, San Diego: Academic Press, pp. 197–174.

Szabolcsi, A. (1997). Strategies for scope taking, *in* A. Szabolcsi (ed.), *Ways of scope taking*, Dordrecht: Kluwer Academic Publishers, pp. 109–155.

Tang, S.-W. (1999). Some speculations about the syntax of noun phrases. in: F. Del Gobbo & H. Hoshi (eds) *UCI Working Papers in Linguistics* 5, 135–154.

Thráinsson, H. (1991). Long distance reflexives and the typology of NPs, *in* J. Koster & E. Reuland (eds), *Long-Distance Anaphora*, Cambridge: Cambridge University Press, pp. 49–75.

Travis, L. d. & Lamontagne, G. (1992). The Case Filter and licensing of empty K, *Canadian Journal of Linguistics* **37**(2): 157–174.

Tremblay, M. (1991). The syntax of possession, *in* K. Leffel & D. Bouchard (eds), *Views on Phrase Structure*, Dordrecht: Kluwer Academic Publishers, pp. 57–81.

Valois, D. (1991). *The Internal Syntax of DP*, PhD thesis, UCLA.

Valois, D. (1996). On the structure of the French DP, *Canadian Journal of Linguistics/Revue canadienne de linguistique* **41**(4): 349–375.

Wiese, H. (1998). Zur Parallelität von Nominalgruppen und Sätzen: Semantiche Argumente für den "sentential aspects of noun phrases", *Sprache und Pragmatik* **45**.

Williams, E. (1981). Argument structure and morphology, *The Linguistic Review* **1**: 81–114.

Williams, E. (1985). PRO and the subject of NP, *Natural Language & Linguistic Theory* **3**: 297–315.

Winston, P. (1992). *Artificial Intelligence, 3rd ed.*, Reading, MA: Addison-Wesley.

Wright, W. (1981). *A Grammar of the Arabic Language*, Beirut: Librairie du Liban. (reissue of the third edition of 1896).

Zribi-Hertz, A. (1997). On the dual nature of the 'possessive' marker in Modern English, *Journal of Linguistics* **33**(2): 511–538.

Zwarts, J. (1992). *X′-syntax – X′-semantics: On the interpretation of functional and lexical heads*, PhD thesis, Utrecht Institute of Linguistics OTS.

Summary

The main purpose of this thesis is to develop a minimalist model of the Arabic noun phrase. The minimalist program as proposed by Chomsky (1995, 1998, 1999) is an outline for developing a linguistic theory, but it is not a fully worked-out theory. One central idea of the model is that tree structures are purely hierarchical structures with no linear relations. This raises the question of how the linear structure is derived from the hierarchical structure.

The generally adopted answer to this question has been formulated by Kayne (1994), but his antisymmetric model, based on the LCA, has several problems. On the conceptual side, the LCA is hardly minimalist: it assumes the presence of elements such as XP, X′ and X (which in Kayne's theory is not the same as the terminal element x) which violate the principle of Full Interpretation. The LCA also does achieve what it was supposed to achieve, because it does not really predict a fixed word order. All it predicts is that the specifier and the complement of a head X will appear on opposite sides of the head in the linear string.

Furthermore, when one applies the antisymmetric model to languages with a predominant OV order in (part of) their structure, one is forced to assume a large number of undesignated functional heads and unmotivated movements, which is also not very minimalist in nature.

For this reason, I develop a different way to linearise a tree structure. The operation that is responsible for linearisation works at PF, and uses two parameters. The procedure works as follows:

(1) linearise tree T:
   a. if T is a terminal, spell out T.
   b. otherwise, take subnodes A and B of T.
   c. linearise A then B, or B then A.

In other words, the procedure takes the tree to be spelled out and analyses it. If the tree is compound, it will apply itself recursively to each subnode. If the tree is simplex, i.e., if it only contains a terminal element, this element is spelled out. Because of its recursive nature, I have dubbed the procedure RLin, for *recursive linearisation*.

Each time RLin encounters a compound node, it must decide in which order it will apply itself to the subnodes. This decision has direct consequences for the order of the terminal elements in the linear string, because the order in which the elements end up in the linear string is the order in which they are encountered by RLin when it examines the tree.

In order to decide which of two subnodes to linearise first, RLin makes use of two principles: the Head principle H, which states that heads must be linearised first, and the Selected principle S, which states that selected elements must be linearised first. At the level of a selected specifier and its sister (traditionally an X′), the selected principle forces the specifier to be linearised first. Because of this, specifiers appear before their heads.

At the level of a head and its complement, the two principles clash: H requires that the head be linearised first, S requires that the complement, which is always a selected element, be linearised first. This clash is resolved by a parameter that orders the two principles with respect to each other. If the ordering is $H > S$, the linear order becomes head-complement, if the ordering is $S > H$, the linear order becomes complement-head.

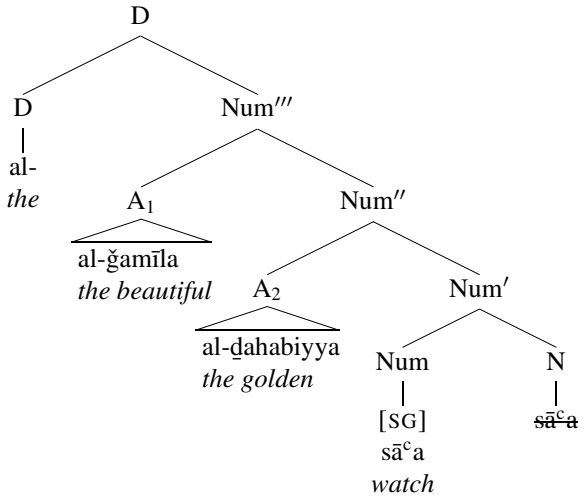
Neither of the principles S and H covers adjuncts, since adjuncts are not selected and they are also not heads. Therefore, RLin needs another parameter, the *adjunct parameter*, in order to decide whether to linearise adjuncts first or second in their nodes.

It is not the case that each category in a language must have the same settings for both parameters. It is in principle possible that two categories have different settings, such as the C and V heads in Dutch: the former is head-initial, which means $H > S$, whereas the latter is head-final, that is, $S > H$. However, we must limit the categories that can determine their own settings to those heads that have independent morphological realisation: only for those heads can a child observe the settings directly in the linguistic input. Heads that are phonologically null or are only affixal must adopt the settings of their complements.

We can demonstrate how all this works with a few examples:

(2) a. al-sā^{c}a -l-ḏahabiyya -l-ǧamīla
the-watch the-gold the-beautiful
'the beautiful gold watch'

b.



In this tree, the N and D heads can set their own parameters. The Num head is not marked,[1] which means it will adopt the settings of its complement, N.

Looking at this tree, we see that the head D must have a setting of $H > S$: its complement, Num‴ (and everything in it), follows the determiner in the linear string in (2a). The setting for the adjunct parameter of D cannot be determined from this structure, because D does not have an adjunct.[2] The adjunct parameter can be observed with the Num head: the two adjectives are spelled out after the noun *sā^{c}a* 'watch'. This is obtained by a setting of adjunct second.

The fact that Num has a setting of adjunct second means that when RLin reaches the node Num‴, it will linearise its subnode Num″ before it linearises A_1. This means that everything in Num″ will appear in the linear string before A_1. This includes the hierarchically lower adjective A_2 *ḏahabiyya* 'gold'. As a result, we obtain a mirror image order for the adjectives.

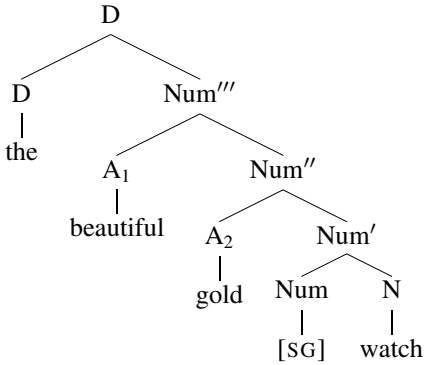
[1] And in cases where it is marked, e.g., in the dual and the plural, it is affixal.

[2] In fact, there does not appear to be any structure in Arabic in which D takes an adjunct.

Let us compare the Arabic phrase in (2) with its English counterpart in (3):

(3) a. the beautiful gold watch

b.



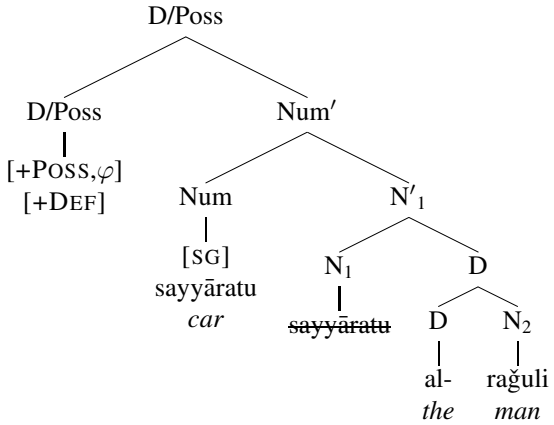
In English, the D head also has an ordering of $H > S$, because the determiner precedes the noun. The Num head, however, has a different setting of the adjunct parameter: it has adjunct-first. In (3), when RLin encounters the node Num‴, it will linearise the adjunct, i.e, the adjective A_1, before the node Num″, resulting in a linear string in which *beautiful* precedes everything in Num″. This gives us the adjective order *beautiful gold*, the exact opposite of the Arabic order.

The Arabic noun phrase has a number of features: CASE, DEF, POSS, NUM and GEN. Of these features, DEF, POSS and NUM can project independent heads. CASE and GEN do not: they are present on the noun and are inherited by the functional heads, but they do not project heads themselves.

When the POSS feature has the value [+POSS], i.e., when the head noun has a genitive complement, the DEF and POSS feature project on a single head. When this head is taken from the lexicon, it has the value [+POSS, DEF: $\emptyset$]. Because the DEF feature is unvalued, it must be valued during the derivation. This happens when the combined D/Poss head probes its complement for a match to value the set of φ-features, which it also has:

(4) a. sayyārat-u -l-raǧul-i
car-NOM the-man-GEN
'the man's car'

b.



The head Poss probes its complement for a match and finds one in the complement noun *al-raǧul* 'the man'. As a result, the unvalued φ-features and DEF feature of the head noun are valued.

Note that the head noun of a genitive structure in Arabic does indeed not have any morphological marking for definiteness, while nouns that do not have a genitive complement are always marked for definiteness or indefiniteness.

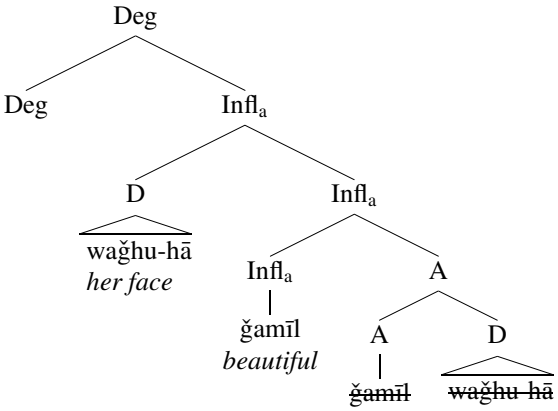
There is a particular structure in Arabic that shows that modifying adjectives can have a DegP-internal subject:

(5) ra'aytu -mra'-at-an ǧamīl-an waǧh-u-hā
I.saw woman-F-ACC.INDEF beautiful.M-ACC.INDEF face.M-NOM-her
lit. 'I saw a woman beautiful her face'
'I saw a woman with a beautiful face'

In (5), the adjective *ǧamīl* is masculine, even though the noun it modifies, *imra'a* 'woman' is feminine. The subject of the adjective is *waǧh-u-hā* 'her face'. The adjective agrees with its DegP-internal subject in gender and number. In definiteness and case, it agrees with the head noun, in (5) *imra'a* 'woman'. The DegP-internal subject has a possessive pronoun, which acts as a resumptive pronoun: it refers back to the head noun.

This means that the structure of the adjective phrase must be along the lines of (6):

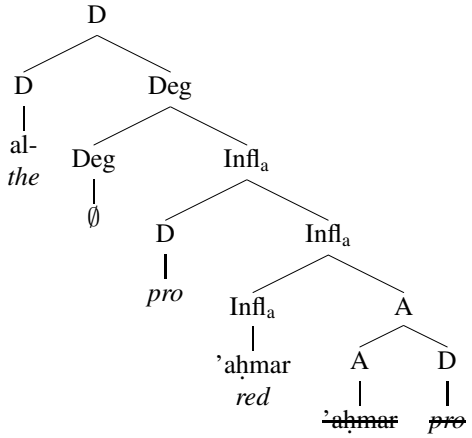
(6)



We can use the same structure to describe modifying adjectives that do not have an overt DegP-internal subject:

(7) a. al-baytu -l-'aḥmaru
the-house-NOM the-red-NOM
'the red house'

b.



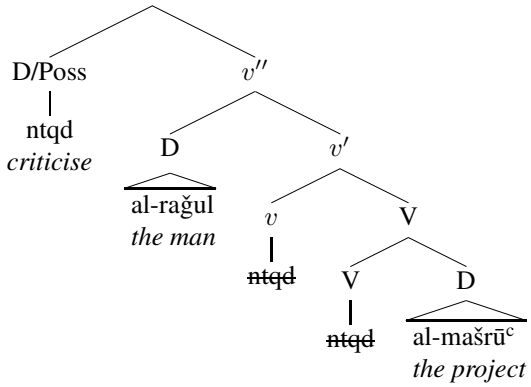
The adjective phrase in (7) also has a DegP-internal subject, but in this case it is not overt, but *pro*. This *pro* has the same function as the pronoun *-hā* 'her' in (6): it acts as a resumptive pronoun referring back to the head noun.

Arabic has a very productive set of deverbal nouns, called *masdars*. These nouns closely resemble English gerunds, in that they can assign accusative case and allow adverbials. At the same time, they also show a distinct nominal behaviour.[3] In other words, deverbal nouns in Arabic function both as complex event nominals and as simplex event/result nominals, in the sense of Grimshaw (1990).

We can give a uniform description of the uses of this type of deverbal noun if we assume that it is possible for an extended projection to switch category at some point from V to N. That is, at the point where in a clause a certain category would be projected, its nominal counterpart is projected instead. So for example instead of projecting T, the structure projects Poss, taking v as complement:

(8) a. intiqād-u -l-rağul-i -l-mašrūᶜ-a
criticising-NOM the-man-GEN the-project-ACC
'the man's criticising the project'

b.



In (8), the POSS feature is projected on a combined head with the DEF feature, because the deverbal noun has an argument, or rather two arguments in this case: both a subject and an object. The subject in (8), *al-rağul* 'the man', is in the specifier position of v. Because it is a selected specifier (it is s-selected by v), it will be linearised first in its node. This means that it will appear before the object *al-mašrūᶜ* 'the project' in the linear string.

[3] That is to say, traditionally, the term *maṣdar* is actually only used for the gerund-like use of these forms. But the same word forms can be used in a 'nominal' sense as well.

Samenvatting

Het doel van deze dissertatie is het ontwikkelen van een minimalistisch model van de Arabische noun phrase. Het minimalistisch programma zoals dat voorgesteld is door Chomsky (1995, 1998, 1999) is een opzet voor het ontwikkelen van een syntactische theorie, maar het is niet een volledig uitgewerkte theorie. Een van de kernideeën van dit model is dat boomstructuren puur hiërarchische structuren zijn, zonder enige lineaire relaties. Dit roept de vraag op hoe de lineaire structuur afgeleid wordt van de hiërarchische structuur.

Het algemeen geaccepteerde antwoord op deze vraag is geformuleerd door Kayne (1994). Zijn antisymmetrische model, dat gebaseerd is op de LCA, heeft echter enkele problemen. Conceptueel gezien is de LCA nauwelijks minimalistisch: het is gebaseerd op de aanname van het bestaan van elementen zoals XP, X′ en X (dat in Kaynes theorie niet hetzelfde is als het terminale element x). Deze elementen schenden het principe van *Full Interpretation*. De LCA doet bovendien niet waar het voor ontworpen is, omdat het niet een volledige voorspelling voor de woordvolgorde doet. Het enige dat het voorspelt is dat de specifier en het complement van een hoofd X aan weerszijden van dat hoofd zullen verschijnen in de lineaire string.

Bovendien, wanneer het antisymmetrische model wordt toegepast op talen met overwegend een OV-volgorde in (een deel van) de structuur, wordt men gedwongen om een groot aantal onbestemde functionele hoofden en ongemotiveerde verplaatsingen aan te nemen. Ook dit is niet in overeenstemming met het minimalistische principe.

Om deze redenen ontwikkel ik een alternatieve manier om een boomstructuur te lineariseren. De operatie die verantwoordelijk is voor de linearisatie werkt op PF, en maakt gebruik van twee parameters. De procedure werkt als volgt:

(1) lineariseer een boom T:
   a. als T een terminaal element is, spel T uit.
   b. zo niet, neem de subknopen A en B van T.
   c. lineariseer A en dan B, of B en dan A.

Met andere woorden, de procedure neemt de uit te spellen boom en analyseert die.

Als de boom samengesteld is, dan past de procedure zichzelf toe op beide subknopen. Als de boom niet samengesteld is, dat wil zeggen als het alleen een terminaal element bevat, dan wordt dit element uitgespeld. Vanwege haar recursieve aard noem ik de procedure RLin, *recursive linearisation*.

Steeds wanneer RLin een samengestelde knoop tegenkomt, moet het bepalen in welke volgorde het de twee subknopen behandelt. Deze beslissing heeft directe gevolgen voor de volgorde van de terminale elementen in de lineaire string, omdat de volgorde waarin de elemententen in de lineaire string terechtkomen, precies de volgorde is waarin RLin ze tegenkomt in de boom.

Om te kunnen bepalen welke subknoop eerst gelineariseerd moet worden, maakt RLin gebruik van twee principes: het Hoofdprincipe H, dat zegt dat hoofden eerst gelineariseerd moeten worden, en het Geselecteerd-principe S, dat zegt dat geselecteerde elementen eerst gelineariseerd moeten worden. Op het niveau van een geselecteerde specifier en zijn zusterknoop (traditioneel een X′), bepaalt S dat de specifier eerst gelineariseerd moet worden. Hierdoor verschijnen specifiers voor hun hoofd.

Op het niveau van een hoofd en zijn complement conflicteren de twee principes: H vereist dat het hoofd eerst gelineariseerd wordt, S vereist dat het complement eerst gelineariseerd wordt. Dit conflict wordt opgelost door een parameter die de twee principes ordent. Als de ordening $H > S$ is, dan wordt de lineaire volgorde hoofd–complement, als de ordening $S > H$ is, dat wordt de lineaire volgorde complement–hoofd.

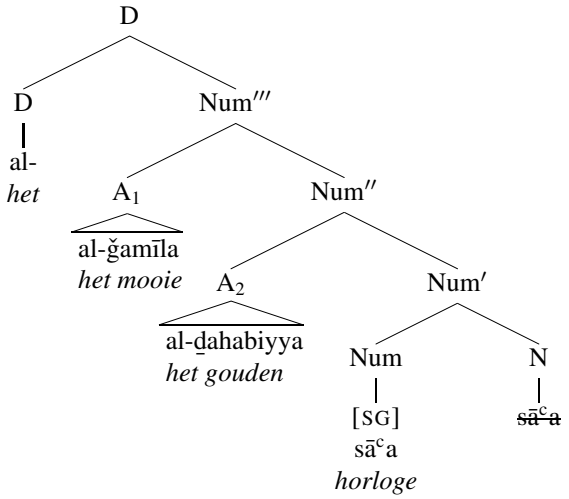
Geen van de principes S en H zegt iets over adjuncten, omdat adjuncten niet geselecteerd zijn en evenmin hoofden zijn. Daarom heeft RLin nog een parameter nodig, de *adjunctparameter*, om te kunnen bepalen of een adjunct als eerste of als tweede gelineariseerd moet worden.

Het is niet zo dat elke categorie in een taal dezelfde setting voor beide parameters moet hebben. Het is in principe mogelijk dat twee categorieën verschillende settings hebben, zoals de hoofden C en V in het Nederlands: de eerste is hoofd-initieel, wat neerkomt op $H > S$, terwijl de tweede hoofd-finaal is, wat neerkomt op $S > H$. We moeten echter wel de hoofden die hun eigen settings kunnen hebben beperken tot de hoofden die een zelfstandige morfologische realisatie hebben: slechts voor die hoofden is het mogelijk voor een kind om de settings direct uit de linguïstische data af te leiden. Hoofden die fonologisch nul zijn of slechts uit een affix bestaan moeten de settings van hun complement overnemen.

We kunnen dit alles demonsteren aan de hand van een paar voorbeelden:

(2) a. al-sā^{c}a -l-ḏahabiyya -l-ǧamīla
het-horloge het-gouden het-mooie
'het mooie gouden horloge

b.



In deze boom kunnen N en D hun eigen parameters zetten. Het Num hoofd is niet gemarkeerd,[1] wat betekent dat het de settings van zijn complement N moet overnemen.

Als we deze boom bekijken, zien we dat het hoofd D een setting van $H > S$ moet hebben: het complement van D, Num‴ (en alles wat zich daaronder bevindt) volgt op de determiner in de lineaire string in (2a). De setting voor de adjunctparameter van D kan niet uit het voorbeeld gehaald worden, omdat D geen adjunct heeft.[2] De adjunctparameter kan wel worden bepaald voor het hoofd Num: de twee adjectieven worden na het nomen *sā^{c}a* 'horloge' uitgespeld. Dit betekent dat Num de setting adjunct second heeft.

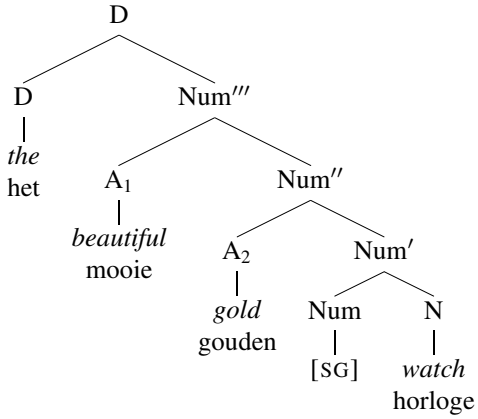
Het feit dat Num een setting van adjunct second heeft, betekent dat wanneer RLin de knoop Num‴ bereikt, het de subnode Num″ lineariseert vóór A_1. Dit betekent dat alles wat in Num″ staat, voor A_1 verschijnt in de lineaire string, inclusief het hiërarchisch lager staande adjectief A_2 *ḏahabiyya* 'gouden'. Het resultaat is dat we een volgorde krijgen voor de adjectieven die het spiegelbeeld is van de volgorde in het Engels of het Nederlands.

[1] In de gevallen waarin Num wel gemarkeerd is, bv. in de dualis en de pluralis, is het een affix.
[2] Er is sowieso geen structuur bekend in het Arabisch waarin D een adjunct heeft.

Laten we ter vergelijking het Engelse equivalent bekijken van (2):

(3) a. the beautiful gold watch
het mooie gouden horloge

b.



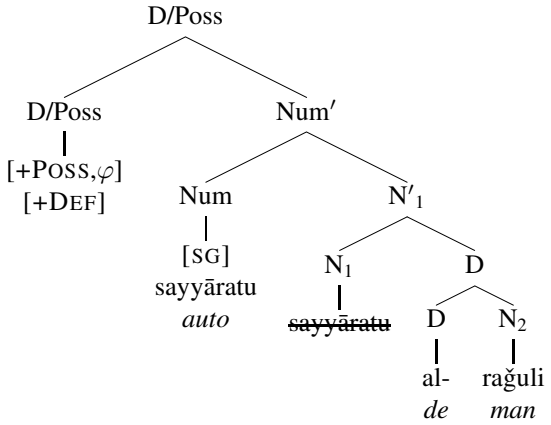
In het Engels heeft het hoofd D ook een setting van $H > S$, omdat de determiner voorafgaat aan het nomen. Het hoofd Num, daarentegen, heeft een andere setting voor de adjunctparameter: het heeft adjunct-first. Wanneer RLin in (3) de knoop Num‴ tegenkomt, zal het eerst het adjunct, dat wil zeggen het adjectief A_1, lineariseren, voor het de knoop Num″. Dit geeft een lineaire string waarin *beautiful* 'mooi' voorafgaat aan alles in Num″. Dit betekent dat we een volgorde voor de adjectieven krijgen van *beautiful gold*, precies het tegenovergestelde van de volgorde in het Arabisch.

De Arabische noun phrase bevat een aantal features: CASE, DEF, POSS, NUM en GEN. Van deze features kunnen DEF, POSS en GEN zelfstandige hoofden projecteren. CASE en GEN projecteren geen zelfstandige hoofden, maar zijn als features aanwezig op het nomen en percoleren omhoog naar de andere functionele hoofden.

Als het feature POSS de waarde [+POSS] heeft, dat wil zeggen als het nomen een genitiefcomplement heeft, dan projecteren DEF en POSS samen op een gecombineerd hoofd. Op het moment dat dit hoofd uit het lexicon wordt genomen, heeft het de waarde [+POSS, DEF: $\emptyset$]. Omdat het DEF-feature niet gevalueerd is, moet het een waarde krijgen tijdens de derivatie. Dit gebeurt wanneer het gecombineerde D/Poss-hoofd in zijn complement gaat zoeken naar een match om de ongevalueerde φ-features, die het ook heeft, te valueren:

(4) a. sayyārat-u -l-raǧul-i
auto-NOM de-man-GEN
'de auto van de man'

b.



Het hoofd Poss zoekt in zijn complement naar een match en vindt die in de vorm van het complement *al-raǧul* 'de man'. Hierdoor krijgen de ongevalueerde φ-features en ook het ongevalueerde DEF-feature op het kernnomen een waarde.

Merk op dat het kernnomen van een genitiefstructuur in het Arabisch inderdaad geen morfologische markering heeft voor definietheid. Dit in tegenstelling tot nomina zonder genitiefcomplement, die altijd gemarkeerd worden voor definietheid of indefinietheid.

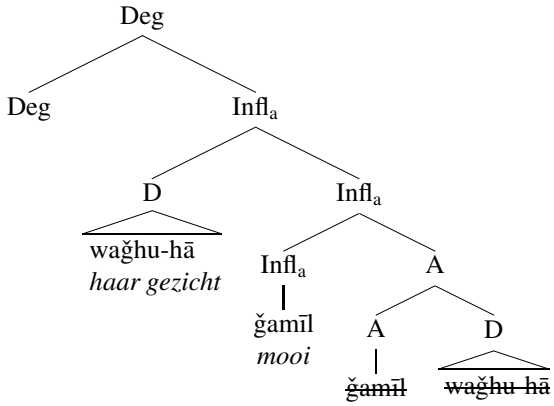
Er is een bepaalde structuur in het Arabisch die aantoont dat een modificerend adjectief een subject in de DegP kan hebben:

(5) ra'aytu -mra'-at-an ǧamīl-an waǧh-u-hā
ik.zag vrouw-F-ACC.INDEF mooi.M-ACC.INDEF gezicht.M-NOM-haar
lett. 'ik zag een vrouw mooi haar gezicht'
'ik zag een vrouw met een mooi gezicht'

In (5), het adjectief *ǧamīl* is mannelijk, terwijl het nomen dat het modificeert, *imra'a* 'vrouw', vrouwelijk is. Het subject van het adjectief is *waǧh-u-hā* 'haar gezicht'. Het adjectief congrueert met dit DegP-interne subject in geslacht en getal. In definietheid en naamval congrueert het met het kernnomen, *imra'a* 'vrouw' in (5). Het DegP-interne subject heeft een possessief pronomen, dat fungeert als resumptief pronomen en terugverwijst naar het kernnomen.

Dit betekent dat de structuur van de adjective phrase eruit moet zien zoals in (6):

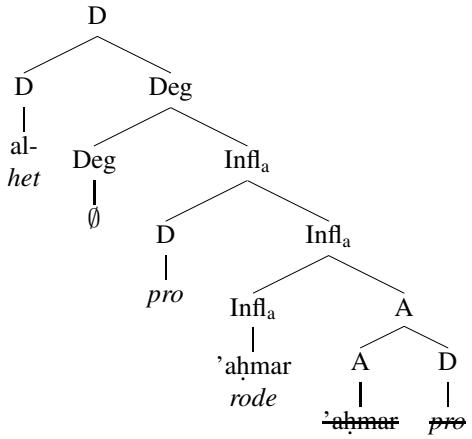
(6)



We kunnen dezelfde structuur ook gebruiken voor modificerende adjectieven die niet een overt DegP-intern subject hebben:

(7) a. al-baytu -l-'aḥmaru
het-huis-NOM het-rode-NOM
'het rode huis'

b.



De adjective phrase in (7) heeft ook een DegP-intern subject, maar in dit geval is dat subject niet overt maar *pro*. Dit *pro* heeft dezelfde functie als het pronomen *-hā* 'haar' in (6): het fungeert als resumptief pronomen dat terugverwijst naar het kernnomen.

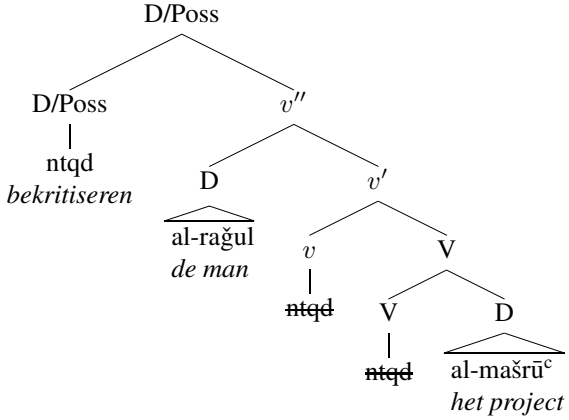
Het Arabisch heeft een zeer productieve set deverbale nomina, *masdars* genaamd. Deze nomina lijken sterk op GERUNDS in het Engels, in de zin dat ze accusatief toe kunnen kennen en adverbia bij zich kunnen hebben. Tegelijkertijd hebben ze ook een

duidelijk nominaal gebruik.[3] Met andere woorden, deverbale nomina in het Arabisch kunnen zowel complex event nominal als simplex event/result nominals zijn, in de zin van Grimshaw (1990).

We kunnen een uniforme beschrijving geven van de verschillende gebruiksmogelijkheden van deze deverbale nomina als we aannemen dat een projectie kan switchen van categorie V naar categorie N. Dat wil zeggen, op het punt waar in een zin een bepaald hoofd zou worden geprojecteerd, wordt in plaats daarvan het nominale equivalent van dat hoofd geprojecteerd. Bijvoorbeeld, in plaats van een T-hoofd te projecteren, projecteert de structuur een hoofd Poss, dat v als complement neemt:

(8) a. intiqād-u -l-raǧul-i -l-mašrūᶜ-a
bekritiseren-NOM de-man-GEN het-project-ACC
''s mans bekritiseren van het project'

b.



In (8) wordt het feature POSS geprojecteerd op een gecombineerd hoofd samen met DEF, omdat het deverbale nomen een argument heeft. Of beter gezegd, het heeft twee argumenten: zowel een subject als een object. Het subject in (8), *al-raǧul* 'de man', staat in de specifierpositie van v. Omdat het een geselecteerde specifier is (het wordt ge-s-selecteerd door v), wordt het eerst gelineariseerd. Dat betekent dat het in de lineaire string voor het object *al-mašrūᶜ* 'het project' verschijnt.

[3] In feite wordt de term *maṣdar* alleen gebruikt voor het *gerund*-achtige gebruik van deze deverbale nomina. Maar precies dezelfde woordvorm kan ook gebruikt worden in een meer nominale betekenis.

Curriculum Vitae

Joost Kremers was born in Nijmegen on 22 April 1974. He attended the Geert Groote College in Deventer from 1986 to 1992 and graduated at the level of "gymnasium bèta" in 1992.

In 1992 he started studying Arabic at the University of Nijmegen, with a specialisation in Arabic linguistics. He graduated with honours in 1997. His MA thesis entitled *When Arabs speak to each other about themselves: a study of* nafs *and* baᶜḍ *in Modern Standard Arabic* received the University MA thesis award for the Arts 1997.

After his graduation, he joined the Arabic-Dutch, Dutch-Arabic Dictionary Project that was conducted at the University of Nijmegen. In 1999, he started as a PhD student at the same university. The work carried out there resulted in the present dissertation.