

Univerzita Karlova v Praze

Filozofická fakulta

Ústav obecné lingvistiky

Bakalářská práce

Adam Pospíšil

Idáfová vazba v arabštině a její morfosyntaktické chování

The Idafa construction in Arabic and its morphosyntactic behaviour

Poděkování

Děkuji Janu Křivanovi za podnětné diskuse věnované otázkám posesivity a jeho trpělivé vedení mé práce.

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 14. srpna 2015

Adam Pospíšil

Klíčová slova (česky)

arabská gramatika, status constructus, posesivita, idáfa, head-marking

Klíčová slova (anglicky):

Arabic grammar, construct state, possessive construction, idafa, head-marking grammar

Abstrakt (česky):

Takzvaná idáfová konstrukce (idáfa) je základní prostředek arabštiny pro zachycení atributivní relace mezi dvěma jmennými frázemi. Její sémantika není přesně specifikována, ale obecně odpovídá typickému mezijazykovému konceptu posesivity. V této práci je předložena analýza způsobu značení této konstrukce, za použití základního rozlišení mezi dependent-markingovými a head-markingovými rysy, které představila Nichols (1986). Ukazuje se, že tradiční pojetí idáfy jako genitivní konstrukce může být zavádějící, neboť morfosyntaktická úloha genitivní koncovky je spíše pochybná a přítomnost atributivní relace je signalizována spíše head-markingovými rysy. Tyto rysy jsou ale poněkud méně zřejmé, jelikož nespočívají v připojení afixu k hlavě konstrukce, ale v určitých omezeních, která jsou na hlavu kladena (hlava musí být holou jmennou frází a nesmí být specifikována co do neurčitosti).

Protože je analýza zaměřena na mluvenou podobu moderní spisovné arabštiny, v této práci je brána v úvahu i socio-pragmatická variabilita spočívající v možnosti redukování jmenných koncovek v mluveném projevu, jíž se velice přínosně zabýval Magidow (2009). Tyto redukce poskytují poněkud odlišný strukturní pohled na morfosyntax této konstrukce.

V souvisejících teoretických úvahách je pojednáno o problému autonomie idáfy jakožto plně gramatikalizované head-markingové konstrukce společně s diachronními otázkami ohledně jejího původu. Pro srovnání jsou uvedeny i některé obdobné konstrukce v jiných semitských jazycích.

V některých kapitolách práce předkládá podněty k možnému empirickému zkoumání některých jevů, u nichž pravděpodobně dochází k určité variabilitě morfosyntaktického značení. Jde například o některé periferní typy užití idáfy, kdy je jejím dependentem vedlejší věta.

Abstract (in English):

The so called Idafa construction is the central means for coding an attributive relation between two nominal phrases in Arabic. The semantics of the relation is rather vague, but it generally corresponds to a typical cross-linguistic concept of possession. This paper presents an analysis of the way this construction is marked, using the fundamental distinction between dependent-marking and head-marking features, which is outlined in Nichols (1986). It turns out that the traditional label classifying Idafa as a genitive construction can be misleading, since the morphosyntactic role of the genitive ending is rather dubious and it is rather the head-marking features that signal the presence of the relation. However, these features seem a little obscure since they do not dwell in adding an affix to the head noun, but rather in imposing certain restrictions on the head (only bare heads are allowed and the head is left unspecified for the category of definiteness).

In addition, as the analysis is aimed at the natural spoken use of Modern Standard Arabic, I take into account the socio-pragmatic variation presented by the possibility of reducing the nominal endings in speech, approached insightfully in Magidow (2009). These reductions give rise to a slightly different structural view of the morphosyntax of this construction.

In the related theoretical considerations, the problem of the autonomy of the Idafa as a fully grammaticalised head-marking pattern is debated together with the diachronic issue of its origin. For comparison, some equivalent constructions from other Semitic languages are mentioned. In some chapters the paper presents ideas for possible empirical research, pointing mainly to phenomena which are expected to be subject to some variation concerning the morphosyntactic marking. These are for example certain peripheral uses of Idafa, in which a subordinate clause serves as dependent.

Obsah

1 INTRODUCTION	8
1.1 Idafa and the typological setting	8
1.2 The nature of the examples in the present paper	9
2 THE SOCIOLINGUISTIC BACKGROUND	10
2.1 Internal variation within MSA	11
2.2 A two-fold model for MSA	12
3 WHAT IS THE IDAFA CONSTRUCTION	14
3.1 The semantics of Idafa	14
3.1.1 Possession and beyond	14
3.1.2 Nominalised verbs in the head position	15
3.2 Basic morphosyntactic properties	16
3.2.1 The bare head vs. unrestricted dependent	16
3.2.2 A coordination of multiple bare nouns as the head	17
3.2.3 The behaviour of the Idafa construction with respect to definiteness	18
3.2.4 The autonomy of the head	19
3.3 Other grammatical constructions with Idafa as their basis	20
4 ARABIC MORPHOLOGY INVOLVED IN MARKING IDAFA	22
4.1 Tradition and alternative approaches	22
4.2 The grammatical state	22
4.3 Classes of nominal forms in Arabic morphology	24
4.3.1 An overview of the declensions	24
4.3.2 The Diptotes	29
4.3.3 The states for CA and LSA	29
5 THE MORPHONOLOGY OF MARKING THE IDAFA CONSTRUCTION	31
5.1 Dependent-marking features	31
5.2 Head-marking features	32
5.2.1 The definiteness rules	32
5.2.2 The differing construct forms	33

5.2.3 Anticipation of the marking due to sandhi	34
6 INTERNAL AND EXTERNAL SYNTACTIC PROPERTIES	36
6.1 Iterated Idafa	36
6.2 Adjectival modification of the Idafa construction	36
6.3 Determination	39
6.4 Some syntactic and semantic considerations related to the behaviour of multiple attributes	41
6.5 More sophisticated elements as dependents	42
6.5.1 Subordinated clauses introduced by the subjunctors 'anna and 'an	42
6.5.2 Specifying expressions introduced by the subordinating pronoun mā	44
6.5.3 Sentence-like expressions with no subjunctive	45
6.5.4 Salience of subordinate clauses in the dependent position	46
6.6 Alternative constructions and multiple dependent	47
6.6.1 The relational adjective	47
6.6.2 Analytical substitution of the Idafa	48
6.6.3 NPs competing for the position of the dependent member	48
7 DIACHRONIC AND CROSS-LINGUISTIC PERSPECTIVE	51
7.1 The concept of Idafa in Semitic	51
7.2 Grammaticalisation pathways for the Idafa	52
7.3 The degree of grammaticalisation	53
7.4 Marking or allomorphic variation?	54
7.5 Head-marking pattern borrowed	56
8 CONCLUSION	57
References	58

1 INTRODUCTION

1.1 Idafa and the typological setting

The so called Idafa (or more precisely *ʔidāfa*) is the central possessive construction in Arabic nominal syntax, coding the syntactic relation between a noun and its nominal modifier. The first is the head of the construction, determining its external properties, while the latter is its dependent.

The goal of this paper is to provide a theoretical discussion of how the Idafa construction should appropriately be approached in a typological perspective, namely in terms of the distinction between head-marking and dependent-marking as outlined by Nichols (1986).

Nichols pointed to the typological importance of the way the relation between the head and the dependent constituents of syntactic constructions is marked in different languages. As for the *locus* of marking nominal attributive constructions, there are basically four options:

a) a dependent-marking construction

CZECH
kancelář ředitel-e
office director-GEN
the director's office“

b) head-marking construction:

HUNGARIAN
a tanár iskolá-ja
DEF teacher school-3.SG.PERT
„the school of the teacher“

c) double-marking construction:

TURKISH
çocu-ğın kitab-ı
child-GEN book-PERT
„the book of a child“

d) no morphological marking, construction marked by word order:

ENGLISH
Charles University

In the preceding examples, there is always an affix which signals the presence of the relation. When the affix appears at the dependent, it is termed *genitive* which is a traditional term in European linguistics because the dependent-marking pattern is prevalent in the Indo-European family. However, when the affix appears at the head of the construction, the consensus as to its label has barely been achieved. Dixon (2009: 267) notes that *genitive* has sometimes been used in both senses, which has led to confusion. He proposes the term *pertensive* for the head-marking affix and I decided to adopt it in the glossing of certain examples, abbreviating it as PERT.

As for the *Idafa* construction in Arabic, its marking pattern is not as evident as in the above examples. Even though the classical version of the language has a *genitive* ending, its real syntactic value is rather sporadic. On the other hand, there appears to be no *pertensive* marker, but it turns out that several head-marking features can be observed in the construction.

Thus, the paper is intended to analyse the marking of *Idafa* in view of its morphosyntactic behaviour and elucidate the head-marking features which are not otherwise overly emphasised.

These considerations are going to take into account also the variation perceived in Modern Standard Arabic (MSA) since the spoken form of this language is subject to significant reductions and can be expected to exhibit patterns slightly different from the norms of the classical language.

1.2 The nature of the examples in the present paper

It should be emphasised that this paper is a theoretical one and that one of its purposes is to identify phenomena and questions on which empirical research should concentrate. Therefore, many of the sections are ended with the outline of an open problem which could be approached by means of quantitative and qualitative analysis based on natural data.

Unless the source of the examples used is indicated, they are artificially created for the purpose of illustrating certain phenomena in the grammar. In most cases they are structural phenomena for which such examples are suited. However, in some cases more sophisticated or peripheral constructions are presented, and are accompanied by my note that their relevance had better be supported by natural data.

2 THE SOCIOLINGUISTIC BACKGROUND

Any linguistic enterprise that targets the Arabic language and has the ambition of contributing to general linguistic theory needs to deal with certain sociolinguistic issues before any actual statement about Arabic is uttered. This is due to the rather intricate sociolinguistic conditions that pertain in the Arab world and that have provided fruitful ground for sociolinguistic thinking since the beginnings of the discipline.¹

The well known sociolinguistic concept associated with Arabic is that of diglossia offered by Ferguson (1959) which tries to capture the dichotomy between the variety of Classical Arabic on one hand and the various local Arabic dialects on the other. The coexistence of the different varieties of Arabic has since received more attention and more refined classifications have been proposed. Badawi (1973) proposes a hierarchy of five levels that range from *Classical Arabic* to *Illiterate spoken Arabic*. It is of extreme importance to emphasize that these layers should be conceived of not as discrete entities existing in reality, but as mere approximations of the real linguistic behaviour which is naturally scalar.

This thesis is almost exclusively concerned with the variety called Modern Standard Arabic (MSA), which is essentially a slightly (especially lexically) modernized version of the Classical Arabic of the Qur'an. This narrowing of the perspective is not motivated by an a priori advantage of such choice, but by the need to resort to some sort of discretisation to allow for any analysis to be carried out at all. The discretisation dwells in picking a single variety of Arabic that can be assumed to be definable as a more or less uniform entity. This appears to be a well tenable assumption due to the generally accepted view of Classical Arabic (or MSA) as a rigid variety anxiously preserved as intact as possible by the tradition that treats it as the holy language of Qur'an, in contrast with the Arabic dialects that have not received normative attention.

Nevertheless, we need to bear in mind the possible consequences that such an approach has for the analysis itself and, more importantly, for the impact that any conclusions can have in the domain of linguistic theory.

The crucial question that needs to be dealt with and that Owens addresses in his article on Arabic sociolinguistic (Owens 2001) is the potential of Arabic to be used for cognitively oriented linguistic research which heavily depends on the availability of a native speaker's intuition. Since MSA is

¹

Of course, even before the field of modern sociolinguistics was established, concerns of sociolinguistic nature had occupied scholars dealing with Arabic (Owens 2001).

(one could even say by definition) not a native language for anyone, it could be argued that it should be ruled out from any such research for not being a proper variety to offer insight into natural linguistic behaviour.

Undoubtedly, the absence of a native speaker's intuition does harshly decrease the scientific value of MSA in many respects. However, a notable degree of relevance of MSA is supported by the simple fact that the language has been serving natural communication in certain contexts and thus probably bears comparison with such languages as medieval Latin (which lacked native speakers, too).

Therefore, I argue that (i) choosing MSA is a defensible way of picking a particular version of Arabic that can be viewed as a single variety and (ii) that such an enterprise can contribute to general linguistic theory. At the same time, both of these assumptions require considering all along the way the traps one can face while trying to ignore some aspects that do not fit neatly into the chosen approach (either by being 'not Classical Arabic enough' or by being of doubtful real value for general linguistic theory for some particular reasons).

2.1 Internal variation within MSA

However, once we decided to leave the world of Arabic dialects aside and concentrate on the area labelled MSA, we need to zoom in even closer to look at the shape of this variety. A little more detailed view easily discovers the fact that even within MSA there is a certain degree of variability. The highly conspicuous example of it is the presence or absence of the vowel endings on verbs where they mark the three different moods and on nouns where they mark the three cases (nominative, genitive or accusative), the latter being of importance for Idafa as we are dealing here with nominal syntax. Besides marking the case, the ending can also include the final *-n* which is usually described as an indefiniteness marker.²

Example (1a) shows a sample sentence in which every possible ending is realised, while (1b) shows how these endings can be dropped.

(1) a.

²

See 4.1 for further discussion about the function of this affix.

ya-ktub-u	Muḥ ammad-u-n	kitāb-a	ta [°] līm-i-n
3.SG-write-IND	Muhammad-NOM-INDF	book-ACC	learning-GEN-INDF

(2) b.

ya-ktub	Muḥ ammad	kitāb	ta [°] līm
3.SG-write	Muhammad	book	learning

'Muhammad is writing a textbook.' (literally 'a book of learning')

The occurrence and distribution of the endings in speech is treated in a very insightful way in Magidow (2009). For us, one of his conclusions is crucial. He argues that the nominal endings, which are used rather sparsely, have ceased to be used as case markers and have acquired a function of marking highly salient nouns besides other discourse functions, which reach far beyond syntax.

This conclusion, which I decided to adopt, needs to be considered in our analysis, since it can help us arrive at a linguistically more realistic view of the morphosyntax occurring in Idafa than the view which would be offered by looking purely at Classical Arabic with the case inflection untouched. In other words, considering the classical prescribed variant would be easier since we would formally work with a well-organized elegant-looking pattern. However, there would be a higher risk of reducing the real linguistic contribution of such analyses. This goes back to the hypothesis that speaking MSA without the vowel endings is more natural for real communication, and even if the endings occur, as Magidow argues, their function is different from marking syntactic relations.

2.2 A two-fold model for MSA

For the purpose of this thesis, I decided to propose a dichotomy between two subvarieties of MSA. The first subvariety would be the truly Classical Arabic (CA), obeying the prescriptive tradition by preserving the case endings.

The second subvariety would be the 'lower' variant of MSA (LSA), characterized mainly by the absence of the nominal endings.

It should be noted that, in harmony with Magidow's article, these subvarieties are not to be seen as separate varieties of a language, but as sets of alternative variants that speakers have at their disposal. Proposing this dichotomy is primarily aimed at capturing the variability within MSA in a

discrete manner, so that we can try to find out whether there are some differences in patterns within the morphological paradigms yielded by the two subvarieties (see Chapter 5).

I am aware of the fact that this dichotomy is based on intuition and would certainly call for empirical research like that of Magidow, which would concentrate on those aspects of using Arabic which have to do with using the nominal endings and which could support the plausibility of proposing our dichotomy or they would discover what aspects are neglected by such approximation. From the cognitive point of view, the question to be asked is what 'having both of the variants at one's disposal' means for the cognitive processing of language, or, put more concretely, what the difference that we discover between the morphosyntactic patterns of the two subvarieties really means for the cognitive representation of the language in view of the fact that both of the subvarieties are somehow stored in the speaker's mind with no discrete boundaries separating them. This would not be a problem, if we could think that the two poles of the dichotomy are independent languages in which case discovering different patterns would simply mean two typological ways of organizing the morphosyntactic marking. In MSA, however, we need to treat the two varieties as coexisting alternatives within one system.

3 WHAT IS THE IDAFA CONSTRUCTION

In this chapter a sketch of some basic properties of Idafa is drawn. They will be referred to in a more detailed analysis of its morphological as well as syntactic features in the following chapters.

(3) Definition

In most general terms, Idafa can be defined as a syntactic construction which joins two noun phrases into one syntactic unit. It is a case of nominal attribution, thus one of the noun phrases serves as the head of the construction and the other as its dependent.

3.1 The semantics of Idafa

The semantics of this relation is rather unspecified and would cross-linguistically count as a typical attributive possessive construction. What is actually marked is the presence of any relation rather than its type. However, the semantics is not completely arbitrary either and has naturally received attention in descriptive works of Arabic grammar. A profound analysis is found in El-Ayoubi (2001). However vague the semantics of the relation, the construction certainly conveys some information on the directionality of the relation, which is reflected in the various functions it can assume.³

3.1.1 Possession and beyond

It should be noted that one of the primary functions of Idafa is expressing possession. Even though this function is prototypical in some sense, calling the Idafa a possessive construction requires due caution regarding the actual linguistic meaning of possession. This, however, is a more general problem of the semantics of possessive expressions in general linguistic theory. In his introduction to a chapter on attributive possession, Heine characterises this concept in syntactic terms:

³ Heine (1997:3) notes that the term of *control* has often been used in the description of possessive relations. It can partly capture the directionality of the relation. Typically the possessor exercises control over the possessee as in ⁵ilm-u rağul-i-n 'the wisdom of a man' However, also the opposite can be the case, as in rağul-u ⁵ilm-i-n 'a man of wisdom'. The directionality serves as one of the basis axes for classification in El-Ayoubi (2001).

“...attributive possession (or nominal, or phrasal possession) consists essentially of two noun phrases linked to one another in a specific way.” (Heine 1997: 143)

This is reflected in the above definition (3).

Dixon (2009: 262) claims that “[e]very language has—in its grammar—a 'possessive construction' within an NP.” As for its function, he notes that the range of concepts which are expressed by means of the possessive construction vary among different languages and provides a list of the types of relationship which are typically coded by it. These are a) *ownership*, b) *whole-part relationship*, c) *kinship relationship*, d) *an attribute of a person, animal or thing*, e) *a statement of orientation or location* and f) *association*. It is worth noting that *Idafa* counts as a highly universal construction in this regard as it covers all of these relationships. As will be mentioned in 3.3, it has also been further grammaticalised to code certain quantified and comparative constructions.

Thus, many of the semantic aspects are cross-linguistically relevant and apply to corresponding attributive constructions in other languages, including typical Indo-European genitive constructions. They are not going to be addressed here unless they play a role in the morphosyntax of the *Idafa*.

3.1.2 Nominalised verbs in the head position

Unsurprisingly, the semantic scope of *Idafa* construction exceeds the realm of possession since it also serves the function the expression of verbal complements once the head of it is a deverbal noun.⁴ In such case, the dependent member represents the subject of the verb if the verb is intransitive and its object if it is transitive, as shown in (4) and (5), respectively. This, however, is another property commonly shared by possessive constructions throughout different languages.

(4) bukā²-u-l-²aṭfāl-i
crying-NOM DEF-children-GEN
'the crying of the children'

(5) širā²-u -l-bayt-i
buying-NOM DEF-house-GEN

⁴ Dixon (2001) lists this function as f) *nominalisation*.

'the buying of the house'

3.2 Basic morphosyntactic properties

3.2.1 The bare head vs. unrestricted dependent

The main structural feature that distinguishes *Idafa* as a construction is the requirement that its head must be a bare noun.⁵ On the contrary, the dependent member of the *Idafa* can be a noun phrase of any kind or even an affixed personal pronoun.⁶ In (6) the head is *bāb* 'door', while the dependent is another *Idafa* construction, with *bayt* 'house' as head and the affixed pronoun as dependent.

- (6) *bāb-u* *bayt-i-hi*
 door-NOM house-GEN-3.SG.POSS
 'the door of his house'

The requirement imposed on the head excludes not only noun phrases composed of multiple lexemes but also the marking of any grammatical features on the head noun except the category of number and case.⁷ This means that the head is left morphologically unspecified for definiteness (see 3.2.3 and it cannot take possessive personal pronoun suffixes (unless, of course, the affixed pronoun itself is the dependent member).

The bareness of the head implies that an *Idafa* construction can only be 'branched to the right', i. e. a further embedded *Idafa* construction or otherwise complex syntagm can only occupy the slot of the dependent member, as in (6) (see 6.1). In fact, the above-mentioned exclusion of affixed personal pronouns is nothing but a consequence of this rule.

⁵ For the sake of introductory discussion, I neglect here the option of multiple coordinated bare heads, since this is a rather secondary development of the *Idafa*. It is discussed in detail in 3.2.2.

⁶ In a general treatment of *Idafa* affixed pronouns are naturally viewed as dependents, even though they differ from separate noun phrases in some respects. Besides their functional specificity, they are obviously also morphologically more closely tied to the head, simply because they are affixes and not independent words, which is testified by vowel harmony and the forms of the pronouns not occurring independently (except forms of third person in all three numbers, in which, however, vowel harmony applies).

⁷ In the case of number, this is rather unsurprising and does not really weaken the 'bareness' of the head as it is a semantically salient category and moreover is often rather lexical in Arabic, due to the so called broken plurals, as shown in 4.3. In cases where there is a plural suffix, the position of head renders it in a reduced form which again contributes to the 'bareness' of the form (see 4.3).

On the contrary, the category of case is functionally void and is absent in LSA (see 2.2). Its presence in CA can be a little surprising from a comparative Semitic perspective since for example in Akkadian the bareness of the head implies the absence of the case endings (see Retsö (2009:5)).

3.2.2 A coordination of multiple bare nouns as the head

The condition that the head must be a bare noun is essential, but to a certain degree it is a simplification since in some cases the head member of an *Idafa* construction can also be a coordination of several nouns, as in (7).

- (7) širā²-u wa bay^ʃ-u l-bayt-i
 purchase-NOM and sale-NOM DEF-house-GEN
 'the purchase and sale of the house'

In such case the bareness condition applies on all of the coordinated nouns and thus the non-final members of the coordination are also in the bare form, even though they do not precede immediately their dependent. This is worth consideration from the diachronic perspective and will be mentioned in 7.3.

However, these constructions are often avoided and replaced by a coordination of multiple *Idafa* constructions in which the common dependent is represented in the non-initial constructions by an affixed pronoun, thus yielding (8) as a syntactically more elegant equivalent of (7).

- (8) širā²-u -l-bayt-i wa bay^ʃ-u-hu
 purchase-NOM DEF-house-GEN and sale-NOM-3SG.M.POSS
 the purchase and of the house and its sale'

There are natural reasons for this preference due to the fact that the *Idafa* apparently evolved diachronically from an immediate head-dependent chain (see 7.2), and thus coordinating multiple head nouns seems rather artificial. At the same time there are also prescriptive tendencies which reflect this and may encourage MSA users to avoid these structures and replace them in the described way. Nevertheless, El-Ayoubi (2003: 523) notes that it is unjustified to assume that the option of coordinated head nouns is an innovation of MSA, since examples of it can be found in Classical Arabic as well.

3.2.3 The behaviour of the Idafa construction with respect to definiteness

The essential property of Idafa which closely relates to the actual head-marking features is its behaviour with respect to definiteness. It can be described as follows.

The head of the Idafa is not marked for definiteness and the construction as a whole receives the definiteness value of its dependent member.

Table 1 presents the different possibilities of definiteness status of nouns in independent position and within Idafa.

Table 1. The determination properties of nouns in independent position and in Idafa

indefinite		definite	
Independent noun			
bayt-u-n		ʔal-bayt-u	
house-nom-INDF		DEF-house-NOM	
'a house'		'the house'	
Nouns in Idafa			
bayt-u	mudīr-n	bayt-u	-l-mudīr-i
house-NOM	director-GEN-INDF	house-NOM	DEF-director-GEN
'a house of the/a director'		'the house of the director'	

As can be seen, the option of a definite head combined with an indefinite dependent cannot be expressed by Idafa. This rather improbable combination has to be coded by a prepositional phrase, yielding *ʔal-bayt-u li mudīr-i-n* 'the house of a director' (literally 'for/to a director' (see 6.6.2). The semantics of different definiteness options in Idafa are certainly worth a more detailed inquiry which is not the aim of this paper. As far as the marking of Idafa is concerned, it is noteworthy that when the dependent is marked as indefinite it can be interpreted semantically as definite, or, put otherwise, when the Idafa construction as a whole is intended to be indefinite, the indefiniteness marker *-n* has to be attached to the dependent, whether the latter is semantically definite or not. Therefore, as is noted in El-Ayoubi (2011: 507), even nouns which are otherwise typically definite,

such as abstract nouns, material nouns or deverbal nouns, occur with the *-n* marker once they modify a semantically indefinite head, as in (9) and (10).

(9) *qitʿat-u* *sukkar-i-n*
 piece-NOM sugar-GEN-INDF
 'a piece of sugar'

(10) *šāhid-u* *ʿiyān-i-n*
 witness-NOM observation-GEN-INDF
 'an eye-witness'

3.2.4 The autonomy of the head

The discrepancy between the location of the indefiniteness marking and its semantic reference as mentioned in the previous subsection is related to the problem of how closely the two nouns in *Idafa* are connected and to what extent the head noun has preserved its syntactic and semantic independence. This is a scale ranging from a connection of arbitrary lexemes like *bayt-u Farīd-i-n* 'the house of Farid' to compound-like collocations like *biṭāqat-u huwwīyat-i-n* 'identity card'.

A criterion of the autonomy of the head can be the extent to which modifications which morphosyntactically pertain to the construction as a whole, or even to the dependent, can be analysed as semantically modifying the head only or the whole construction.⁸ The phrasal modification is exemplified by adjectival modification and determination in (11) and the dependent-located modification by the *-n* in (10).

(11) *šāhid-u* *l-ʿiyān-i* *l-ḡadīd-u* *hādā*
 witness-NOM DEF-observation-GEN DEF-new-NOM DET.M.SG
 'this new eye-witness'

If the modification semantically affects the whole phrase, it can be argued that the connection is very close and in many cases such collocations are highly lexicalised and could be analysed nearly as composition. This possibility is even more plausible in LSA where the absence of the case endings weakens the morphological independence of the head.

⁸ Adjectival modification and determination are treated in more detail in 6.2 and 6.3, respectively.

On the other hand, true composition would require the definite article not to appear in front of the dependent, thus interrupting the would-be compound. However, as Retsö (2009: 59) shows, this can be observed in cases like the one presented in Figure 1. In this example we see a transition between an ordinary Idafa construction and a compound. The indefinite singular form shown in the left column can be treated in the definite singular form and in plural either as an Idafa construction or as a separate lexeme. In the latter case, its autonomy is testified by the formation of its own broken plural.

Figure 1: Idafa vs. compound

			Idafa		compound
			<i>ra²s</i>	<i>²al-māl</i>	<i>²ar-ra²smāl</i>
			head	DEF-capital	DEF-capital
<i>ra²s</i>	(-)	<i>māl</i>			'the capital'
head		money	<i>ru²ūs</i>	<i>a²mwāl</i>	<i>rasāmīl</i>
			heads.PL	money.PL	capitals.PL
			<i>ru²ūs</i>	<i>a²l-a²mwāl</i>	<i>²ar-rasāmīl</i>
			head.PL	DEF-capital.PL	DEF-capital.PL
		'a capital'			'the capitals'

Due to the fact that in CA indefinite Idafa phrases are marked by the *-n* affix at the dependent, this affix could structurally be seen as a phrasal clitic, even though it is phonologically attached to the dependent.

Further syntactic considerations concerning the independence of the head are treated in 6.4

3.3 Other grammatical constructions with Idafa as their basis

It should be noted that in a broader perspective a wider range of constructions actually belong to the realm of Idafa in syntactic terms. This includes mainly quantitative expressions in which the quantifier functions as head and the quantified entity as dependent, as in (12)

- (12) *ba^ʕq-u* *l-buyūt-i*
 some-NOM DEF-houses-GEN

'some houses'

A special case of Idafa are also interrogative expressions in which the head slot is occupied by the interrogative pronoun ²*ayy-u* (fem. ²*ayy-at-u*), where the dependent member is required to be indefinite, as in (13).

- (13) ²*ayy-u bayt-i-n*
which-NOM house-GEN-INDF
'Which house?'

Also many comparative expressions are in fact Idafa constructions, with the *elative* form of the adjective in the role of head⁹, as in

- (14) ²*akbar-u* *bayt-i-n*
big.elative-NOM house-GEN-INDF
'the biggest house'

However, in these grammaticalised constructions, further conditions are applied to yield the comparative semantics.¹⁰

In an even more general and rather diachronic approach, also any prepositional phrase can be viewed as an Idafa construction with the preposition in the head position.

⁹ The *elative* is an adjectival form which conveys the meaning of magnifying the property expressed by the adjective, and is used in several comparative constructions.

¹⁰ For example, an indefinite noun in the dependent position attached to an elative head gives a superlative meaning as in (14).

4 ARABIC MORPHOLOGY INVOLVED IN MARKING IDAFA

The purpose of this chapter is to provide a rather technical background necessary for dealing with the morphonology of Idafa in terms of the various nominal classes which show slightly different patterns of morphological marking.¹¹

4.1 Tradition and alternative approaches

Even though there are strong traditions of describing the morphology of Classical Arabic, in modern scholarship there are several approaches which treat certain phenomena in alternative descriptive ways. This significantly concerns the marking of definiteness, in particular the status of the marker *-n*¹², and, in general, the important notion of grammatical state.¹³

It is not my ambition here to analyse the competing views, and in the following sketch of relevant morphology I shall adhere to the traditional account of *-n* as an indefiniteness marker, which Retsö (1984-86) relates to what he labels the “Franco-German school”¹⁴.

At the same time, I am aware of the importance of such a debate, especially concerning the possibility that the traditional approach to the notions of state, determination and definiteness can be misleading in many respects. These issues are efficiently outlined in Retsö (1985), where he argues that there are in fact no plausible reasons for analysing the *-n* suffix as an indefiniteness marker and that “the nunated/diptote form of the Arabic noun is a neutral form, unmarked as regards definiteness : indefiniteness” (Retsö 1985:343)

4.2 The grammatical state

As has been mentioned in the introduction, the notion of *state* has been central in Semitic linguistics, even though rather more prevalent in languages other than Arabic. The traditional view

¹¹ See (Fischer) or other descriptive works for a complex view of Arabic morphology.

¹² The application of this marker is traditionally called *nunation*, in Arabic *tanwīn*, that is adding the letter *nūn*.

¹³ See 1.1 and 3.

¹⁴ Retsö acknowledges that this approach is predominant in the field, being pursued in contemporary scholarship among others by Fischer (1972). The alternative view rejects the assumption that the nunation per se conveys any indefinite meaning. “It is represented by a few authors, mostly with American background” (Retsö:1984-1986).

assumes the opposition between the *absolute state* (*status absolutus*) and the *construct state* (*status constructus*), the first denoting nominal forms in independent position, whereas the latter denoting dependent forms which are part of a larger syntagm,¹⁵ namely that of constructions which correspond to the *Idafa* in Arabic and are often referred to as genitive constructions.¹⁶

The forms in construct state tend to be reduced morphologically in some way due to the loss of syntactic independence which prevents them also from being marked with the definite article.¹⁷

In Arabic, the notion of state has not played a prominent role since the actual phonological difference between dependent and independent forms is less conspicuous than in other Semitic languages. It is these subtle differences which are to be covered in the morphological survey in 4.3, based on the classification by Retsö (1984-86).

In an attempt to give the term *state* a more precise definition and apply it in a more adequate description of Classical Arabic, Retsö (1984-86) proposes a three-way categorisation of states as forms occurring in different syntactical environments (see 4.3.3).

In more general perspective, the notion of state has been used more broadly for description of phenomena in nominal morphosyntax, both within the realm of Semitic languages and in other families. Creissels (2009) provides a helpful discussion of different related terms in descriptions of several African languages from different families, which deal with nominal forms resembling those encountered in the Semitic languages. He notes that for example in Berber linguistics, the term *annexed state* is used, and he sheds light on the imprecision of such usage, since the Berber forms in question are not used in the role of heads, but in that of dependents. This makes the use of the term overtaken from the Semitic background misleading, since Creissels righteously takes head-marking as crucial for any application of this notion. This follows from his view on the fundamental difference between construct state and case, which he phrases as follows:

“Case forms and construct forms are conditioned by the syntactic status of nouns, but case encodes the role of NPs as elements of broader constructions, irrespective of their internal structure, whereas

¹⁵ See Gesenius 197

¹⁶ As mentioned in the introduction, I see the term *genitive* as highly misleading in the description of this kind of constructions. However, here the complex notion of *genitive construction*, can be accepted as a conventional term for noun-noun attributive constructions in general, without reference to the morphosyntactic marking, as discussed in Chyba: *Zdroj odkazu nenalezen*.

¹⁷ Gesenius 197 This feature will be treated in detail with focus on Arabic in 3.2.3.

construct forms encode information on the internal structure of noun phrases. Case is a particular variety of dependent marking, whereas construct forms are an instance of head marking.”

(Creissels (2009:75))

Creissels himself proposes using the term *construct form*, since “it is not entirely clear whether, in the Arabic or Hebrew grammatical traditions, *construct state* primarily refers to the construction itself, or to the particular form taken by the head noun in this construction.” (Creissels (2009:2))

I nevertheless decided to use the two terms *status absolutus* (SA) and *status constructus* (SC) as practical technical terms for labeling the respective forms in examples of different Idafa constructions.

More discussion about the interpretation of the notion of state is provided in 7.4.

4.3 Classes of nominal forms in Arabic morphology

In the following overview of nominal forms I have adopted the classification by Retsö because it covers the different classes exhaustively. However, not all of the distinctions are of equal importance for the treatment of the marking in Idafa, thus I will try to comment on features which I consider more relevant for each of the classes while providing their complete overview in Table 2.

4.3.1 An overview of the declensions

A) 1st declension (strong)

The 1st declension covers an overwhelming majority of Arabic lexicon and includes the majority of masculine singulars as well as many broken plural forms. In CA they take the case endings and in state III they are marked by the final *-n*. In LSA its members have a single form without any ending.

B) Feminine declension

This class covers forms which are marked by the feminine suffix *-a(t)*.¹⁸

¹⁸ These do not have to be feminine in terms of natural gender. Many broken plural forms formally belong to this class.

It behaves identically to the 1st declension class within the CA variety. However, its behaviour is remarkably different within LSA (as will be emphasised in chapter 5). The final *-t* of the ending disappears in LSA, but is preserved when the noun appears as the head of *Idafa*, i.e. in the construct form, thus creating a different marking pattern in LSA.¹⁹

C) Animate plurals

This class contains external masculine animate plurals which are formed by the ending *-ūn(a)* in nom. and *-īn(a)* in gen./acc. The final vowel of these endings disappears in LSA.²⁰ In the construct form, a reduction of the ending takes place, yielding *-ū/-ī* (in both CA and LSA). In LSA the case distinction is often abandoned and only *-ī(n)* appears in nominative as well. This is better seen as disappearance of the case distinction in LSA.

D) The dual

Similarly, the class of the dual forms is characterized by the ending *-ān(i)* in nom. and *-ayn(i)* in gen./acc. Again, the final vowel disappears in LSA or in pausal forms. In the construct form, only *-ā/-ay* remains.

E) Affix *-āt*

This class covers the majority of animate feminine plurals and plurals of lexemes of both genders which do not form broken plurals, being also a default plural marker for inanimate borrowings. It is formed by the ending *-āt*, to which in CA the nom. ending *-u* and gen/acc ending *-i* is attached, followed by the nunation, which renders it inflectionally similar to the 1st declension.

F) Weak nouns – 2nd and 3rd declension

¹⁹ In fact, even within CA the final *-t* vanishes together with the following case ending and nunation, or, more precisely, is weakened to *-ah*, in the so called pausal forms, that is occurring at the end of an utterance or a verse.

²⁰ As with the feminine affix *at*, the omission of the final *a* of *-ūna/-īna* also occurs in pausal forms within CA.

This class covers certain forms of nouns having a weak vowel, i. e. *w* or *y* as the third radical.²¹ Namely those forms, in which the final weak consonant brings about a reduction of the final syllable once the marker *-n* is attached.

Thus, for phonological reasons, a difference arises between the indefinite form marked by *-n* and the form which either takes the definite article ²*al-* or is followed by another noun, being the head of an *Idafa* construction.

In accordance with Retsö (1984-86) I distinguish two subclasses of weak nouns, labelled 2nd and 3rd declension. The first one merges the nominative and genitive forms, while the latter only has a single form for all three cases.²²

Even though this is a structurally remarkable feature, I consider the weak nouns rather marginal for the purposes of this paper, since the endings do rarely occur in the prescribed variants in spontaneous speech, especially once the nunation is eroded. I admit that I also lack an appropriate apparatus for dealing with the weak nouns in terms of the inner variability within MSA. The problem is that the weak nouns require some additional rule to change from CA inflected forms for LSA reduced forms, besides the simple omission of the case ending and the nunation.²³ This is because in the weak nouns the original case ending is phonologically merged with the final weak radical in nominative and genitive nunated forms in the 2nd declension, and in all cases in the 3rd declension. Thus, dropping it simply together with the nunation is not possible. Presumably, such forms end up being identical with the two other states, having the final vowel prolonged once the final *-n* is not present. Unfortunately, it remains unknown how speakers actually treat the weak declensions while speaking MSA, and it is another problem which requires empirical research

²¹ As is known about Arabic morphology in general, an overwhelming majority of lexemes are built by imposing a vocalic pattern on a three-consonantal root. The three consonants are sometimes called radicals, and in some cases their phonetic quality can cause modifications to the applied vocalic pattern, as is the case with the so called weak vowels *w* and *y*.

²² For the account of different morphonological classes to be exhaustive, it should be noted here that also diptotic forms occur in this class, because a root ending in a weak consonant can also naturally appear in a pattern that is diptotic. However, in the case of weak nouns of the 2nd declension it is only the accusative form which is diptotic whereas it is the merged nominative and genitive that can take the nunation.

In case of the 3rd declension, the uninflected forms which are classified as diptotic are in fact not weak nouns since the final *-ā* is not the third radical but an obsolete feminine suffix which cannot take the nunation.

²³ By “rule” I do not mean any rigid phonological command that applies universally when speakers intend to speak some version of MSA, which I am modeling here by LSA. It is probable that individual speakers deal with it differently and it is not but another language problem which they face while using MSA.

carried out on spoken data.²⁴ As far as the variability model based on the opposition of CA vs. LSA, the indefinite forms of weak nouns within LSA are thus uneasy to define and my choice of the LSA forms is rather arbitrary here.

G) The five-noun declension²⁵

This is a class which includes five items, namely *ʔab*, 'father', *ʔaḥ*, 'brother', *ʕam*, 'uncle', *fū*, 'mouth', and the relational pronoun *dū*. Retsō also treats it as a subclass of the 1st declension. However small this class may be, it is worth attention, similarly to the feminines in LSA, for having a special form when occurring as head of *Idafa*, with the prolonged final vowel, i. e. *ʔabū -l-walad-i* 'the father of the boy' and *ʔal-ʔab-u* 'the father'.

Table 2. The endings of the different noun classes

		CA				LSA	
		state	I	II	III	I	II
	case			dipt.	tript.		
A) 1st declension	nom	-u	-u	-un	-u	∅	
	gen	-i	-i	-in	-a		
	acc	-a	-a	-an	-a		
B) Feminine declension	nom	-at-u	-at-u	-at-un	-at-u	-a	-at
	gen	-at-i	-at-i	-at-in	-at-a		
	acc	-at-a	-at-a	-at-an	-at-a		

²⁴ The problem of weak nouns is not specifically treated even in Magidow even though his discussion of different variants is otherwise very comprehensive.

²⁵ This set of nouns is traditionally called *ʔal-ʔasmāʔ ʔal-ḥamsa* 'the five names'.

C) Animated plural	nom	-ūna	-ū	-ūna		-ūn (-īn)	-ū (-ī)
	gen/acc	-īna	-ī	-īna		-īn	-ī
D) Dual	nom	-āni	-ā	-āni		-ān (ayn)	-ā (ay)
	gen/acc	- ayni	-ay	-ayni		-ayn	-ay
2nd declension (weak)	nom	-ī	-ī	-in	-in		-ī
	gen	-ī	-ī	-in	-in		
	acc	-iya	-iya	-iyan	-iya		
3rd declension (weak)	nom/gen/acc	-ā	-ā	-an	-ā		-ā
Five-noun class	nom	-u	-ū	-un		∅	-ū
	gen	-i	-ī	-in			-ī
	acc	-a	-ā	-an			-ā

My version of the classification differs from Retsö's in treating separately the class of feminine forms, which in Retsö's classification is subsumed under the 1st declension. I prefer this because of the special behaviour of the feminines with respect to *Idafa* from the perspective of the variation between CA and LSA.²⁶ I also treat separately the so called five-noun class.

Retsö uses the term *declension* for referring to the various nominal classes in accordance with the tradition in Indo-European linguistics.²⁷ Unlike typical Indo-European declension classes, however, the Arabic declensions are rather classes of forms which cover different vocalic patterns in which lexemes occur or even forms with an affix marking the category of number. This means that the broken plural forms are treated as separate lexemes.²⁸ Thus, it is not that one nominal lexeme as such belongs to one declension class, but its singular form can inflect according to one class and its plural form according to another one.

²⁶ This is of little concern to Retsö since his article only accounts for the CA variety.

²⁷ I decided to replace it with the form *declension*, which seems to be better established in English-language discourse.

²⁸ The so called broken, or internal, plurals are plural forms which are formed by changing the vocalic pattern of the word and morphologically behave as the singulars while syntactically showing agreement with feminine singular forms. Indeed, plural formation in case of broken plurals is rather a lexical than inflectional category.

4.3.2 The Diptotes

Within the first three declensions, a subclass of the so called diptotes is subsumed, opposed to the remaining majority known as triptotes.²⁹ The diptotes are forms which cannot take the *-n* marker and have a single ending marking both the genitive and accusative.³⁰ Their behaviour with respect to definiteness is worth attention because once a diptote noun turns definite, either by the prefixed definite article *ʔal-* or by being the head of an *Idafa* construction, it behaves like a triptote, taking all three case endings.³¹

They cover both singular and plural forms. However, I consider the specific behaviour of the diptotes marginal because it can only be observed when case endings are used. Therefore, it is relevant for the CA subvariety only, while in LSA no diptote class can be defined. It is thus not of such concern to the focus of this paper.

4.3.3 The states for CA and LSA

For each declension Retsö proposes three *states* which are forms in which the given word occurs in different syntactic environments. In each of the states, always three case forms are available.³² This creates a paradigm in which state and case are orthogonal categories, the first given by the clause-internal environment, the latter by the clause-external environment.³³

²⁹ The class of feminines is almost completely triptotic, with the exception of some proper names like *Fātimat-u* 'Fatima', which are not going to be of any concern for this paper.

³⁰ Having only two case forms is the reason why they are called diptotes, as opposed to the triptotes, which have three case forms.

³¹ Again, this is a wording rooted in the traditional account of Arabic morphology. In a more modern approach, proposed by Retsö (1984-86), there is no need for assuming that a diptote lexeme turns a triptote in certain conditions. The triptote/diptote distinction does not function here on the level of the lexicon, but it is rather a certain set of forms within the paradigm which can be diptotic for certain lexemes. Namely, it is the indefinite forms, or, in Retsö's description, the state III forms in the first three declensions. For the sake of precision, it should be noted here that in the second (weak) declension, it is only the accusative that shows the diptotic ending *-a*, whereas the genitive falls together with the nominative and they both have the nunated form ending *-in*.

³² Including some case syncretism, given by the diptote subclass.

³³ See Creissels' quotation in 4.2.

State I is defined as “the form of a nominal nucleus dependent on a determinator” By *determinator* Retsö refers to a set of 'peripheral morphemes', including the definite article *ʔal-*.³⁴

State II 'is the form of the nominal nucleus dependent on a following modifier. The modifier can be a peripheral bound morpheme following the nucleus (in Arabic these are the pronominal suffixes) or another nominal nucleus. Indeed, this is the case of an *Idafa* construction. As he notes, “also whole sentences may function as modifiers” (see 6.5).

State III is “the form of a nominal nucleus without preceding determinator or following modifier.” Thus, the nunated forms as well as diptote forms are seen as independent forms forming a syntagm on their own.

The LSA variety can be accounted for by means of only two states which can be labeled *absolute* and *construct* in accordance with the traditional terminology.

³⁴ The other markers in this set are the vocative particle *yā* and the general negation particle *lā*, taking the nominative and accusative case forms in Retsö's state I, respectively, e.g. *yā malik-u* 'oh king', *lā malik-a* 'there is no king'.

5 THE MORPHONOLOGY OF MARKING THE IDAFA CONSTRUCTION

This chapter is concerned with the description of morphological marking in Idafa construction. We begin with the dependent-marking features, and proceed to a detailed view of the more delicate head-marking features which are probably less visible and familiar to an Indo-European eye. The description is in fact nothing but highlighting the given features, which have already been technically described in the previous chapter and summarised in Table 2, from the perspective of the head- and dependent-marking.

5.1 Dependent-marking features

The dependent-marking element is the genitive ending attached to the dependent member. Its form is *-i* and occurs as an agglutinative morpheme in the 1st declension and in the feminine declension and after the *-āt* suffix.³⁵ In the LSA variety the ending is dropped.

In the weak 2nd declension the ending is merged with the third radical into *-in* or *-ī*, and there is no difference between the nominative and accusative forms. In the weak 3rd declension no case marking is present. In LSA, the final *-ī* is not dropped, but it would be highly inadequate to ascribe any genitive function to this morpheme within LSA. Indeed, it only represents the third weak radical there.³⁶

For diptotic forms within the 1st declension the genitive ending has the allophone *-a*, which renders it identical with the accusative.

In the plural class the gen./acc. ending is *-īn(a)* as opposed to the nom. ending *-ūn(a)*. The final vowel *a* is in parentheses because it is dropped in LSA. Moreover, there is an important additional variation that sees the gen./acc. *-īn(a)* often replacing *-ūna* (or *-ūn*) even in the nominative case, which means that the case distinction disappears here, as it does in other declensions through the omission of the endings. See (Magidow 2009)

The behaviour of the dual forms is identical to that of the animate plurals, the actual endings being gen./acc. *-ayn(i)* vs. the nom. *-ān(i)*. The final *i* is dropped in LSA.

³⁵ In the case of the *-āt* suffix, however, it is homonymous with the accusative ending.

³⁶ See for discussion of the difficulties with accounting for the weak nouns in LSA.

We see that in LSA the genitive ending is not preserved in any of the declensions, except the animate plural and dual classes. And even within these two classes its preservation is rather debatable as it tends to merge with the nominative ending.³⁷

These facts point to the complete loss of case marking. An essential insight into the functional nature of the endings is offered in Magidow (2009) when he argues that the syntactic validity of the ending has been replaced by other functions in discourse (see 2.1).

In conclusion, it is evident that the dependent-marking elements are either not present at all (this case being represented by LSA) or they can be present, but their syntactic validity, that is actually marking the syntactic function of the genitive which signalizes a relation to the head noun, is dubious.

5.2 Head-marking features

5.2.1 The definiteness rules

The crucial head-marking feature of Idafa is grounded in its behaviour concerning definiteness, which was outlined in 3.2.3. Since the Idafa construction as a whole receives the definiteness value of its dependent member, its head noun cannot be preceded by the definite article, nor followed by the indefiniteness marker *-n*.³⁸

The absence of either of the markers, i. e. the suffix *-n* and the definite article *ʔal-*, however, is not equally significant from the perspective of its marking potential, which follows from taking into account both the SA and LSA subvarieties. The absence of the definite article is a trait that dominates both of the varieties. To the contrary, the *-n* marker is generally disappearing in LSA, where the indefinite form is left unmarked (i.e. without affixes). Therefore, the absence of the *-n* marker is relevant only for SA.

³⁷ In fact, it's an arbitrary structural decision of mine to preserve the distinction of *-ūn/-īn* and *-ān/ayn* within the model of LSA, and probably it would be more adequate to pick the *-īn* and *-ayn* variants only. I have preserved it because the model of LSA is essentially based on the apocope of final segments.

³⁸ See 4 for discussion of alternative views on the nunation.

5.2.2 The differing construct forms

a) The feminines

As has already been shown, the feminines provide valuable evidence for the presence of the head-marking in the construction. While in CA they behave identically to the nouns in the 1st declension, in LSA the final *-t* of the feminine suffix is retained in the construct form, differentiating it from the absolute ending *-a*. A diachronic explanation is offered in 17.2. Yet the synchronic account

b) The animate plural and dual ending

These two pairs of endings exhibit an overt differentiation between the absolute and construct form, being truncated in the latter. This opposition is preserved in LSA, since it is only the final vowel of the absolute ending that is dropped. Thus absolute *-ūna/-īna* opposed to construct *-ū/-ī* in CA yield absolute *-ūn/-īn* opposed to unchanged construct *-ū/-ī* in LSA. The same scheme applies in the case of the dual, where CA *-āni/-ayni* and LSA *-ān/-ayn* are both opposed to construct *-ā/-ay*. As has been mentioned earlier, the opposition of the two case forms, divided by a slash above is often neglected in spoken Arabic and it is probably more adequate to treat it rather as free variation in LSA, with the prevalence of the ending corresponding to the gen./acc. form in CA, i. e. absolute *-īn* and construct *-ī*. As for dual, its adequate modeling in LSA is rather elusive since in general speakers tend to avoiding it in spontaneous speech and it is rather reserved for the written language. Nevertheless, how they deal with it when they do use it is another empirical question that would require empirical research, similarly to that of the weak nouns, as mentioned in 4.3.1.

c) Other special forms

There is a minor group of lexemes for which the absolute and construct form differ phonologically. They are results of individual idiosyncratic diachronic paths. I shall only show two examples, namely the word for word for 'God' and 'woman'. In the sense of the monotheistic god always appears in a definite form *ʔallāh-u*, in which the definite article *ʔal-* has merged with the root. The indefinite form which could refer to any god from potentially multiple deities is *ʔilāh-u-n*. Once the word serves as the head of Idafa, the basis *ʔilāh* is used which gives expressions like *ʔilāhī* 'my God' or *ʔilāh-u -l-muslimīna* 'the Good of the Muslims'. Synchronically, this can be viewed as marking the head of the Idafa by having a special form for it. This form is identical to the indefinite form of

the word, but since in the sense of the single God the lexeme always occurs as definite, this form can be seen as marked for being the head of Idafa, opposed to the independent form ²*allāh*.³⁹

The word for 'woman' is a structurally similar example. Its definite form is ²*al-mar²at-u*, opposed to the construct form (²*i*)*mra²at-u*, in which the initial syllable is a “prothetic” one in the sense that it is dropped once the form is preceded by a word ending in a vowel. The basis of the construct form is identical to the one for the indefinite form (²*i*)*mra²at-u-n*.

These examples, however minute the nuances are, could probably be seen as cases of morphological suppletion which involves marking the Idafa. It is not surprising that it occurs with such frequent lexemes.

5.2.3 Anticipation of the marking due to sandhi

Since we are concerned with the possible cognitive aspects of the particular marking patterns, the following morphonological detail is worth consideration. Arabic has a set of sandhi rules that concern most prominently the definite article ²*al*, but also other linking phenomena. While a larger set of them is strictly applied in the classical quranic language, in spoken MSA, which is modeled here by LSA, they do not need to be observed so rigorously and should probably be viewed as rather optional. However, the reduction of the definite article naturally occurs massively due to its high frequency.

The rule says that if the article is preceded by a word ending in a vowel, it drops its own vowel and merges with it into one syllable. Moreover, if the preceding vowel is long, the syllabic structure of Arabic which only exceptionally tolerates closed syllables with long vowels, causes the vowel to be shortened. This phenomenon is mainly visible in prepositional phrases where a preposition ending in a vowel is merged with the article into one syllable. A prototypical example would be the preposition *fī* meaning 'in'. It is used in the following scheme (15) which shows the syllable merger as well as the corresponding indefinite phrase and a prepositional phrase with Idafa (the endings which occur in CA are in parentheses).

- (15) a. *fī* *bayt(-i-n)*
 in a house

³⁹ In specific usage the lexeme ²*ilāh* can behave also regularly, yielding a definite form ²*al-²ilāh*, which can refer to a single god, but not necessarily the one of the monotheistic religions.

- b. *fī* + ²*al-bayt(-u)* > *fī-l-bayt(-i)*
in the house in the house
- c. *fī* + *bayt(-u) l-mudīr(-i)* > *fī bayt(-i) l-mudīr(-i)*
in house of the director in the house of the director

I prefer not to consider either the indefinite nor the definite noun phrase the unmarked one and do not want to embark on any transformational enterprise which would consider the definite phrase a derivation of the indefinite one or vice versa. The fact that is relevant for a cognitively oriented view is that once the noun occurs as the head of an *Idafa* construction that is definite as a whole by its definite dependent member, it needs to be marked by the absence of the definite article. Moreover, in the case of the syllable merger the necessity to execute the marking is shifted even one segment to the front. This is due to the elision of the article which is accompanied by rendering the preceding vowel its original length. Even though phonologically it remains the same segment (the marking concerns the syllable preceding the noun), morphologically this actually means marking the *Idafa* as early as on the preposition. I believe that this is a worthy contribution to the analysis of the role anticipation plays in marking syntactic relations. The phenomenon has not been empirically researched so far, but the personal experience with acquiring *Idafa* which I have gained as a learner of Arabic points to its possible significance.

On the other hand, I am aware that this phenomenon can barely be analysed as an actual grammatical marking of the *Idafa* on the preposition, as it is only a consequence of the actual marking, i. e. the absence of the definite article ²*al-*, which is confronted with the given morphonological sandhi rule. This, however, doesn't make it irrelevant for a comprehensive account of the way *Idafa* is marked, as speakers need to process it in production.

6 INTERNAL AND EXTERNAL SYNTACTIC PROPERTIES

6.1 Iterated Idafa

As outlined in 3.2.1, the dependent member of an Idafa construction can itself be an Idafa construction, as in 1. The iterated embedding of Idafa constructions is not limited and chains of several embedded attributes are not uncommon.

- (16) [maktab-u]^{SC} [mudīr-i]^{SC} l-bank-i
office-NOM director-GEN DEF-bank-GEN
'the office of the director of the bank'

Naturally, the potentially infinite embedding is subject to cognitive factors, and as El-Ayoubi (2011:512) notes, chains containing more than four members seldom occur. There are other syntactic possibilities for coding sequences of multiple embedded attributes, such as attributive prepositional phrases. This alternative would give (17) as an equivalent of (16). It is treated in 6.6.2.

- (17) ʔal-maktab-u li-[mudīr-i]^{SC} l-bank-i
office-NOM to-director-GEN DEF-bank-GEN
'the office of the director of the bank'

Regarding the head-marking issue which is in the focus of this paper, the hypothesis which I make is that one of the cognitive factors at play in speakers propensity to avoid more elaborate Idafa chains is the fact that the head-marking, namely the absence of the definite article, requires anticipation of the structure of the syntagm which is being produced (see also 5.2.3). It appears that for this hypothesis to be tested, empirical research would have to be conducted on a sample of languages that show both types of marking with control of a range of other possible factors including word order.

6.2 Adjectival modification of the Idafa construction

The behaviour of modifying adjectives in a way reveals the internal coherence of Idafa as a construction. An adjective modifying the head always comes only after the dependent member, that

is after the Idafa construction as a whole, as shown in (18).⁴⁰ This is again a consequence of the principle invoked in 3.2.1, i. e. the required 'bareness' of the head.

In syntactic terms this means that an adjective modifying the head of an Idafa construction is less tightly connected to it than the noun phrase functioning as its dependent within Idafa. It is in many cases natural to view the adjective as modifying the Idafa construction as a whole. See 6.4 for further theoretical treatise.

- (18) a. ²*al-maktab-u* *l-ğadīd-u*
 DEF-office-NOM DEF-new-NOM
 'the new office'
- b. [*maktab-u*]^{SC} *l-mudīr-i* *l-ğadīd-u*
 office-NOM DEF-director-GEN DEF-new-NOM
 'the new office of the director'

An important consequence of this property is the occasional ambiguity which arises when the head and dependent nouns share the same values of agreement categories, i. e. number and gender and case, as in (19). In such case, the adjective can be interpreted as modifying either the dependent member, or the head noun.⁴¹

- (19) *fī* [*maktab-i*]^{SC} *l-mudīr-i* *l-ğadīd-i*
 in office-GEN DEF-director-GEN DEF-new-GEN
 'in the new office of the director' or 'in the office of the new director'⁴²

It is worth noting that even though the category of case is an agreement category and one would argue that it is the case agreement between the head noun and the adjective that is often apt to make the attribution unambiguous, as with the nominative in (18), in fact, agreement in case is a rather illusory phenomenon. As has been argued earlier in 2.1, the so called case endings barely serve any syntactic marking, unless one imagines this within an artificial situation of deliberately analysing the syntax of a sentence with the help of these endings.

⁴⁰ In Arabic, modifying adjectives are always postponed.

⁴¹ The second option, i. e. modifying the head, represents here also the possibility of modifying the whole Idafa construction, as invoked in 6.4.

⁴² All prepositions in Arabic take the genitive.

Thus, only the categories of gender and number can in reality serve to avoid the ambiguity. Indeed, even (18) is ambiguous if we regard the use of different endings as free variation or simply look at this example in LSA, as shown in (20)

- (20) [maktab]^{SC} ʔal-mudīr ʔal-ǧadīd
office DEF-director DEF-new
'the new office of the director'

Therefore, this ambiguity is relatively frequent and can occur even in iterated Idafa constructions. It is might be one of the reasons why Idafa constructions are sometimes avoided when more complicated phrases would arise. This is further treated in 6.6.3.

If the number and gender values are not shared this ambiguity disappears, as in (21).

- (21) [maktab-u]^{SC} l-mudaráʔ-i l-ǧudud-i
office-NOM DEF-ředitel-GEN DEF-new.PL-NOM

It should be added that there is no syntactic rule preventing an adjectival modification of both the head and the dependent. El-Ayoubi (2011: 521) provides examples of this, like in (22).

- (22) [qiṭāʔ-āt-u]^{SC} š-šaʔb-i t-tūnisīy-i l-muḥtalif-at-u
fraction-PL-NOM DEF-people-GEN DEF-Tunisian-GEN DEF-various-FEM-NOM
'the various fractions of the Tunisian people'⁴³

In general, it seems that even though the adjectival modification is not syntactically restricted by the properties of Idafa, the sequential complications and ensuing semantic ambiguity or processing difficulties may lead to its avoiding by speakers.

⁴³ The adjective is in the feminine form since inanimate plurals such as *qiṭāʔ-āt* behave syntactically as feminine singulars.

6.3 Determination

The way modification by demonstrative pronouns is carried out reveals, similarly to the adjectival modification, the internal coherence of the Idafa. In this case, however, even a change in the ordering of segments appears. While in non-Idafa noun phrases the demonstratives always precede the head noun, once an Idafa construction is formed, a demonstrative pronoun modifying the head is moved to the very end of the construction.⁴⁴ The difference is shown in (23a-b).

- (23) a. *hādā* *-l-bayt-u*
 DET.M.SG DEF-house-NOM
 'this house'
- b. [*bayt-u*]^{SC} *-l-mudīr-i* *hādā*
 house-NOM DEF-director-GEN DET.M.SG
 'this house of the director'

If the Idafa is modified by an adjective, the demonstrative comes only after it, claiming the phrase final position, as in (24).⁴⁵

- (24) [*bayt-u*]^{SC} *-l-mudīr-i* *l-ğadīd-u* *hādā*
 house-NOM DEF-director-GEN DEF-new-NOM DET.M.SG
 'this new house of the director'

When the demonstrative is modifying the dependent member, it can precede it as there are no restrictions applying on the dependent. As El-Ayoubi (2003:521) points out, this means that the demonstrative in this position is the only element which can occur between the head and the dependent of the Idafa, as in (25).⁴⁶

⁴⁴ El-Ayoubi 522 notes that the demonstrative normally precedes a noun phrase only in company with the definite article and since in Idafa no article can occur at the head, the demonstrative cannot retain its preverbal position.

⁴⁵ This suggests the adjectival nature of its behaviour in Idafa, since the elements surrounding the noun are expected to be ordered in the same hierarchical sequence whether they unfold to the left or to the right. Hence, the demonstrative as the most external modifier is shifted to the very end once it cannot precede the noun.

⁴⁶ El-Ayoubi (2003:521) righteously notes, besides demonstratives, also quantifiers can come before the dependent member, as in *nata²iğ-u akṭar-i -l-džuhūd-i* 'the results of most of the efforts' and *nata²iğ-u kull-i l-džuhūdi* 'the results of all the efforts'. However, as has been pointed out in 3.3, these are cases which can be treated as Idafa constructions themselves with the elative form ²*akṭhar* 'most' and the quantifier *kull* 'all' as heads.

- (25) [bayt-u]^{SC} hāḏā -l-mudīr-i
 house-NOM DET.M.SG DEF-director-GEN
 'the house of this director'

Theoretically, one can easily imagine an example in which both the head and the dependent are modified by a demonstrative, in which case the one modifying the head would be separated from it by the whole dependent NP including its demonstrative. Unfortunately, I am not able to provide natural examples for such situation, so the following example (26) is to be seen only as an illustration for this case, made even more obscure by the identical number and gender of the members.⁴⁷

- (26) bāb-u hāḏā -l-bayt-i hāḏā
 door-NOM DET.M.SG DEF-house-GEN DET.M.SG
 'this door of this house'

If we let the number and gender differ, a more transparent scheme (27) appears, even though rather unlikely to occur either.

- (27) bāb-u hāḏīhi -l-maktabat-i hāḏā
 door-NOM DET.FEM.SG DEF-library-GEN DET.M.SG
 'this door of this library'

An empirical question which arises is that of the real occurrence of more or less intricate *Idafa* constructions with demonstrative modifiers in spontaneous spoken MSA. Even though these constructions can easily occur in written text, the necessity to postpone the demonstrative raises concerns about the cognitive accessibility of such constructs. This would however require thorough empirical analysis of spoken data.⁴⁸

⁴⁷ As claimed in the introduction, many of the examples in this paper are artificial, that is I have created them for the purpose of illustrating certain phenomena when these were so uncontroversial that such an approach could be defended. However, this is a phenomenon which is certainly marginal, if occurring at all, and thus its relevance should be supported by empirical evidence, which I am not able to provide. Therefore, I only mention it as a syntactical option which speakers rarely resort to.

⁴⁸ Again, Magidow (2009) can be taken as a helpful model for such enterprise.

6.4 Some syntactic and semantic considerations related to the behaviour of multiple attributes

This elementary property of dealing with multiple attributes might have some interesting consequences for more theoretical syntactic considerations. In fact, from the point of view of immediate constituent analysis, the mere possibility that the head of an *Idafa* construction is modified by an adjective, is in contradiction with my claims about the 'bareness' of the head in 3.2.1, since a noun modified by an adjective is in no way a bare noun, but a noun phrase comprising multiple lexemes. However, this depends on whether we conceive of the adjective as modifying the head noun only, or as modifying the noun phrase formed by the *Idafa* construction as a whole. In the latter case, the problem is solved and indeed, very often it is highly natural to adopt this phrase modification perspective, especially if the semantic ties between the constituents of the *Idafa* construction are strong enough or even resemble composition (as treated in 3.2.4), like in (28)

- (28) [ittifāqīy-at-u]^{SC} s-salām-i l-ğadīd-at-u
 treaty-FEM-NOM DEF-peace-GEN DEF-new-FEM-NOM
 'the new peace treaty'

However, in many cases the semantics of the involved constituents and the adjective calls rather for a head-only modification perspective as in

- (29) [maktab-u]^{SC} Farīd-i-n ʔal-ğadīd-u
 kancelář-NOM DEF-Farid-GEN-n DEF-nový-NOM
 'the new office of Farid'⁴⁹

In some cases, both semantic interpretations are possible, as in (30).

- (30) [maktab-u]^{SC} l-mudīr-i -l-ğadīd-u
 office-NOM DEF-director-GEN DEF-new-NOM

In one perspective the NP can be understood as generally referring to a director's office as a type of office, without reference to a concrete director, which makes it plausible to have the adjective

⁴⁹ The fact that the affix *-n* appears at proper names like Farid though they are semantically - and here even syntactically – definite, is one of the reasons why Retsö refuses its analysis as an indefiniteness marker. This is why in this example I did not label it with the INDF gloss.

modify the whole phrase. In the other, there is a particular director who has a new office, in which case it is more plausible to view the adjectival attribute as more closely related to the head noun.

These are, of course, theoretical issues that emerge from adopting the perspective of a particular kind of syntactic analysis and do not need to be seen as relevant universally. For example, in terms of dependency syntax, this problem does not arise as both of the modifiers, one within the Idafa construction, and one adjective, are simply analysed as attributes, with no need to establish any hierarchical relation between them.

However, the semantic issues that are related to such considerations, as those of the semantic proximity shown in 3.2.4, are of undeniable importance. This concerns the issue of solving the Idafa and the competition for the dependent position, treated in 2.5.

6.5 More sophisticated elements as dependents

A syntactically intriguing phenomenon, however relatively marginal it may be, are cases in which the dependent member of an Idafa construction is not a noun phrase but a subordinated clause. It can either be a clause introduced by the subjunctors *ʔanna* or *ʔan*, or it can be headed by the subordinating pronoun *mā*. It can also be free of any introducing particle, being a mere quotation in attributive function. Each of these is commented on in its subsection of this chapter. The material which they can offer for theoretical considerations is discussed later in 6.5.3.

6.5.1 Subordinated clauses introduced by the subjunctors *ʔanna* and *ʔan*

These constructions are typically headed by abstract nouns which can have a predicative syntagm as their attribute. These are expressions like 'the idea that...', 'the principle that...', 'the assumption that...'. Very often these are deverbalised nouns, in which case the clausal attribute represents the complement of the original verb. In (31) the word *mabdaʔ* 'principle' is serving as the head for the subordinated clause introduced by *ʔanna*.

- (31) *mabdaʔ-u* *ʔanna* *l-ʔidāfat-a* *tuṭīru* *l-ǧadal-a*
principle-NOM that DEF-Idafa-ACC provokes DEF-debate-ACC⁵⁰

⁵⁰ The word *l-ʔidāfat-a* is in accusative because the subjunctive *ʔanna* always requires the subject of the subordinate clause to be in accusative.

'the principle that Idafa provokes debate'

Interestingly, not every construction of this kind behaves as an Idafa. For example the frequently appearing verbal noun *qawl* derived from *qāla* 'to say', thus meaning 'claim', 'utterance', rarely has the subordinate clause as its dependent, and it either carries the definite article, as in (32a) or has a dependent referring to the subject, i.e. the source of the utterance, as in (32b). This appears to be an example of the competition of multiple dependents, even though one would expect that at least in cases where the subject of the verb is not mentioned, an Idafa should be formed with the utterance as its dependent. This, however, is not the case.⁵¹

- (32) a. ²*al-qawl-u* ²*inna* *l-²iḏāfat-a* *tuṭīru* *l-ḡadal-a*
 claim-NOM that DEF-Idafa-ACC provokes DEF-debate-ACC
 'the claim that Idafa provokes debate'
- b. *qawl-u* *l-mudīr-i* ²*inna* *l-²iḏāfat-a*
 claim-NOM DEF-director-GEN that DEF-Idafa-ACC
 tuṭīru *l-ḡadal-a*
 provokes DEF-debate-ACC
 'the director's claim that Idafa provokes debate'

Unlike many other phenomena which are worth a quantitative analysis, this construction could be researched by means of corpora as its marking features are partly captured by the Arabic script. Namely, the absence of the definite article at the head noun implies that the author intended the construction with the subordinate clause as an Idafa construction. It would thus be possible to find out how often different speakers writing in Arabic neglect the Idafa in such cases by using the definite article with the head noun and whether the reason might be that the dependent slot is occupied by something else.

⁵¹ One could speculate that the behaviour of *qawl* could be special since this verb prescriptively takes a different subordinating particle, ²*inna*, whereas the majority of other verbs take the default ²*anna*. However, this is of no relevance for the contemporary use of Arabic since these particles are often interchanged.

6.5.2 Specifying expressions introduced by the subordinating pronoun *mā*

These are constructions in which the dependent member specifies a general temporal or local expression. In many cases these are frequent collocations, like in (33), but the construction is productive, like in (34).

(33) [*ʔaʃr-u*]^{SC} *mā* *qabla* *-t-tārīkh-i*
 age-NOM what before DEF-history-GEN
 'prehistory'

(34) [*fatrat-u*] *mā* *baʕda* *-l-ʔinqilāb-i*
 period-NOM what after DEF-coup-GEN
 'the period after the coup'

Syntactically, the dependent member is a nominal clause with the pronoun *mā* as its subject and the given prepositional phrase as its predicate. Thus, the examples would literally be translated as 'the age of what is before history' and 'the period of what is after the revolution'.

These specifying expressions can likewise be seen as rather peripheral but they can be useful for minute observations of the actual marking behaviour and tendencies in these peripheral cases (as will be debated in 6.5.4. This is due to the feminine nouns which in LSA mark the construct form by the final *-t*, as treated in 5.2.2. Thanks to this overt marking, constructions headed by feminine nouns can easily whether the speaker uses an expected *Idafa* or resorts to mere apposition. The two possibilities are given in (35) . In the second case the head *fatra* would be expected to carry the definite article as in (35c), but I have encountered the case shown in (35b) which points to the weakening of the *Idafa* strategy, but is not made definite by the article either.

(35) a. *fī* *fatra-t* *mā* *baʕda* *l-ʔinqilāb*
 in period-PERT what after DEF-coup
 'in the period after the coup'

b. *fī* *fatra* *mā* *baʕda* *l-ʔinqilāb*
 in period what after DEF-revolution
 'in the period after the coup'

- c. *fi l-fatra mā ba^ʿda l-²inqilāb*
 in period what after DEF-revolution
 'in the period after the coup'

Unfortunately, I am not able to provide any reliable data for this phenomenon, since it requires spoken data which would either have to be searched for examples of these constructions, or they would have to be carefully elicited.

Indeed, this is one of the phenomena which are not captured by the Arabic orthography since the feminine affix *a(t)* is written by a special grapheme which gives no clue whether the affix is reduced or not. Therefore, it is a phenomenon observable only in spoken Arabic and it can be argued that speakers who only use MSA for writing do not necessarily have to make any decision concerning the problem of the reduction of the final *-t*.

6.5.3 Sentence-like expressions with no subjunctor

These are usually direct quotations serving as attributes. The head nouns are similar to those in 6.5.1. The following three examples are taken from El-Ayoubi 493, including the quotation marks.⁵² While (36) is a prototypical abstract noun, in (37) and (38) the expressions ^ʿ*alā a²sās-i* and *min bāb-i* can be already considered as partly lexicalised into prepositions.

- (36) *mabda²-u* “*lā šay²un ya²tī min lā šay²in*”
 principle-NOM “nothing comes from nothing”
 'the principle of “nothing comes from nothing”'
 (the principle saying “nothing comes from nothing”)

- (37) ^ʿ*alā* *a²sās-i* “*ḥuḍ-hu ^ʾaw ^ʾutruk-hu*”
 on basis-GEN “take it or leave it”
 'on principle “take it or leave it”

- (38) *min* *bāb-i* “*kullu šay²in mumkinun*”
 from door-GEN “everything is possible”
 'based on (the principle that) “everything is possible”'

⁵² In El-Ayoubi examples are in the Arabic script, but even there quotation marks are used because apparently they were used in the original sources which are Arabic newspapers..

Most prototypical for this construction is probably (37) since the imperative clause which serves as dependent cannot otherwise be syntactically incorporated into another construction (except, of course, quoting constructions like 'He said: "take it or leave it"'). To the contrary the dependents in 36 and 38 have the form of statements and could also be introduced by *ʔanna* and yield constructions of the type presented in 6.5.1.

6.5.4 Salience of subordinate clauses in the dependent position

Seen from a comprehensive perspective on *Idafa* and especially on its head-marking behaviour, these constructions are peripheral in some sense, because there does not appear to be a semantic reason why a subordinated clause should necessarily be coded as a dependent of an *Idafa*, and indeed, this is not always the case. From a diachronic point of view, it seems probable that this use of *Idafa* is rather secondary, emerging as a projection of a syntactic construction established primarily as a conjunction of two nouns (see 17.2). This raises the question to which extent the *Idafa* exclusively serves for coding these constructions and whether users of Arabic tend to neglect it sometimes. I consider this question relevant because in contrast to the prototypical noun-noun attributive constructions which embody *Idafa*, in these peripheral cases there is less need for actually marking the relation between the head and the dependent. The reasons are that it is either marked already by the subordinating conjunction or the relation can be expected to go unmarked on the basis of cross-linguistic considerations. For example, in many languages which have a dependent-marking construction, the genitive inflection cannot apply to entire clauses and thus the genitive construction is ruled out in these cases and the subordinate clauses are only marked by subjunctors or the attributive relation is rendered by mere apposition.⁵³

Moreover, in such constructions it is more likely than in prototypical noun-noun possessive constructions that more elements are going to compete for the available attributive positions. These can typically be complements of the underlying verbs if the head is a deverbal noun or NPs with other semantic roles related to the quotation, like in *John's idea to go*. Once the dependent position

⁵³ In English, a quotation in attributive position often preferably occurs before the head noun, like in *the "take it or leave it" principle* which is preferred over the opposite word order *the principle "take it or leave it"*. This can be interpreted as a behaviour resembling that of typical English appositions, in which the head is preceded by the modifier. Therefore, it shows a tendency to treat more complicated or vague attributes in the same way as more prototypical ones.

is unique as is the case in *Idafa*, such elements appear to be more salient and apt for acquiring it than the rather vague clausal attribute.

6.6 Alternative constructions and multiple dependent

The rich spectre of semantic functions expressed by *Idafa* can also be conveyed by alternative constructions.

6.6.1 The relational adjective

Apart from apposition which is not going to be treated here, an obvious alternative to *Idafa* is using the highly productive relational adjective formed by the affix *-ī*. An *idafa* construction and its equivalent formed in this way is shown in

- (39) a. *muškil-at-u* *s-siyāsat-i*
 problem-FEM-NOM DEF-politics-FEM-GEN
 'the problem of politics'⁵⁴
- b. *ʔal-muškil-at-u* *s-siyās-īy-at-u*
 problem-FEM-NOM DEF-politics-ADJ-FEM-NOM
 'the political problem'⁵⁵

However, thus formed adjectival attributes are not always equivalent due to the not completely predictable semantics of the derived relational adjective, as seen in

- (40) a. *ʕašāʔ-u* *l-ʕamal-i*
 dinner-NOM DEF-work-GEN
 'business dinner'
- b. *ʔal-ʕašāʔ-u* *l-ʕamal-īy-u*

⁵⁴ To avoid any confusion, it should be emphasised that there is no gender agreement in this example. Only by accident both of the head and the dependent are feminine nouns.

⁵⁵ Here, indeed, there is gender agreement between the head and the adjectival modifier.

DEF-dinner-NOM DEF-work-ADJ-GEN
 'practical dinner'

6.6.2 Analytical substitution of the Idafa

The most efficient way of joining two noun phrases while avoiding Idafa is connecting them by a preposition, that is adjoining the second noun phrase as an attribute within a prepositional phrase. The most common preposition to serve this purpose is *li*, which is a very universal preposition with dative, directional or benefactive meaning. Other prepositions with locative or ablative meaning are also used. (41a-b) shows an Idafa construction modified by an adjective (repeated from (18b)) and its equivalent yielded by using the preposition *li*. Such alternative makes it possible to switch the order of the two attributes.

- (41) a. [*maktab-u*]^{SC} *l-mudīr-i* *l-ġadīd-u*
 office-NOM DEF-director-GEN DEF-new-NOM
 'the new office of the director'
- b. ²*al-maktab-u* *l-ġadīd-u* *li* *l-mudīr-i*
 DEF-office-NOM DEF-new-NOM for DEF-director

The motivation for using such analytical constructions is achieving higher syntactic flexibility. This is relevant especially when there are multiple attributes that need to be coded, as in (41). Some problems related to multiple attributes have already been raised, both for attributes on one syntactic level thus modifying one head and chains of embedded attributes (see 6.1-6.4).

6.6.3 NPs competing for the position of the dependent member

If we think of the analytical constructions presented in the previous subsection as substituting for corresponding Idafa constructions, it entails the assumption that the Idafa is primary in some sense. In the case of solution by prepositions this is justified because of a great number of expressions which occur on their own an Idafa constructions and are solved once an additional dependent

appears. (42a) shows the compound-like collocation meaning 'news agency'. When a proper name is added, the collocation is “dissolved” and it is the proper noun which occupies the dependent position, while the original dependent is shifted to the prepositional phrase, as seen in (42b). This is preferred over (42c) which treats the two attributes inversely.

- (42) a. wikālat-u l-ʔanbāʔ-i
 agency-NOM DEF-news-GEN
 'news agency'
- b. wikālat-u Reuters li l-ʔanbāʔ-i
 agency-NOM Reuters for DEF-news-GEN
 'Reuters news agency'
- a. wikālat-u l-ʔanbāʔ-i li Reuters
 agency-NOM DEF-news-GEN for Reuters
 'the news agency of Reuters'

This is a rather intriguing feature since the semantic connection between the members of the quasi-compound is certainly closer than between the head noun and the proper noun. I am not able to provide an explanation of this. It would be worth finding a way how to conduct some quantitative research on spoken data to find out to which extent this strategy of dealing with multiple attributes is really used in spontaneous speech as it is not excluded that it is mainly occurring in set phrases like the names of institutions.

Another issue is that of a chain of embedded attributes. In such case, there are multiple possibilities of how the chain is interrupted, i.e. which segment of it is shifted to the prepositional phrase and which remains coded by *Idafa*. (43) is an extension of (16) and (17), showing two possible ways of portioning the original *Idafa* chain.

- (43) a. [maktab-u]^{SC} [mudīr-i]^{SC} -l-bank-i
 office-NOM director-GEN DEF-bank-GEN
 'the office of the director of the bank'
- b. ʔal-maktab-u li [mudīr-i]^{SC} l-bank-i

office-NOM for director-GEN DEF-bank-GEN
'the office of the bank's director'

- c. *[maktab-u]^{SC}* *l-mudīr-i* *li* *l-bank-i*
office-NOM DEF-director-GEN for DEF-bank-GEN
'the director's office of the bank'

The English translation of the two options 43a-b tries to convey the semantic difference between them. However, the factors which influence speakers' preference of one of them in spontaneous speech remain to be searched for by means of empirical research.

7 DIACHRONIC AND CROSS-LINGUISTIC PERSPECTIVE

This chapter is intended to provide a brief outline of some ideas about the origin of Idafa and also set it into the context of some similar constructions found in other Semitic languages. Such a comparative and diachronic perspective is necessary if one strives for better understanding of the way Idafa is marked.

7.1 The concept of Idafa in Semitic

A construction resembling Idafa, both formally and functionally, can be observed throughout the Semitic family. However, it is not obvious what characteristics should be used for the definition of such construction. In his treatise on the parallel behaviour of different types of nominal attribution in Semitic, i. e. modification by a noun phrase, an adjective and a relative clause, Retsö (2009) uses the distinction between *annexation* and *juxtaposition*. The latter is roughly equivalent to what is usually termed apposition, being characterised by there being “no formal indicator of the link between the constituents; the modifier is simply placed after the head as a comment.” On the contrary, *annexation* “denotes a tighter connection between the constituents with a clear morphological and/or syntactical marking of the link.” (Retsö 2009: 4) In this setting, the Idafa represents annexation while adjectival attribution represents juxtaposition. This is certainly adequate in the synchronic perspective. However, if a diachronic insight is attempted, another article by Retsö is thought-provoking. In Retsö (1993) he righteously challenges the traditional view of Arabists on the fate of Idafa, namely its substitution by prepositional phrases in Arabic dialects, as a transition from a synthetic/morphological marking to analytical/syntactic marking. The following passage is worth quoting:

“The conceptual confusion among Arabists on this point becomes apparent when one realises that also the ²idāfa construction in fact is an analytical construction like those with *dī*, *taba^s* etc. since the ²idāfa consists of syntactic ordering of at least two base-morphemes. The contrast between HA and the vernaculars in this field is not at all the one between synthetic morphological marking and analytical syntactic marking, at least not according to the common linguistic definition of the terms. The main device for marking the genitive construction in Arabic as well as in Semitic as a whole consists of different forms of syntactic grouping of free morphemes. Some languages like e.g.

Akkadian, Ugaritic and HA have an additional marking by a bound morpheme which is actually redundant.'⁵⁶

(Retsö 1994: 337)

Indeed, by the redundant bound morpheme the genitive ending is meant. Thus, this dependent-marking feature is certainly not going to be included in a comprehensive definition. On the other hand, Retsö points out that even the head-marking features, which are so crucial for Arabic, are not uniform across the Semitic family and differ as to the degree to which they are grammaticalised.

7.2 Grammaticalisation pathways for the Idafa

As we have seen, though Idafa is not to be viewed as mere apposition synchronically, its origin may certainly have to do with simply putting the two nouns next to each other. A possible grammaticalisation path that might have led to such a construction is that of an apposition of two nouns becoming grammaticalised through phonological means. As Gesenius (1881) briefly notes in his description of Biblical Hebrew, the stress is shifted to the second constituent which brings about some additional changes to the first constituent, i. e. the head. Generally, this means that the head noun is not in an independent form anymore but in a special reduced form, called *status constructus*. In Hebrew, the construct form is often morphologically different from the absolute form, due to the phonological and prosodic rules that apply once the word is immediately followed by its dependent. An example can be *dāḇār* 'a word', opposed to *dāḇar ʔelohīm* 'the word of God'.

In Arabic, the form of the given lexeme serving as head remains unchanged and the grammaticalised attributive relation is only marked by the absence of the definite article, as treated in 5.2. The several cases in which there is a difference in the form of the head noun itself, mentioned in 5.2.2, can indeed be seen as consequences of the phonological and prosodic process of convergence of the two nouns. This is true for the feminines within LSA, even though in that case it is not a reduction of the head that has been brought about by the annexation of the dependent, but rather it is the retention of the final *-t* which would otherwise have been dropped due to an universal apocope operating in LSA. This is not an unnatural result since the apocope can easily be imagined not to apply in the case of a sequence of closely connected nouns. In the case of the dual and animate plural endings, an explanation in terms of the reduction due to the annexation of another

⁵⁶ HA is the abbreviation for *Hocharabisch*, that is Classical Arabic in the terminology of this paper.

noun seems possible as a truncation of the full ending, but the diachrony of these endings is probably more complicated and its analysis is beyond the scope of this paper.⁵⁷ The case of the five-noun class is special because the five lexemes contained in it have a long final vowel in the construct form, opposed to the short vowel occurring otherwise. This might be due to the shortness of these lexemes and a need to connect them with the following word. Or it can be a retention of an otherwise reduced form.

7.3 The degree of grammaticalisation

The fact that Arabic leaves the forms themselves untouched can be considered a conservative feature. The changes occurring in Hebrew are innovations which make the established syntactic construction more easily observable and arise, similarly to the feminines in LSA, inquiries about a proper analysis of its morphonology. Gesenius explains his view in the closing paragraph of his introduction to the issue as follows:

'Uebrigens aber geht aus Obigem zur Genüge hervor, dass der Stat. Constr. strenggenommen nicht als eine syntaktisch-logische, sondern vielmehr zunächst als eine rein lautliche, rhythmische (durch die Tonverhältnisse bedingte) Erscheinung zu betrachten ist.'

(Gesenius 1881: 197)

I think that this formulation probably reflects to some extent a slight Indo-European bias for it appears to deny a morphosyntactic status to the construction. Whether this is justified or not depends on the extent to which we believe that such a connection of two nouns is grammaticalised. There are cases which agree with Gesenius' scepticism as to the grammatical autonomy of the construction, namely the use of the construct forms even in other contexts than immediately preceding a following noun, like in (44) where it appears before a prepositional phrase or in (45) where it appears before an adjectival modifier.

- (44) [Šimḥa-t]^{SC} *ba-qqāsīr*
 joy-PERT at.DET-harvest
 'the joy of the harvest' (Gesenius 1881: 261)

⁵⁷ See Retsö 1997

- (45) Kneset ha-gedolā
 congregation DEF-great.F
 'the Great Congregation' (Retsö 2009: 15)

An inversed type of counterexamples to challenge the degree grammaticalisation degree of the noun-noun construction is that of the head noun being marked by the definite article. Even though such cases do occur, for Biblical Hebrew Gesenius (1881: 253) himself notes that mostly these are only apparent and are due to incorrect analysis of the data. On the other hand, such examples occur in other Semitic languages. Ouhalla (2009) quotes example (46) from Spanish Arabic:

- (46) fī al-arḍ al-qāḍī al-maḍkūra
 in DEF-land DEF-judge DEF-said
 'in the said land of the judge'

In the case of Spanish Arabic, the absence of a synthetic Idafa construction might be an influence from Spanish, as Ouhalla argues. Nonetheless, structurally it shows that the range of options how an Idafa-like construction can be shaped in the realm of Semitic languages.

However, in the case of Arabic there is no doubt about a high degree of grammaticalisation, if only due to the obligatory absence of the definite article. Moreover, other evidence can be presented for the autonomy of the Idafa as it is marked in Arabic. It is the occurrence of cases in which the head member is not a single bare noun but a coordination of multiple nouns as shown in 3.2.2. In these cases the construct form, i. e. one without the definite article, is used even for the non-final constituents, that is in phonological environment which wouldn't require a construct form on its own. Therefore, we see that we are dealing with a specific morphosyntactic construction and not a mere phonologically conditioned morphonological variation. A different rather cross-linguistic piece of evidence is mentioned in 7.5.

7.4 Marking or allomorphic variation?

The problem of the autonomy of Idafa as a construction is also connected to the interpretation of the notion of state as outlined in 4.2. While Retsö (1997) argues that no semantic load should be associated with the different states, he goes even further in seeing this category as mere allomorphic variation, conditioned by the syntactic position of the word. In the case of CA, his conclusion is

“dass der Status im Arabischen als eine allomorphische Variation der nominalen Flexionsmorpheme definiert werden kann, nicht ungleich der Variation zwischen starker und schwacher Adjektivflexion in den germanischen Sprachen” (Retsö (1997: 69)).

Undoubtedly, Retsö is right in that structurally the phenomenon can indeed be viewed as allomorphic variation and indeed the above-mentioned scenario of its grammaticalisation shows how such variation could have come about. However, I believe that such a description challenges the syntactic importance of the construction synchronically, or more precisely, does not allow for reversing the causality of the observed phenomena in the sense that if a position-conditioned variation occurs it cannot end up being interpreted as actively marking the different syntactic situations which are given by the positions. To put it more concretely, if we view the different states as mere allomorphic variants, we cannot easily say that e.g. the construct state marks the present attributive relation as it is rather interpreted as a consequence of the presence of this relation. However, as the previous sections show, I tend to view the processes which happen to the head, turning it into the construct form, as ordinary head-marking of the relation. If we consider the comparison with cases of allomorphic variation, such as the variation of strong and adjectives in the Germanic languages, we see that the differing inflection of the respective adjective classes does not point to a different syntactic construction being used. To the contrary, the occurrence of the construct form signals the presence of a nominal attribute. Therefore, I believe that it is reasonable to assign to the variation a grammatical function.

I should add that the conflict between Retsö's emphasis on the allomorphic nature of the states, and the preceding ideas of mine is in fact much less significant than it may seem, since in Retsö's definition. the *state* refers only to the very form of the lexeme which occurs in a specific environment, while the presence or absence of the article is not a part of the given state, but constitutes one of the environments. Thus, there is no reason why the absence of the article could not be a marker of the relation. And indeed, as seen in section 5.2, the absence of the definite article is the essential means of marking the relation. Some controversy between the view proposing allomorphic variation and my proposal of an active syntactic function can only occur in case of the plural and dual ending , in which the absolute and construct forms overtly differ and in the case of Hebrew where the phonological form of the lexemes changes. This, of course, is a matter of theoretical analysis and probably both of the approaches can lead to adequate description of the data.

7.5 Head-marking pattern borrowed

The problem of the nature of the marking pattern observed in *Idafa* and its comparison to more prototypical head-marking patterns in which there is a pertensive affix, can be elucidated by the pattern which has evolved in Khuzestani Arabic, as outlined by Shabibi. This variety of Arabic spoken in Iran has been massively influenced by Persian, including its pattern of marking nominal attribution. In Persian there is one way of marking attribution for both adjectival and nominal attributes. It is head-marked by the *-e* affix attached to the head, as seen in (47).⁵⁸

- (47) a. *pesār-e* *modīr*
 boy-PERT chief
 'the chief's son'
- b. *pesār-e* *jāvān*
 boy-PERT young
 'the young boy'

In Khuzestani Arabic, this unified pattern has been projected into the Arabic morphology in such a way that also adjectival attribution is signaled by the absence of the definite article *ʔal-*, as seen in (48).

- (48) *pardāt* *l-əmlawwən-āt*
 curtains DET-colourful-PL
 'colourful curtains'

This can be counted as evidence for claiming that the Arabic marking pattern is indeed a head-marking one and can cross-linguistically be compared to more standard patterns with overt pertensive marking like that of Persian.

⁵⁸ This marker is traditional *_y* named *ezafet*, which is a terminological borrowing from Arabic, being nothing else than *Idafa*.

8 CONCLUSION

In the preceding chapters several issues related to the marking of *Idafa* have been raised in the perspective of both the classical version of the language and the changes that occur in usual spoken use of the standard language.

The morphological aspects have been studied and it has been argued that there are in reality no traits of dependent-marking. This might sound a little surprising, given that the concept of the genitive occurs in any normative grammar of Arabic. However, the simple reason for the absence of dependent-marking is the fact claimed by Magidow (2009) that the so called genitive has become void of any syntactic function and in the spoken version of MSA is dropped.

On the other hand, the head-marking features are less conspicuous but their importance in the behaviour of the construction is obvious. The important theoretical question which has been debated is that of a proper description and interpretation of the morphological pattern of *Idafa*. While it is possible to treat the absolute and construct forms of nouns as allomorphic variation, such view appears not to acknowledge the grammaticalised syntactic function of the construct form. I believe that there are reasons for treating the changes applying to the head nouns as head-marking features, even though the marking is not carried out by an independent affix, but rather by certain syntactic restrictions imposed on the head. A nice piece of possible evidence for the autonomy of such head-marking pattern is provided by Khuzestani Arabic, which has borrowed the structure of the Persian head-marking pattern using exactly these Arabic head-marking features as are found in *Idafa*, i.e. the absence of the article.

References

- Badawi, Elsaid, M. G. Carter, Adrian Gully. *Modern Written Arabic. A Comprehensive Grammar*. London - New York: Routledge. Taylor & Francis Group, 2004.
- Badawi, El-Said. *Mustawayat al-arabiyya l-mu'assira fi misr*. Cairo: dar al-ma'arif. 1973.
- Brockelmann, Carl. *Grundriss der vergleichenden Grammatik der semitischen Sprachen. II. Band: Syntax*. Berlin: Verlag von Reuther et Reichard, 1913.
- Creissels, Denis. 2009. Construct forms of nouns in African languages. In K. Austin, Oliver Bond, Monik Charette, David Nathan and Peter Sels (ed.), *Proceedings of Conference on Language documentation and linguistic theory 2*. 73-82. London: SOAS.
- Diem, Werner. „Alienable und inalienable Possession im Semitischen.“ *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 136 (1986), s. 228-291.
- Dixon, R. M. W. 2009. *Basic Linguistic Theory, Volume 2*. Oxford: Oxford University Press.
- El-Aoubi, Wolfdietrich Fischer, Michael Langer. *Syntax der Arabischen Schriftsprache der Gegenwart: Teil I, Band I. Das Nomen und Sein Umfeld*. Wiesbaden: Reichert Verlag, 2011.
- Ferguson, Charles. 1959. Diglossia. *Word* 15: 325–340.
- Fischer, Wolfdietrich. *Grammatik des Klassischen Arabisch*. Wiesbaden, 1972.
- Gesenius, Wilhelm, Emil Kautzsch, Emil Rödiger. *Gesenius' hebräische Grammatik*. 23. Auflage. Leipzig: Verlag von F. C. W. Vogel, 1881.
- Heine, Bernd. 1997. *Possession: Cognitive Sources, Forces and Grammaticalization*. Cambridge: Cambridge University Press.
- Hoyt, Frederick. „Noun Phrase.“ In Versteegh, Kees (ed.). *Encyclopedia of Arabic language and linguistics. Vol. III*. Leiden: Brill, 2007, s. 428-434.
- Ivanova, Valerija, Adam Pospíšil. Užívání koncovek v mluvené spisovné arabštině. Seminární práce ke kursu Sociolingvistika. Nепublikováno, 2013.
- Luschützky, Hans Christian. *Uvedení do typologie jazyků*. Praha: FF UK, 2003.
- Magidow, Alexander. 2009. *The Semantic and Pragmatic Role of Case Marking in Formal Spoken Arabic*. Austin: University of Texas (Master Thesis.)
- Nichols, Johanna. „Head-Marking and Dependent-Marking Grammar“. *Language* 62.1 (1986), s. 56-119.
- Owens, Jonathan. 2001. *Arabic Sociolinguistics*. Arabica 48, 419-469

Ouhalla, Jamal. 2009. Variation in Genitive Noun Phrases: Moroccan Arabic, Shamaliya Arabic and Spanish Arabic. In Watson, Janet & Retsö, Jan (ed.), *Relative clauses and genitive constructions in Semitic*. 195-217.

Retsö, Jan. 1984-1986. State, Determination, and Definiteness in Arabic. A Reconsideration. *Orientalia Suecana* XXXIII-XXXV. 341-346.

Retsö, Jan. 1994. 'rāb in the forebears of modern Arabic dialects. In Dominique Caubet and Martine Vanhove (ed.), *Actes des premières journées internationales de dialectologie arabe de Paris*. 333-343. Paris: INALCO.

Retsö, Jan. 2009. Nominal Attribution in Semitic. Typology and Diachrony. In Watson, Janet & Retsö, Jan (ed.) 2009, *Relative clauses and genitive constructions in Semitic*. 3-34.

Watson, Janet & Retsö, Jan (eds.) 2009, *Relative clauses and genitive constructions in Semitic*. Journal of *Semitic Studies*. Supplement 25. OUP