THE COMPOSITE NATURE OF *llaḏī, ḥayṯu, man/mā-min* DISTRIBUTIONAL, TYPOLOGICAL AND PRAGMATIC ASPECTS OF OLD ARABIC RELATIVE MARKERS

Francesco Grande

Università Ca' Foscari, Venezia

Abstract

This paper offers a distributional account of the morphology and semantics of the Old Arabic relativizers *llaḏī, ḥayṯu, man/mā-min*, showing that they constitute a multi-layered pattern of complementary distribution, based on the semantic opposition {-/+restrictive}, and on its audible counterpart. Far from being a suprasegmental opposition {+/-pause} (cp. English), the latter opposes *lla, ḥay* to *min*, which are analyzed accordingly as replacive morphemes of {+/-pause}. Such a replacive allomorphy is also given a pragmatic characterization, and explained as an adaptive behavior of Old Arabic, relative to its oral-poetic ecological conditions. Deviations from the aforesaid pattern are explained here by invoking typological factors such as heaviness and Jespersen cycle which, in turn, are triggered by an instance of phonological reduction described by Arab Grammarians, and targeting *min*. This phenomenon is arguably part and parcel of a more general shift from analytical to synthetical language, well-known to typologically-oriented studies on (Old) Arabic.

0. Introduction and aim

This paper is designed to contribute to a better understanding of the morphology and semantics of the Arabic relative markers *llaḏī*, *ḥayṯu*, and *man/mā*[1] (which in its purely relative function co-occurs with *min*)[2]. The paper takes as its starting point

[1] This paper is based on some materials presented at the 4th Congress of the Coordinamento dei Dottorati Italiani in Scienze Cognitive (Rome, June 7-9, 2010) and the 26th Congress of the Union Européenne des Arabisants et Islamisants (Basel, September 12-16, 2012). It ideally complements the research on Old Arabic relative markers and clauses carried out in Grande (2013: Chs. 2, 3), but can be read independently of it. Abbreviations include the following: ADP = Adposition, C = consonant, F = feminine, N = Noun, NP = Noun Phrase, OA =Old Arabic, P = Preposition, PP = Prepositional Phrase, RC = Relative Clause, V = Vowel, # = word-boundary, – = morpheme-boundary.

[2] See end of Sect. 3 for details.

THE ARABIST. BUDAPEST STUDIES IN ARABIC 32 (2013)
https://doi.org/10.58513/ARABIST.2013.32.3

the idea often found in the literature that *llaḏī* and *ḥayṯu* are complex linguistic entities in which a relative stem is expanded through addition of some morphemic material, and aims at presenting a critical discussion and improvement of it.

By way of illustration, Brockelmann (1910:123) claims that *llaḏī* can be divided into the smaller meaning units *l, la, ḏī* – an article, a reinforcer and a relative stem, respectively: "Dans toutes les langues sémitiques les relatifs provennient des démonstratifs. […] en arabe classique on se sert de la forme renforcée par la et l'article".

Nonetheless, even if one were to concede that the quite vague notion of 'reinforcement' appropriately describes the morpheme *la*, this still would not explain *why* the piece of information in question is associated with the relative stem *ḏī*, so further research is needed on this topic.

The discussion is organized as follows. Section 1 sets up the historical and geographical context of this study, whereas the data relevant to it are presented in Sections 2 and 3. Sections 4 and 5 introduce and develop a recent semantic analysis of the OA relative markers. Section 6 deals with some phonological and syntactic issues raised by such a semantic analysis, and provides a solution to them, which results, to a good extent, in a satisfactory multi-layered pattern of complementary distribution for the OA relativizers. Section 7 presents some observations concerning the morphology and pragmatics of the OA relativizers, in order to refine the pattern at issue. Finally, Section 8 further refines the latter, proposes a functional rationale for its emergence, and offers the main conclusions of this study.

1. Historical and geographical context

The investigation carried out here on the morphology and semantics of the Arabic relative markers focuses on pre-Classical Arabic or, according to a different terminology, Old Arabic. In the wake of Owens (2006:63-64, 88, 198-199) this is a stage of language documented from about 300 CE to 800 CE, and located in the Arabian Peninsula. Under this definition, the *terminus a quo* of Old Arabic is the so-called Nemara inscription (328 CE), its *terminus ad quem* is represented by the collection of forms, words, constructions and sentences/utterances reported in the written sources stretching as far as the IX century CE, among them:

- the Koran (*al-Qur'ān*) and especially its readings (*qirā'āt*) known in Islamic scholarly tradition as the 'ten readings' (i.e. non-aberrant);
- the grammatical treatise *Kitāb* by Sībawayhi (d. 177/793) and works dating back to the same period, as well as late grammatical treatises and dictionaries, which repeat and take extracts from Sībawayhi's work, etc.

From a qualitative viewpoint, these two groups of sources differ in a fundamental respect, namely as to the kind of philological evidence found in them, which can be either internal or external (Monroe 1972:6 *et passim*). The former group only consists of internal evidence – the Koran and its readings representing in themselves a corpus of Old Arabic data –, while the latter also includes external evidence: as well known, in addition to gathering *linguistic* data from Old Arabic, Sībawayhi's work *et similia* furnish these latter sources with a *metalinguistic* description.

This being said, creating a distinction between primary sources in terms of internal and external evidence raises the problem of the authenticity of the internal evidence presented by Sībawayhi etc., who often cite lines of pre-Islamic poetry to exemplify a given linguistic fact of Old Arabic. In this case, the *vexata quaestio* on the authenticity of pre-Islamic poetry, however, has no bearing on the reliability of the Old Arabic data themselves because, as Rabin (1951:15) sharply points out: "The verses [...] used by the philologists themselves, are meant to serve as examples, not as evidence for the existence of the phenomenon they illustrate. For that reason in most cases it does not matter much whether they are genuine or not."

On the whole, the linguistic materials attested during this period qualify as an (idealized) uniform stage of language, since the lexicon and syntax used from the Nemara inscription onward do not differ noticeably from those used in the written sources of 750-800 CE. For instance, both share the demonstrative *tī* 'this.F', and the relative marker *ḏū*, while the same features are not found in Classical Arabic (Versteegh 1997:32). Hence, similar features define these linguistic materials as Old Arabic (OA henceforth) and set it apart from Classical Arabic.

2. The data: non-systemic relative markers

OA was fairly rich in relative markers, but as far as we can tell from the extant sources, they were used with different degrees of frequency by the speakers.

A group of relative markers (*ḏī, ḏū, llāy*, *llā'i*) is found, whose use appears to be confined to particular geographical or tribal areas, and only exceptionally do they enter the supra-dialectal language (*Kunstsprache*). This can be plausibly inferred from the external and internal evidence offered by the OA sources mentioned in the previous section. To begin with, Early Arab Grammarians (*apud* Rabin 1951:39, 205) classify the relativizers *ḏī, ḏū, llāy*, *llā'i* as diatopic variants of *llaḏī*, stating that *ḏī* was heard only in Yaman and Ḥiğāz, *ḏū* among the Ṭay' tribe, *llāy* among the Qurayš tribe[3]. This external evidence is confirmed by the internal evidence found in the Koran, which makes no or very low usage of such relative markers:

[3] See also fn. 8 below on *ḏī*.

according to *The Quranic Arabic Corpus*[4], the relativizers *ḏī, ḏū, llāy* never occur in it[5], whereas *llā'i* does four times only (*Koran* XXXIII, 4; LVIII, 2; LXV, 4)[6,7].

Moreover, if Pre-Islamic poetry may be credited as being reliable, it would confirm this linguistic scenario: according to Monroe (1972:9, 14), a dialect form exceptionally enters the OA (poetic) *Kunstsprache* for metrical reasons – if and only if it allows the formula it is part of to fit the meter. A telling example is the relative construction attested in a line of poetry cited in *Šarḥ al-Kāfiya* (III, 23), i.e. *l-mar'i ḏū ǧā'a sāᶜiyan* 'the man who came in a hurry', where the relative marker *ḏū*, otherwise infrequent in the *Kunstsprache*, is called for by the meter and cannot be metrically replaced by *llaḏī* or *man…min*.

In sum, the evidence found in primary sources arguably shows that the relativizers *ḏī, ḏū, llāy*, *llā'i* were not frequent in OA. In one sense, the lack of frequency of a given linguistic entity is testimony to its being an element not fully integrated within a linguistic system: if, in fact, it can be hardly denied that a linguistic system is a network of oppositions (cp. the Saussurean axiom of differentiality, as per Saussure 1916:163-165), it is also fairly uncontroversial that these oppositions, in order to be detectable, must occur *frequently* in the linguistic system they belong to, as argued by Saussure (1916:215): "Dans deux mots tels que *maison : ménage*, on est peu tenté de chercher ce qui fait la différence des termes, soit parce que les éléments différentiels (*-ezõ* et *-en-*) se prêtent mal à la comparaison, soit parce qu'aucun autre couple ne présente une opposition parallèle. Mais il arrive souvent que les deux termes voisins ne diffèrent que par un ou deux éléments faciles à dégager, et que cette même différence se répète régulièrement dans une série de couples parallèles".

That high frequency is a clue to systemicity of linguistic entities also emerges from some of Saussure's (1916:138, 155) formulations of his axiom of *langue*, whose constitutive elements are at once endowed with high frequency ("cette forme, souvent répétée, et acceptée par la communauté, est devenue un fait de langue") and systemicity ("la langue ne peut être qu'un système de valeurs pures").

---

[4] An online annotated linguistic resource developed by Kais Dukes, School of Computing, University of Leeds: http://corpus.quran.com/

[5] Whereas *ḏū* occurs as a possessive marker in the Koran: *ḏū l-faḍli* (e.g. *Koran* II, 105; III, 74; VIII, 29; LVII, 21) etc.

[6] http://corpus.quran.com/search.jsp?q=lem%3A%7Bl~a%60%5E_%23iY

[7] If one were to follow the accepted readings by Abū ᶜAmr ibn al-ᶜAlā' (d. 154/770) and Ibn Kaṯīr al-Makkī (d. 120/737, as transmitted by al-Bazzī, d. 250/864), who read *llāy* instead of *llā'i*, the situation would look slightly different: *ḏī, ḏū, llā'i* never occur in the Koran, whereas *llāy* does, four times. This does not affect the main point that these relative markers as a whole are rarely or not used at all in the Koran. See Rabin (1951:154) for details.

It seems therefore reasonable to state that the relativizers *ḏī, ḏū, llāy*, *llā'i* did not enter the system of OA relativizers due to their low frequency.

3. The data: systemic relative markers

In contrast to *ḏī, ḏū, llāy*, *llā'i*, the relative markers *llaḏī*, *ḥayṯu*, and *man/mā* appear to be quite frequent in the OA sources, as it will become clear throughout this section.

Addressing the grammatical sources first, this scenario can be proven on the basis of external evidence – notably, the description of the relative markers *llaḏī*, *ḥayṯu*, *man/mā* made by Early Arab Grammarians. Such a line of inquiry seems promising in view of the fact that, generally speaking, grammatical description by its very nature relies upon data taken from *langue* (Saussure 1916:13), whose defining characteristic is *frequent* usage, as discussed in Section 2.

It is interesting to note in this respect that a keyword search of the relative markers *llaḏī*, *ḥayṯu*, *man/mā, ḏī, ḏū, llāy, llā'i* in *Kitāb*[8] reveals that Sībawayhi's grammatical description of the OA relative clause mentions the relativizers *llaḏī*, *ḥayṯu*, *man/mā*, while no mention is made of the relativizers *ḏī, ḏū, llāy, llā'i*: from a Saussurean perspective, this points to the frequent nature of one group of relative markers, as opposed to the infrequent nature of the other.

In greater detail, Versteegh (1993:152) remarks that the technical definition of the OA relative clause (RC henceforth) found in *Kitāb* is composed of three conceptual elements: a syntactic criterion (i.e., an antecedent noun: *tamāmu l-ismi*), a semantic criterion (the ability to complement the meaning: *bi-hi yatimmu*), as well as a dedicated term denoting the OA RC as a whole (*ṣila*). For expository purposes, these concepts will be referred to here as 'antecedent requirement', 'complementation', and 'technical term' respectively. Versteegh cites in this connection *Kitāb* I, 87: "'the one I saw (was) so-and-so' ... *ra'aytu* 'I saw' is the *tamām* 'complement' of the name and by it *yatimmu* 'it [= the noun] is complete'.

[8] Conducted through a searchable version of *Kitāb*, based on Hārūn's edition and available at the following link: http://al-mostafa.info/data/arabic/depot/gap.php?file=0017 11-www.al-mostafa.com.pdf. The only instances of *ḏū* and *ḏī* found in this text denote possession (cp. fn. 5 above). A relative stem *ḏū* also occurs in the form *llaḏūna* (*Kitāb* III, 411), where, however, it is a syntactically conditioned allomorph of *llaḏī* (plural subjecthood).

It is neither a *ḫabar* 'predicate' nor a *ṣifa* '[…] adjective' " [9, 10], and: "in your complement to the noun, i.e., the relative clause" [11, 12].

Versteegh and Sara (cp. fn. 10, 12 above) interpret the term *ism* as 'antecedent noun' although its interpretation as 'relativizer' – i.e., as an expression elliptical for *al-ism al-mawṣūl* – cannot be ruled out in principle. The plausibility of Versteegh's and Sara's interpretation, however, lies in a distributional argument: Sībawayhi contrasts the construction *ism* + *tamām* with the constructions *ism* + *ḫabar* and *ism* + *ṣifa*, a contrastive paradigm that can be made sense of if, and only if, the linguistic entity to which the term *ism* refers is able to co-occur with all of the constituents *tamām*, *ḫabar*, *ṣifa* depending on the context. Plainly, while in Arabic grammar a relativizer is never described in this way, a noun is, which confirms Versteegh's and Sara's interpretation of *ism* as (antecedent) noun in the passage of *Kitāb* in question.

It is of particular relevance here that in this passage Sībawayhi exemplifies his definition of the OA RC by means of the construction *llaḏī ra'aytu fulānu*, for at least two reasons. In the first place, the OA RC that Sībawayhi concretely has in mind when defining a RC in *Kitāb* I, 87, is the one featuring the relative marker *llaḏī*. Secondly, implicit in the same example is the ability of an OA antecedent noun to have a covert realization (cp. Italian *chi*, *quanto* standing for *colui il quale*, *ciò che* respectively). Accordingly, Sībawayhi's technical definition of the OA RC encompasses not only the three conceptual elements mentioned by Versteegh, but also the syntactic property of 'covertness of antecedent'.

This fact is important because it allows to draw a comparison between the definition of OA RC discussed above and other definitions of it, as given elsewhere in *Kitāb*, where the antecedent noun is *explicitly* said to be covert. A case in point is *Kitāb* II, 105, which reads:

"This is the chapter about the [construction] in which a [covert] noun appears in the same syntactic position as the relative marker *llaḏī* that typically occurs in a definite context [and that in this case rather occurs in an indefinite context][13]. Under these circumstances and if the [covert] noun is a predicate, the noun in question is indefinite[14], standing for *raǧul* ['a man'], and [as such] requires a complement. For instance, you say:

---

[9] *llaḏī ra'aytu fulānu … ra'aytu tamāmu l-ismi bi-hi yatimmu wa-laysa bi-ḫabarin wa-lā ṣifa*

[10] Sara's (2006:44) translation, with transliteration adapted.

[11] *fī-mā atmamta bi-hi l-ismi yaᶜnī ṣ-ṣila*

[12] Versteegh's (1993:152) translation.

[13] This addition is required by the following context. See also the footnote below.

[14] Remarkably, Badawi, Carter and Gully (2004:506) offer a similar description for the relativizer whose antecedent is covert in Modern Standard Arabic – the most recent incar-

I. *hāḏā man aᶜrifu munṭaliqan* [‘this is *A MAN* about whose departure I know’]

II. *hāḏā man lā aᶜrifu munṭaliqan* [‘this is *A MAN* about whose departure I do not know’], i.e.,

III. *hāḏā llaḏī ᶜalimtu annī lā aᶜrifu-hu munṭaliqan* [‘this is *A MAN* about whose departure I admit I do not know’],

IV. *hāḏā mā ᶜindī muhīnan* [‘this is *SOMETHING* I consider insulting’]

The sentences in (I, II, IV) are relative clauses introduced by either *man* or *mā*. Such relative clauses are complement to *man* and *mā*, which means that in (I, II, IV) these two relative markers also work as antecedent nouns. The sentence in (III) and *llaḏī* stand in the same relationship of complementation”[15, 16].

Two remarks can be made concerning this definition of OA RC. First, Sībawayhi exemplifies it using some constructions featuring *man/mā* (as well as *llaḏī*). Second, the definition of OA RC, as given in *Kitāb* II, 105 includes three out of the four conceptual elements found in the definition of *Kitāb* I, 87:

- a semantic criterion, i.e. complementation (cp. the expression *yatimmāni bi-hi*)
- a syntactic criterion, i.e. antecedent requirement (cp. the expression *bi-manzilati raǧul*)
- a syntactic criterion, i.e. covertness of antecedent (cp. the expression *fa-yaṣīrāni-sman*)

That said, the two definitions differ in the technical term employed by Sībawayhi for ‘relative clause’: the former has *ṣila*, the latter *ḥašw*.

Besides *Kitāb* II, 105, another definition of OA RC that shares with *Kitāb* I, 87 the description of a covert antecedent noun is *Kitāb* III, 56-59:

“This is the chapter on the conditional clause. […] Among the particles introducing the conditional clause is […] *ḥayṯumā* [‘wherever’]: the particles *ḥayṯu* and *iḏ* cannot introduce the conditional clause unless combined with the particle *mā*, in which case the resulting combination behaves as a single particle. […] What prevents *ḥayṯu* from introducing the conditional clause is that [the verb following it, e.g.] *takūnu* in the sentence *ḥayṯu takūnu akūnu* [‘where you are, I am’] is a relative clause, [not a protasis]. In fact, this [sentence] is semantically equivalent to

---

nation of Classical Arabic. They in fact define this kind of relativizer as “indefinite”, when it manifests itself as *man/mā*.

[15] *hāḏā bābu mā yakūnu l-ismu fī-hi bi-manzilati llaḏī fī l-maᶜrifa iḏā buniya ᶜalā mā qabla-hu wa-bi-manzilati-hi fī l-iḥtiyāǧi ilā l-ḥašwi wa-yakūnu nakiratan bi-manzilati raǧul wa-ḏālika qawlu-ka hāḏā man aᶜrifu munṭaliqan wa-hāḏā man lā aᶜrifu munṭaliqan ay hāḏā llaḏī ᶜalimtu annī lā aᶜrifu-hu munṭaliqan wa-hāḏā mā ᶜindī muhīnan wa-aᶜrifu wa-lā aᶜrifu wa-ᶜindī ḥašwun la-humā yatimmāni bi-hi fa-yaṣīrāni-sman kamā kāna llaḏī lā yatimmu illā bi-ḥašwi-h*

[16] Translation mine. Capital italics stand for a covert element in the Arabic data. On the phrasal verb *buniya ᶜalā* in the sense of ‘(to act as) a predicate’, see Levin (1985:333).

l-makānu llaḏī takūnu fī-hi akūnu ['I am in the (same) place you are']. […] By contrast, if you make an interrogative sentence, what follows the [interrogative particle] is not a relative clause. The point relevant [for the description of *ḥayṯu* and *ḥayṯu-mā*] in this regard is that *ḥayṯu-mā* behaves as an interrogative particle in that it is not followed by a relative clause. For instance, if you say *ḥayṯu-mā takun akun* ['wherever you are, I am'] the verb [*takun*] is not a relative clause dependent on what precedes it [=*ḥayṯu-mā*], as much as if you say *ayna takūnu* ['where are you?'], the verb [*takūnu*] is not a relative clause dependent on what precedes it [=*ayna*]" [17, 18].

In this passage, Sībawayhi provides a contrastive definition of OA RC, according to which this kind of dependency is obtained when a particle *ḥayṯu* is detectable, which does not occur in the syntactic context of a conditional clause – i.e., when *ḥayṯu* is not combined with *mā* (*ḥayṯu takūnu akūnu*). It is self-evident that in the definition at issue the OA RC is exemplified by *ḥayṯu* – although, for the sake of completeness, it should be also noted that in the same passage another example of OA RC is found, which features the relativizer *llaḏī* (*l-makānu llaḏī takūnu fī-hi akūnu*).

The contrastive definition under discussion also includes three out of the four conceptual elements found in the definition of OA RC given in *Kitāb* I, 87, namely:

- a semantic criterion, i.e. covertness of antecedent (cp. the gloss of *ḥayṯu* as *l-makānu llaḏī fī-hi*)
- a syntactic criterion, i.e. antecedent requirement (cp. the expression *ṣilatin li-mā qabla-hu*)
- a technical term, i.e. *ṣila*

The general picture that transpires from a survey of Sībawayhi's definitions of the OA RC is that they are all construed as a conceptual structure consisting of at

[17] *hāḏā bābu l-ǧazā'i ... wa-mā yuǧāzā bi-hi mina ẓ-ẓurūfi ... ḥayṯumā ... wa-lā yakūnu l-ǧazā'u fī ḥayṯu wa-lā fī iḏ ḥattā yaḍummu ilā kulli wāḥidin min-humā mā ... wa-lākinna kulla wāḥidin min-humā ma^{c}a mā bi-manzilati ḥarfin wāḥidin ... wa-innamā muni^{c}a ḥayṯu an yuǧāzā bi-hā anna-ka taqūlu ḥayṯu takūnu akūnu fa-takūnu waṣlun la-hā ka-anna-ka qulta l-makānu llaḏī takūnu fī-hi akūnu ... iḏā stafhamta lam taǧ^{c}al mā ba^{c}da-hu ṣila fa-l-waǧhu an taqūla l-fi^{c}lu laysa fī l-ǧazā'i bi-ṣilatin li-mā qabla-hu ka-mā anna-hu fī ḥurūfi l-istifhāmi laysa ṣilatan li-mā qabla-hu wa-iḏā qulta ḥayṯu-mā takun akun fa-laysa bi-ṣilatin li-mā qabla-hu ka-mā anna-ka iḏā qulta ayna takūnu wa-anta tastafhimu fa-laysa l-fi^{c}lu bi-ṣilatin li-mā qabla-hu*

[18] Translation mine. On the term *ǧazā'* in the sense of 'conditional clause' in Sībawayhi's technical prose, see Sadan (2012:310-311) and references therein. See also Dévényi (1988:14-17).

least three of the following ingredients: covertness of antecedent, antecedent requirement, complementation, technical term. This is illustrated in Table 1 below.

Table 1: The OA Relative Clause – Sībawayhi's Definitions

Passage	Relativizer		Conceptual elements			
Kitāb			Covertness of antecedent	Antecedent requirement	Complementation	Technical term *ṣila*
I, 87	*llaḏī*		YES	YES	YES	YES
II, 105	*man/mā*	*llaḏī*	YES	YES	YES	NO
III, 56-59	*ḥayṯu*	*llaḏī*	YES	YES	NO	YES

The quadripartite conceptual structure in question qualifies as a genuine grammatical description of the OA RC that in Saussurean terms is plausibly based on the observation of frequent data (see end of Sect. 2). It ensues that the descriptiton that Sībawayhi offers of the OA RC introduced by the relative markers *llaḏī*, *ḥayṯu*, *man/mā* is external evidence proving the high frequency of such relativizers in this diachronic stage of Arabic, as anticipated at the outset of this section.

The same impression of high frequency for the relative markers *llaḏī*, *ḥayṯu*, *man/mā* we gain from the Koran. As discussed in Sect. 1, this religious and literary document is important for the present study because, in addition to being a primary source of OA along with the work of Early Arab Grammarians, it also provides this language with the internal evidence not necessarily found in the old grammatical treatises. This consists of raw statistics of the Koranic occurrences of the relative markers *llaḏī*, *ḥayṯu*, *man/mā* developed through *The Quranic Arabic Corpus*[19], according to which *man/mā* occurs in the Koran 1866 times (602 and 1266 times, respectively)[20], *llaḏī* 1464 times[21], *ḥayṯu* 29 times[22], with the caveat that in all likelihood *ḥayṯu* has a lower frequency relative to *llaḏī* and *man/mā* since its locative meaning can be also realized by the complex *llaḏī* plus *fī-hi*, a point already made by Sībawayhi in his definition of the OA RC headed by *ḥayṯu* (cp. his gloss of *ḥayṯu takūnu akūnu* as *al-makānu llaḏī takūnu fī-hi akūnu* immediately above).

[19] See fn. 4 above.

[20] http://corpus.quran.com/search.jsp?q=pos%3AREL%20stem%3A%D9%85%D8%A7, http://corpus.quran.com/search.jsp?q=pos%3AREL%20stem%3A%D9%85%D9%8E%D9%86

[21] http://corpus.quran.com/qurandictionary.jsp?q=%7Bl~a*iY

[22] http://corpus.quran.com/qurandictionary.jsp?q=Hyv

Hence, as in the case of the infrequent occurrence of the relativizers *ḏī, ḏū, llāy, llā'i*, the frequent occurrence of the relativizers *llaḏī*, *ḥayṯu*, *man/mā* results from the convergence of both external and internal evidence from OA. Furthermore, as in the case of the infrequent occurrence of the relativizers *ḏī, ḏū, llāy, llā'i*, the frequent occurrence of the relativizers *llaḏī*, *ḥayṯu*, *man/mā* must have been – in Saussurean terms – directly proportional to their degree of systemicity in OA. To put it differently, the relativizers *llaḏī*, *ḥayṯu*, *man/mā* are very likely to have entered the system of OA relativizers because of their fairly high frequency.

Putting such markers in their syntactic context yields the representation of the system of OA relativizers depicted in Table 2 below.

Table 2: The System of OA Relativizers in Context

Type	Source	Structure	Example	Translation
(1)	*Koran* VII, 196	*Aḷḷāhu llaḏī nazzala l-kitāba*	N+*llaḏī*+RC	God, who sent down the Book[23]
(2a)	*Koran* XLII, 13	*min-a d-dīni mā waṣṣā bi-hi Nūḥan*	*mā* + RC ; *min* + N	religion that He charged Noah with
(2b)	*Koran* II, 174	*mā anzala-ḷḷāhu min-a l-kitābi*		
(3)	*Ṣaḥīḥ* II, 148	*bi-Ḫayfi Banī Kināna ḥayṯu taqāsamū ᶜalā l-kufri*	N+*ḥayṯu*+RC	at Ḫayf Banī Kināna, where they took an oath of Kufr[24]
Symbols				
+	followed by			
;	followed or preceded by (syntactically unordered)			

4. A recent semantic analysis of (some) Old Arabic relative markers

It would be tempting to claim that the system of OA relativizers *llaḏī*, *man/mā*, *ḥayṯu*, as resulting from the above discussion, is identical to that usually ascribed to Classical Arabic, and its most recent incarnation, Modern Standard Arabic, in view of the fact that Badawi, Carter and Gully (2004:490, 630) define the system of

[23] The English translation of the Koran used here is by Arberry, unless otherwise stated.
[24] Mohsin Khan's translation. See Sect. 5 for further details on this OA example.

relativizers found in the latter variety as consisting precisely of a set of relativizers *llaḏī*, *ḥayṯu*, *man/mā*.

While empirically grounded, this observation should not obscure the fact that the identity between the system of OA relativizers and its Classical/Modern Standard Arabic counterpart is observed on the sound-side, which does not necessarily imply that it holds for the meaning-side too. In this light, even though it is commonly agreed in the literature that Classical/Modern Standard Arabic lacks any semantic distinction between restrictive and non-restrictive relativizers – and, more generally, between restrictive and non-restrictive relative clauses –, it cannot be taken for granted that the same situation be back-projected to OA as well. Therefore, it comes as no surprise that in a recent study Gensler (2004) argues that the distinction at issue did exist in OA, so that the relativizer *llaḏī* introduced a non-restrictive RC and the relativizer *man/mā* a restrictive one.

In order to appreciate this distinction, we must recall (see, among many others, Comrie 1981:138-139) that a non-restrictive RC is a modifier conveying new information and set off by a comma or intonational break from its antecedent, which consists of a noun conveying old information; whereas a restrictive RC is a modifier conveying old information and having as its antecedent a noun conveying new information, with no intervening comma or intonational break.

Gensler (2004) supports his non-restrictive analysis of *llaḏī* mainly with a statistical argument: on a total of 225 Koranic RCs having an overt antecedent and introduced by this relativizer (cp. (1) in Table 2 above), 140 occurrences are non-restrictive, 47 restrictive (38 occurrences providing no clear interpretive context).

Gensler also observes that one third of the non-restrictive occurrences of the relativizer *llaḏī* (48 times) is found in a syntactic context where the overt antecedent is the proper name *Aḷḷāh* or one of its epithets, which results in the structure: theophoric name $_{\text{ANTECEDENT}}$ + *llaḏī*$_{\text{RELATIVIZER}}$ illustrated in (1) above. In particular, an exact string search conducted through the *Tanzil* Koranic corpus[25] reveals that within this structure the proper name *Aḷḷāh* co-occurs with *llaḏī* with well more than chance frequency (32 times)[26], with the consequence that the

[25] An online annotated linguistic resource developed by Hamid Zarrabi-Zadeh, Department of Computer Engineering, Sharif University of Technology: http://tanzil.net/

[26] http://tanzil.net/#search/quran/"الله الذي", http://tanzil.net/#search/quran/"الذي الله"/p 2.

According to Diem (2007:85, 109) the proper name *Aḷḷāh* frequently occurs in combination with the relativizer *llaḏī* in the form of a formulaic structure *al-ḥamdu li-ḷḷāhi llaḏī* in Arabic documentary texts such as "*waqf* documents, appointments of high officials and marriage contracts between persons belonging to the upper class". Diem (2007: 67, 85) also highlights that these documents can be dated to the 9th century CE, and that "they are of so elaborated a style that they lack *any* deviation from the literary language" (italics in the original), which implies that they can be regarded as OA materials, in addition to the written sources mentioned at the end of Sect. 1. Two aspects of the expression *al-ḥamdu li-*

structure *Aḷḷāh*$_{\text{ANTECEDENT}}$ + *llaḏī*$_{\text{RELATIVIZER}}$ can be plausibly regarded as a systemic *fait de langue* in Saussurean terms (cp. Section 2) – and as a 'collocation' in Hallidayan terms (see e.g. Halliday 2002:166 on frequency as a defining characteristic of collocation).

By contrast, according to the *Tanzil* Koranic corpus, the proper name *Aḷḷāh* never co-occurs with the relativizer *man* in the structure * *Aḷḷāh*$_{\text{ANTECEDENT}}$ + *man*$_{\text{RELATIVIZER}}$: while the sequence *Aḷḷāh* + *man* occurs 14 times, it does *not* manifest an antecedent – relativizer relation[27] . Again, the ungrammaticality of this structure is expected from a collocational viewpoint, since according to Halliday (2002:160), the noun *tea* can co-occur with the adjective *strong*, as in *strong*$_{\text{ADJECTIVE}}$ + *tea*$_{\text{NOUN}}$ (cp. *Aḷḷāh*$_{\text{ANTECEDENT}}$ + *llaḏī*$_{\text{RELATIVIZER}}$) but not with the adjective *powerful*: **powerful*$_{\text{ADJECTIVE}}$ + *tea*$_{\text{NOUN}}$ (cp. **Aḷḷāh*$_{\text{ANTECEDENT}}$ + *man*$_{\text{RELATIVIZER}}$). It follows that the structure * *Aḷḷāh*$_{\text{ANTECEDENT}}$ + *man*$_{\text{RELATIVIZER}}$ can be regarded as a systemic gap that enters, along with the structure *Aḷḷāh*$_{\text{ANTECEDENT}}$ + *llaḏī*$_{\text{RELATIVIZER}}$, the morpho-syntactic opposition: proper name$_{\text{ANTECEDENT}}$ + *llaḏī*$_{\text{RELATIVIZER}}$ vs. *proper name$_{\text{ANTECEDENT}}$ + *man*$_{\text{RELATIVIZER}}$. What's more, the morpho-syntactic opposition in question is highly reminiscent of a semantico-syntactic opposition often noticed in literature, namely: proper name$_{\text{ANTECEDENT}}$ + non-restrictive relativizer vs. *proper name$_{\text{ANTECEDENT}}$ + restrictive relativizer (cp. English *John, which* vs. **John that*), as schematized in Table 3 below.

We would like to stress here that a parallelism of this kind between the morphological and semantic levels is too extended to be accidental. It is also worth pointing out that under the Saussurean axiom of the binary nature of linguistic sign, such a parallelism can be plausibly explained as the fact that the sound-strings *llaḏī, man/mā* are the audible counterparts of the sememes *NON-RESTRICTIVE RELATIVIZER, RESTRICTIVE RELATIVIZER*, respectively, and that vice versa, these sememes are the conceptual counterparts of the sound-strings *llaḏī, man/mā*.

ḷḷāhi llaḏī thus characterized are of particular relevance here. On the one hand, its formulaic character bears testimony to its systemicity (see Sect. 5 on the systemic status of formulae); on the other, Diem (2007:91) himself recognizes that "The relative clause of *al-ḥamdu li-llāhi llaḏī* is *non-restrictive*" (italics in the original) as a consequence of the co-occurrence of *llaḏī* with *Aḷḷāh*. In these respects, Werner's study strongly supports the hypothesis entertained here that the proper name *Aḷḷāh*$_{\text{ANTECEDENT}}$ followed by the non-restrictive relativizer *llaḏī* was a systemic *fait de langue* in OA.

[27] http://tanzil.net/#search/quran/"من الله", http://tanzil.net/#search/quran/"من الله"/p2, http://tanzil.net/#search/quran/"من الله"/p3, http://tanzil.net/#search/quran/"من الله"/p4.

Table 3: Two Overlapping Distributional Patterns

GRAMMATICALITY	RC-HEAD	RELATIVIZER	EXAMPLE
OK	Proper name	*llaḏī*	*Aḷḷāh (u/a/i)*$_X$ *llaḏī*$_X$
	Proper name	NON-RESTRICTIVE	*John, which*
*	Proper name	**man/mā*	**Aḷḷāh (u/a/i)*$_X$ *man*$_X$
	Proper name	*RESTRICTIVE	**John that*

In other words, the co-occurrence restriction that affects the exact strings *Aḷḷāh*$_{\text{ANTECEDENT}}$ + *llaḏī*$_{\text{RELATIVIZER}}$, **Aḷḷāh*$_{\text{ANTECEDENT}}$ + *man*$_{\text{RELATIVIZER}}$ in the Koran is adduced here as a distributional argument not only to corroborate Gensler's (2004) statistical argument concerning the non-restrictive meaning of *llaḏī*, but also to support another hypothesis by him, according to which *man/mā* is a restrictive relativizer.

This point is particularly compelling since the arguments made by Gensler in favor of the latter hypothesis seem quite weak. His claims is that *man/mā* has a restrictive meaning on the basis of 11 Koranic occurrences of RC introduced by this kind of relativizer and having as an overt antecedent the PP *min* + N, as exemplified in (2) above – yet, this number appears to be too low to be construed as a statistical argument proving any restrictive meaning of *man/mā*.

Indeed, Gensler maintains that in the eleven relative constructions examined by him the restrictive meaning of *man/mā* is a function of the P *min* that co-occurs with this relativizer and is also part and parcel of the antecedent *min* + N. In this view, the relative construction *min* + N ; *man/mā* + RC would be for all intents and purposes a partitive construction where the partitive P *min* would 'pick out' from a set denoted by the antecedent (e.g. *min-a l-kitābi* in (2b) above) the subset denoted by the RC featuring *man/mā* (e.g. *mā anzala-ḷḷāhu* in (2b) above), so that the antecedent would restrict the reference of its RC.

A semantic argument of this sort, however, is falsified by some examples given by Gensler himself in his investigation of the relative construction *min* + N; *man/mā* + RC, e.g. *min-a l-dīni mā waṣṣā bi-hi nūḥan* (cp. (2a) above) and *mā nansaḫu min āyatin* (*Koran* II, 106), where the context in which *min* occurs does not force its partitive reading. This is palpable in Arberry's translation, which avoids partitive renderings ('religion that He charged Noah with', 'whatever verse We abrogate' instead of 'portion of religion', 'whatever portion of verse' etc.) as well as in a grammatical observation that Ibn Hišām al-Anṣārī (d. 761/1360) in his work *Muġnī l-labīb* (I, 618) attributes to the medieval Muslim scholar Abū l-Baqā' al-

[c]Abkarī (d. 616/1219), according to which the P *min* is semantically unnecessary (*zā'ida*) in the relative construction *mā nansaḫu min āyatin*[28, 29].

On these grounds, the presence of the P *min* in the relative construction *min* + N; *man/mā* + RC is not guarantee of partitive/restrictive semantics and the most effective way to test Gensler's hypothesis that *man/mā* is a restrictive relativizer arguably is a distributional argument, based on the phenomenon of co-occurrence restriction, along the lines of above, rather than a statistical or semantic argument.

A further distributional argument that can be used to test this hypothesis relies upon a phenomenon of complementary distribution involving the word-order of relative constructions, as firstly observed by Lehmann (1984:271-278). After examining a world-wide sample of 80 languages, the German scholar draws the typological generalization that where languages have a choice between the post-nominal position of the RC (= N + RC) and pre-nominal one (RC + N), the post-nominal position will be primarily used for non-restrictive RCs, the pre-nominal position for restrictive RCs, provided that the presence of the sememe of restrictiveness in the pre-nominal RC correlates with the lack of the sememe of tense. In Lehmann's (1984:49, 52 ff.) own terminology, the pre-nominal RC must be a 'Relativpartizip', or, in more traditional terms, a reduced RC[30], as illustrated in Table 4 below.

[28] *qawlu Abī l-Baqā'i al-[c]Abkariyyi fī mā nansaḫu min āyatin inna-hu yağūzu kawnu āyatin ḥālan wa-min zā'idatan*

[29] On the term *zā'ida* in the sense of 'semantically unnecessary', see Versteegh (1993: 145-146), which dates its origin to the earlier period of Arabic grammar (750 CE). Versteegh also shows that at that time this term coexisted with the synonyms *ṣila, laġw*.

[30] In a more recent English formulation of the notion of 'Relativpartizip', Lehmann makes this point very clear: "Since embedded RCs function as adjectivals or nominals, they generally show restrictions on tense/aspect/mood, genitive case on the logical subject and, ultimately, non-finiteness of the verb (relative participle)" (http://www.christianlehmann.eu/ling/lg_system/grammar/nexion/rel_clause.html). It should be also noted that a well-studied subtype of RC endowed with a tenseless verb is the so-called 'reduced' RC, where the tenseless verb is a covert manifestation of *be*, e.g. *John's house in the wood*. Here, the RC-status of the modifier *in the woods* is evidenced by a substitution test: cp. a paraphrase such as *the house of John's which is in the woods* (Taylor 2000:110).

Table 4: Lehmann's Distributional Generalization on Relative Clauses

	Complementary Distribution	
RC	Type A	Type B
Syntax	Pre-nominal	Post-nominal
Semantics	Tenseless	Tensed
	Restrictive	Non-restrictive

Lehmann's generalization plausibly implies that if two given RC-types A and B manifest the two-fold opposition {pre-nominal, tenseless} vs. {post-nominal, tensed}, this is good typological/distributional evidence for positing also an opposition {restrictive} vs. {non-restrictive} between them.

Applying Lehmann's generalization to OA, then, the RC introduced by *man/mā* would be restrictive, if it proved to be at once tenseless and pre-nominal. Concerning the first property, it is the authoritative opinion of Arab Grammarians that in the Koranic passage *mā ṭāba min-a n-nisā'i* I (*Koran* IV, 3), the gnomic context forces an atemporal reading of the suffix-conjugation form *ṭāba* found in the RC introduced by *mā*: for instance, Ibn Yaᶜīš (d. 553/1158) glosses the occurrence of the verb *ṭāba* under scrutiny as *ṭ-ṭayyibu* 'the good' in *Šarḥ al-Mufaṣṣal* (II, 380)[31] (cp. also Arberry's translation 'such women as seem good to you'). As regards the second property, this is immediately apparent from the Koranic example (2b) above, showing that the RC introduced by *man/mā* can fulfill a pre-nominal position, in sharp contrast to the RC introduced by *llaḏī*, which cannot: *min* + N + *man/mā* + RC, *man/mā* + RC + *min* + N, vs. N + *llaḏī* + RC, **llaḏī* + RC + N. What follows from this fundamental asymmetry, illustrated in Table 5 below, is that a RC introduced by *man/mā* is restrictive owing to its ability to combine a tenseless meaning with a pre-nominal syntax (i.e., to precede the complex: *min* + N).

[31] *mā ṭāba min-a n-nisā'i bi-maᶜnā ṭ-ṭayyibu min-hunna*

Table 5: OA Relative Clause – A Word-Order Asymmetry

RELA-TIVIZER	RELATIVI-ZATION	GRAMMA-TICALITY	STRUCTURAL DESCRIPTION					
man/mā	Pre-nominal	OK	*man/mā*	RC	*min*	N		
	Post-nominal	OK			*min*	N	*man/mā*	RC
llaḏī	Pre-nominal	*	*llaḏī*	RC		N		
	Post-nominal	OK				N	*llaḏī*	RC

To summarize the main results of this section, we have used a pattern of complementary distribution observed in the OA RC, and opposing the bundles of features {pre-nominal, tenseless} to the bundle of features {post-nominal, tensed}, as a distributional argument for characterizing the relativizer *man/mā* as originally restrictive. We have independently arrived at the same result by building another distributional argument, based on the phenomenon of co-occurrence restriction between restrictive relativizers and proper Ns, clearly exemplified by the opposition between the systemic gap * $Aḷḷāh_{\text{ANTECEDENT}}$ + $man_{\text{RELATIVIZER}}$ vs. the collocation $Aḷḷāh_{\text{ANTECEDENT}}$ + $llaḏī_{\text{RELATIVIZER}}$. Moreover, the distributional argument at issue also supports a non-restrictive interpretation of the relativizer *llaḏī*, along with a statistical argument developed by Gensler (2004). Generally speaking, the distributional line of reasoning adopted in this section captures Gensler's (2004) intuition that OA, unlike Classical Arabic, drew a distinction between restrictive and non-restrictive relativizers when the OA relativizers are combined with an overt RC-head.

A distributional approach of this kind has the advantage of also including in the analysis the OA relativizers combined with a *covert* RC-head, which in *Kitāb* II, 105 is said to convey the meaning of an indefinite noun such as *rağul* 'a man'(see Sect. 3). If, in fact, we consider that the clausal context in which this type of RC-head occurs, namely *hāḏā man a^{c}rifu munṭaliqan* 'this is *A MAN* about whose departure I know', etc., forces its specific reading[32], and that, in turn, the specific reading of an indefinite noun forces the restrictive reading of a relativizer (cp. the ungrammaticality of the non-restrictive relativizer *, who* when combined with the specific noun *a man* in **I met a man, who was going to St. Ives*, as per Loock 2010:16), it naturally follows from the examples of OA RCs mentioned in *Kitāb* II, 105 that the OA relativizers *llaḏī* and *man/mā* have restrictive semantics when co-occurring with a covert RC-head.

[32] I.e., the referent of *rağul* in this context cannot be identified independently of the embedding of the RC, unlike the generic noun *a house* in *he bought a house* (cp. Loock 2010: 16, 38).

Despite the advantage of a unified semantic analysis of the overt and covert phonological contexts in which *llaḏī* and *man/mā* occur, a distributional approach to the OA system of relativizers along the above lines faces a major limitation in describing *ḥayṯu*: while this relativizer has been shown in Section 3 to be as systemic as *llaḏī* and *man/mā*, the analysis carried out hitherto remains silent on its semantic behavior in terms of restrictiveness or non-restrictiveness vis-à-vis a covert or overt RC-head. This theoretical problem is dealt with in the next section.

5. Extending the proposal: the Old Arabic relative marker *ḥayṯu*

In order to shed light on the semantics of *ḥayṯu*, we should place this relative marker and the RC introduced by it in their appropriate distributional context, which from a phonological standpoint can be either an overt or covert RC-head.

Starting with the first type of context, we shall focus on *Ṣaḥīḥ*, rather than on the Koran, in view of the fact that the former source attests to a construction in which (a RC introduced by) *ḥayṯu* co-occurs with an overt RC-head, whereas the latter does not.

In effect, in reporting a religious schism that took place in a locality known as Ḫayf Banī Kināna or al-Muḥaṣṣab, and that gave rise, within Muḥammad's tribe, to a faction reluctant to accept Islamic revelation, al-Buḫārī (d. 256/870) reproduces a speech of Muḥammad's, which can be regarded as a genuine instance of OA datum due to its chronological location; moreover, it includes the aforementioned kind of relative construction.

Yet there is another reason for concentrating on *Ṣaḥīḥ*: in its authoritative edition by Muḥammad Zuhayr Ibn an-Nāṣir an-Nāṣir, the relative construction in question, in addition to exhibiting an overt RC-head in combination with the relativizer *ḥayṯu*, occurs 7 times with no or slight variation. More specifically, it manifests itself in the identical tokens *bi-Ḫayfi Banī Kināna ḥayṯu taqāsamū ᶜalā l-kufri* (*Ṣaḥīḥ* II, 148; V, 51, 148; cp. also (3) in Table 2 above) and the quasi-identical tokens *bi-Ḫayfi Banī Kināna ḥayṯu taqāsamū ᶜalā l-kufri yaᶜnī ḏālika l-muḥaṣṣaba* (*Ṣaḥīḥ* II, 148), *bi-Ḫayfi Banī Kināna l-muḥaṣṣabi ḥayṯu qāsamat Qurayšun ᶜalā l-kufri* (*Ṣaḥīḥ* IV, 71), *l-Ḫayfu ḥayṯu taqāsamū ᶜalā l-kufri* (*Ṣaḥīḥ* V, 148), *bi- Ḫayfi Banī Kināna ḥayṯu taqāsamū ᶜalā l-kufri yurīdu l-muḥaṣṣaba* (*Ṣaḥīḥ* IX, 140).

Technically speaking, the tokens of relative construction thus characterized qualify on the whole as a 'formula', in Monroe's (1972:15) sense, which defines it as a "verbatim, or nearly verbatim repetition" (with the caveat that the usage of this term here is merely descriptive and non-committal as to any particular formulaic

theory[33]). The relevance of the formulaic nature of this relative construction for our understanding of the phonological and semantic distribution of the relativizer *ḥayṯu* becomes apparent when we adopt a sociolinguistic perspective. According to Monroe (1972:8), among the social conditions that crucially contribute to the rise of a formula is its frequent usage by a given (literate) community, where "by a process of natural selection, a traditional stock of collectively known formulas is elaborated and adopted". Furthermore, Saussure highlighted that on a social level, the same frequent usage of a given expression which is quintessential to formularity correlates, on a linguistic level, with its systemicity, as discussed in Sect. 3.

A sociolinguistic scenario of this sort evidences that the aforesaid tokens, because of their formulaic nature, instantiate a frequent, and therefore systemic type, where the distributional context co-occurring with (the RC featuring) *ḥayṯu* is an overt RC-head that consists of a proper N of place.

We should recall at this point from Sect. 4 that a systemic combination – or, technically speaking, collocation – made up of a proper N and a relative marker points to the latter's non-restrictiveness in semantic terms, so that *Ṣaḥīḥ* bears the marks of a stage of OA where the relativizer *ḥayṯu* combined with an overt RC-head had a non-restrictive meaning.

With this in place, it is now possible to move on to the discussion of the second kind of distributional context in which *ḥayṯu* and related RC manifest themselves, notably a covert RC-head that denotes a place, as explicitly stated in *Kitāb* III, 56-59, where the relative construction *ḥayṯu takūnu akūnu* is glossed as *al-makānu llaḏī takūnu fī-hi akūnu* (cp. end of Sect. 3). In this respect, the description of the covert RC-head that co-occurs with *ḥayṯu*, as found in Arabic grammatical sources, lends independent confirmation to the above description of its *overt* counterpart, which is equally characterized as a noun of place, based on *Ṣaḥīḥ* (cp. the RC-head *Ḫayf Banī Kināna* or *al-Muḥaṣṣab* in (3) in Table 2 above). But this is not the whole of the matter: Arabic grammatical sources, among which is *Šarḥ al-Mufaṣṣal* (III, 114), also highlight that in this distributional context the relativizer *ḥayṯu* has the special syntactic property of governing the RC introduced by it, which behaves accordingly as a genitive complement[34].

It is worth noting at this point that Ouhalla (2000:234-235) interprets the Arabic genitive complement as semantically restrictive. In fact, Ouhalla observes that a restriction of co-occurrence between the definite article and genitive complement (e.g. *l-ahlu, ahlu-l-kitābi* vs. **l-ahlu-l-kitābi*) cuts across the Semitic languages (OA included), and that a restriction of this sort typically is semantically-

[33] See e.g. Ostle (1982) for some differences between Monroe's (1972) formulaic theory and Zwettler's (1978).

[34] *fa-ka-mā kānat iḏ muḍāfatan ilā ǧumlatin tūḍiḥu-hā ūḍiḥat ḥayṯu bi-l-ǧumlati llatī tūḍaḥu bi-hā iḏ*.

conditioned, in that two constituents conveying the same meaning cannot co-occur on the syntagmatic axis (cp. *dog-s, f-ee-t* vs. **f-ee-t-s*). According to Ouhalla, such a cross-linguistic pattern of restriction can be coupled with the standard interpretation of the definite article as restrictive, to derive the semantics of the Arabic genitive complement, the latter being incapable of co-occurring with the definite article because of sharing with it a restrictive meaning.

Thus, the ungrammatical type **l-ahlu-l-kitābi* provides distributional evidence for a restrictive analysis of the genitive complement introduced by the instance of *ḥayṯu* that co-occurs with a covert RC-head. This amounts to saying, from a compositional perspective, that this kind of *ḥayṯu* is the locus of restrictiveness in the distributional context under study or, in other words, that it is a restrictive relativizer, and that as such it is neatly opposed to its counterpart combined with an overt RC-head, which behaves as a non-restrictive relativizer instead, and is exemplified by the type (3) in Table 2 above[35].

In sum, the distributional approach to *ḥayṯu* pursued in this section has in essence relied upon a cross-linguistic and language-specific restriction of co-occurrence (**John that* and **l-ahlu-l-kitābi*, respectively) to provide a unified semantic analysis of the overt and covert phonological contexts in which this relative marker occurs, in terms of (non-)restrictiveness, and, in doing so, it has circumvented the interpretive problems related to *ḥayṯu* raised at end of the previous section. Interestingly, this result affords a description of the system of OA relativizers that is more exhaustive than Gensler's (2004), as illustrated in Table 6 below, which rewrites Tables 2 and 5 in a more abstract fashion, and encompasses the metalinguistic observations by Sībawayhi dealt with in Sect. 3.

[35] As a corollary, the genitive complement analysis that Arab Grammarians worked out for the type *ḥayṯu takūnu akūnu* can be regarded from an epistemological perspective as a way of expressing in syntactic terms the semantic property of restrictiveness.

Table 6: OA Relative Markers – Distribution (First Approximation)

<table>
<tr><td rowspan="2">RC-head</td><td colspan="6">Phonology</td></tr>
<tr><td colspan="3">Overt</td><td colspan="3">Covert</td></tr>
<tr><td>Relativizer</td><td>Semantics</td><td colspan="2">Syntax</td><td>Semantics</td><td colspan="2">Syntax</td></tr>
<tr><td>llaḏī</td><td>non-restrictive</td><td></td><td>post-nominal</td><td>restrictive</td><td></td><td>post-nominal</td></tr>
<tr><td>man/mā</td><td>restrictive</td><td>pre-nominal</td><td>post-nominal</td><td>restrictive</td><td></td><td>post-nominal</td></tr>
<tr><td>ḥayṯu</td><td>non-restrictive</td><td></td><td>post-nominal</td><td>restrictive</td><td></td><td>post-nominal</td></tr>
<tr><td rowspan="3">Examples</td><td colspan="3">(1) Aḷḷāhu llaḏī nazzala l-kitāba</td><td colspan="3">(4) hāḏā (RAĞULUN) llaḏī ᶜalimtu annī lā aᶜrifu-hu munṭaliqan</td></tr>
<tr><td colspan="3">(2a) min-a d-dīni mā waṣṣā bi-hi Nūḥan /
(2b) mā anzala-ḷḷāhu min-a l-kitābi</td><td colspan="3">(5) hāḏā (RAĞULUN) man aᶜrifu munṭaliqan</td></tr>
<tr><td colspan="3">(3) bi-Ḫayfi Banī Kinanāta ḥayṯu taqāsamū ᶜalā l-kufri</td><td colspan="3">(6) ḥayṯu takūnu akūnu (= al-makānu llaḏī takūnu fī-hi akūnu)</td></tr>
</table>

A study of the distribution of *ḥayṯu* along these lines, though, solves one problem, but creates another: the attentive reader will have noticed that the (non-) restrictiveness-based distinction that governs the system of OA relativizers *llaḏī, man/mā*, *ḥayṯu* as a whole enters a pattern of complementary distribution with respect to a *phonological* context consisting of a covert or overt RC-head, in the case of *llaḏī* and *ḥayṯu*, but not of *man/mā* (cp. the dark grey cells in Table 6), since the latter exhibits a restrictive meaning irrespective of the overt or covert nature of the RC-head it co-occurs with.

Matters are further complicated when we adopt a broader perspective, which considers the aforesaid (non-)restrictiveness-based distinction relative to a *syntactic* context of distribution. If, in fact, we recall from Table 4 that tenseless, restrictive RCs and their tensed, non-restrictive counterparts enter a cross-linguistic pattern of complementary distribution, which is sensitive to their pre-nominal or post-nominal position, respectively, the actual distribution of the OA relativizers and related RCs looks typologically odd, in view of the fact that in this language a tenseless, restrictive RC and, more generally, a restrictive RC *tout court*, indeed occurs in a post-nominal position. Such an anomalous behavior, which was labeled

as a 'word-order asymmetry' in Sect. 4 and depicted in Table 5 there, allows for the complementary distribution of the features {pre-nominal, tenseless, restrictive} vs. {post-nominal, tensed, non-restrictive} only insofar as the non-restrictive and post-nominal relativizers *llaḏī, ḥayṯu* and the restrictive and pre-nominal relativizer *man/mā* with an overt RC-head are concerned (cp. Table 6, first column, and the light grey cells in its second column).

Therefore, the research carried out thus far establishes that the restrictive/non-restrictive semantics of the system of OA relativizers *llaḏī, man/mā, ḥayṯu*:

(α) is phonologically context-sensitive, except for the *restrictive man/mā*, able to co-occur with *both* an overt and -unexpectedly- a covert RC-head;

(β) is syntactically context-sensitive, only insofar as the non-restrictive and post-nominal *llaḏī, ḥayṯu* and the restrictive and pre-nominal *man/mā* with an overt RC-head are concerned, the remaining relativizers, all of them *restrictive*, being able to occur in an unexpected post-nominal position;

Briefly put, the system of OA relativizers is rather nebulous especially in *its restrictive manifestations*. The next section offers a solution for this puzzling distributional behavior, which seriously threatens the viability of a semantic analysis of the system of OA relativizers, as developed until now.

6. Potential phonological and syntactic counterexamples to a semantic analysis of the Old Arabic relative markers

The unexpected distribution of the OA restrictive relativizers within the phonological and syntactic contexts in (α, β), i.e. the relativizers tabulated under (2a, 4, 5, 6) in Table 6, calls for a more in-depth examination of their phonology and syntax. For convenience's sake, the relative markers in question will be referred to henceforth as overtly headed *man/mā* (2a), and covertly headed *llaḏī, man/mā, ḥayṯu* (4, 5, 6) respectively – with their restrictive semantics implied in this terminology, unless otherwise stated.

Let's begin our inquiry with the covertly headed relativizers that, according to Sībawayhi, cumulate the functions of a N and a relativizer proper (cp. his metalinguistic glosses *al-ismu fī-hi bi-manzilati llaḏī* and *al-makānu llaḏī takūnu fī-hi* presented in Sect. 3). Such a description can be reformulated, in modern terms, by saying that a word such as *llaḏī, man/mā, ḥayṯu* contains both a zero-morpheme standing for a N which acts as a RC-head, and (a) phonologically-realized morpheme(s) indicating the relativizer, so that the word at issue displays the ratio of two or more morphemes per word. Sībawayhi's description is confirmed in particular by the fact that the covertly headed *llaḏī* has an exceptional indefinite reading (cp. Sect. 3), which can be only the result of an indefinite

antecedent N 'parasitic' onto it. Putting this observation into a typological perspective, the ratio morpheme/word observed in the OA word-like structure: $Ø_N$ + *llaḏī, man/mā, ḥayṯu* or, more formally, #N-grammatical morpheme(s)# can be considered as high with respect to another salient OA structure of the same kind, namely the structure: N + case-ending, which can be exemplified through the Koranic nominal (*l-*)*fulk-i*. Here, in fact, the lack of *sandhi*-internal phenomena between the N and case-ending (e.g. no palatalization between a N-final velar and the oblique case *i*: *fulki* vs. **fulci)* evidences a word-boundary (*fulk#i#*, cp. Corriente 1971a:47), whence a ratio of one morpheme per word, i.e. #N#grammatical morpheme(s)#.

Moreover, two different trends in Semitic historical linguistics, such as the diffusionist and genetic models (see Moscati *et al.* 1964:112-113 and Garbini & Durand 1994: 98-99, 136-137, respectively) dovetail in that the two above structures have a different origin: the structure N + case-ending is a pre-Semitic archaism, whereas its counterpart N + relativizer is a North-West Semitic innovation, at least as far as the relativizer *llaḏī,* is concerned. It therefore stands to reason that OA underwent an incipient drift from analytical (cp. pre-Semitic N + case-ending) to synthetical type (cp. North-West Semitic N + *llaḏī*), in concordance with Corriente (1971a:48-50), and that, specifically for the covertly headed relativizers $Ø_N$ + *llaḏī, man/mā, ḥayṯu*, they underwent a process of synthesization, by means of which the analytical N working as a RC-head became more synthetical via phonological deletion and affixation to the relativizer *llaḏī, man/mā, ḥayṯu*[36]. It ensues that each of the covertly headed relativizers found in the types (4, 5, 6) is better seen as a synthetical N that develops out of a RC-head plus a grammatical marker, namely a restrictive relativizer, as a result of a process of synthesization. This is schematized in Table 7 below.

[36] A well-established trend in Arabic and Semitic linguistics assumes the opposite drift (synthetical-to-analytical) to be at work in the diachronic evolution of Arabic, but typological arguments falsify this interpretive paradigm (see Owens 2006 for details and discussion). Similarly, Petrácek (1981:172) draws from his study of the OA/Classical Arabic verbal system the conclusion that: "La structure de l'arabe apparaît maintenant comme agglutinative plus qu'auparavant".

Table 7: Covertly Headed *llaḏī, man/mā, ḥayṯu* – Process of Synthesization

Index of Synthesis	Analytical Basic Form	Process	Synthetical Derived Form
	#N#grammatical marker(s)#	Synthesization	#N-grammatical marker(s)#
	[*fulk#i*$_X$]	# ⇒ -	[$\emptyset_{\text{INDEFINITE}}$-*llaḏī*$_X$]
Ratio morpheme/word	= 1		≥ 2

Moving from the restrictive relativizer that is covertly headed to its overtly headed counterpart, the same analysis arguably applies. This argument is built as follows:

Firstly, the overt N that heads this kind of restrictive relativizer (phonologically realized as *man/mā*, cp. (2a) above), has a particular phonological behavior in OA, in that not only is it combined with the particle *min* to yield the structure: *min* + N, but also tends to trigger phonological reduction on the particle in question, which accordingly becomes *mi*, and yields a structure: *mi* + N. The relevant *loci probantes* are some lines of pre-Islamic poetry recorded in *Lisān* VI, 4282, which attest to expressions such as *mi-l-kaḏibi* for *min-a l-kaḏibi* (cp. also Esseesy 2010:219)[37].

In second place, such a phonological behavior constitutes appreciative evidence that the particle *mi* was a prefix forming a single word with the N following it, given that phonological reduction is a suprasegmental phenomenon which cross-linguistically diagnoses a word-like unit, and OA is no exception to this general trend; in effect, knowing that an OA construct state such as *ṯalāṯatu darāhima* (cp. *Kitāb* IV, 464) blocks adjective-insertion just as words do ('uninterruptability': cp. *three dogs* vs. **dog-three-s*), we can establish in this way that it is a word-like unit, so that, in the wake of Fleisch (1959) and Benmamoun (2005), we can ascribe precisely to such a word-like status the phonological reduction that affects the construct state *ṯalāṯatu darāhima* in 'contracted' forms such as *ṯalāttu darāhima*, where the vowel *a* preceding the bound *t* is dropped, as reported by Sībawayhi (*Kitāb* IV, 464)[38].

On these grounds, a sharp contrast arises between the structure *min* + N, where the ratio morpheme/word is 1 (analytical #*min*#N#), and *mi* + N, where the ratio

[37] *wa-yaǧūzu ḥaḏfu n-nūni min min wa-ᶜan ᶜinda l-alifi wa-l-lāmi li-ltiqā'i s-sākinayni wa-ḥaḏfu-hā min min akṯaru min ḥaḏfi-hā min ᶜan li-anna duḫūla min fī l-kalāmi akṯaru min duḫūli ᶜan wa-anšada … mi-l-kaḏibi*

[38] *wa-ḥuǧǧatu-hu qawlu-hum ṯalāttu darāhima tuddaġamu ṯ-ṯa'u min ṯalāṯati fī l-hā'i iḏā ṣārat tā'an*

morpheme/word is at least 2 (synthetical #*mi*-N#), a diachronic alternation proving that the overtly headed *man/mā* (2a) underwent a process of synthesization, by means of which the analytical N working as a RC-head became more synthetical, via phonological reduction and affixation to *mi(n)*[39].

The overall scenario borne out from these typological observations in the OA restrictive relativizers *llaḏī, man/mā, ḥayṯu* (2a, 4, 5, 6) is that they occur in a phonological context of distribution far more uniform than previously thought, since the RC-heads with which they co-occur, albeit different on a segmental level (overt particle *mi(n)* + N vs. covert N), share a common property on a suprasegmental level. Such RC-heads are *both* 'synthetical' Ns, i.e. Ns occupying the same high position along the so-called 'index of synthesis', which refers to how many morphemes can occur per word in a given language (Comrie 1981:46-51). This property is regarded here as suprasegmental, in the sense that the high index of synthesis or, practically speaking, the word-like status observed in the overt complex *mi(n)* + N and in the covert N acting as RC-heads is the result of a process of phonological reduction, which is partial in the case of the overt complex *mi(n)* + N (final consonant-dropping, targeting the particle: *min* > *mi*), and total in the case of the covert N (deletion of an entire constituent, affecting N: *rağul* > $ø_{\text{RAĞUL}}$) .

With the overtly and covertly headed *llaḏī, man/mā, ḥayṯu* subsumed under the common category of relativizers headed by a synthetical N, actually the phonological(-segmental) opposition itself between overtly vs. covertly headed relative markers makes no longer sense, and a new phonological(-suprasegmental) opposition arises instead. On the one hand, in the aforesaid types (2a, 4, 5, 6) and, in addition, (2b), the restrictive and post-nominal relativizers *llaḏī, man/mā, ḥayṯu*, as well as the restrictive and pre-nominal relativizer *man/mā*, are headed by a synthetical N, which can be partially or totally reduced (*mi(n)* + N or $ø_{\text{N}}$). On the other hand, in the remaining types (1, 3) the non-restrictive and post-nominal relativizers *llaḏī, ḥayṯu* are headed by an analytical N, whose analytical character is apparent from the lack in it of either *min* > *mi* reduction or N-deletion.

As emerges above, the system of OA relativizers *llaḏī, man/mā, ḥayṯu* is context-sensitive to a suprasegmental-phonological level of linguistic representation, which obeys a pattern of complementary distribution: abstracting away from syntactic factors, to which we shall return shortly, the restrictive relative markers *llaḏī, man/mā, ḥayṯu* co-occur with a synthetical RC-head, whereas their non-restrictive counterparts *llaḏī, ḥayṯu* co-occur with an analytical RC-head.

A more rigorous formulation of the above statement would be that *all* such restrictive relativizers co-occur with *both* the kinds of synthetical RC-head attested

[39] That a CV-sequence in OA is not an independent word, and rather an affix belonging to a larger word, is also shown by graphical words such as <*lizaydin*>, <*bizaydin*>, where the phonological shape taken by the *bona fide* prefixes *li, bi* is precisely a CV-sequence.

in OA, defined above as a partially or totally reduced N (*mi(n)* +N or $ø_N$). While this distributional pattern has been already highlighted with regard to *man/mā* (cp. (2), (5) above), it still requires clarification in the case of *llaḏī, ḥayṯu*, which the discussion thus far has shown to co-occur with a totally reduced RC-head (cp. (4), (6) above), but not with its partially reduced counterpart.

That the restrictive relative marker *llaḏī* can co-occur with a partially reduced RC-head (i.e. with *mi(n)* +N) is proven by its ability to replace the *restrictive* relative marker *man/mā* that co-occurs with a RC-head of this sort (substitution test): by way of illustration, in the Koran a context of post-nominal RC is found, notably ___ *ǧā'a-ka min-a l-ᶜilmi*, which admits both the relativizers *mā* and *llaḏī* (cp. *llaḏī ǧā'a-ka min-a l-ᶜilmi* in *Koran* II, 120, and *mā ǧā'a-ka min-a l-ᶜilmi* in *Koran* II, 145, III, 61, XIII, 37). Furthermore, in *Koran* XLII, 13 we find an instance of pre-nominal RC with a partially reduced head, where the *restrictive* relativizer *mā* is conjoined to the relativizer *llaḏī*, which points to the latter's restrictive meaning in this context (law of coordination of likes): *min-a d-dīni mā waṣṣā bi-hi Nūḥan wa-llaḏī awḥaynā ilay-ka wa-mā waṣṣaynā bi-hi Ibrāhīma*.

Regarding the restrictive relativizer *ḥayṯu*, two significant clues of its ability to co-occur with a partially reduced RC-head comes from OA relative constructions such as *irᶜaw min arḍi-nā ḥayṯu ši'tum* and *yuṣarrifūna ḥayṯu šā'ū min arḍi-him* documented, respectively, in pre-Islamic poetry and Ibn Iṣḥāq's (d. 151/761) materials, as transmitted in the *Sīra* by Ibn Hišām (d. 218/833) (*apud* Reckendorf 1895: 635). In effect, besides the relative marker *ḥayṯu* and the partially reduced RC-head, these constructions display two interesting structural properties: (i) the RC-head contains a possessor-denoting genitive (cp. *nā, him* in *min arḍi-nā, min arḍi-him*) and, (ii) the RC, which is pre-nominal in *irᶜaw min arḍi-nā ḥayṯu ši'tum*, and post-nominal in *yuṣarrifūna ḥayṯu šā'ū min arḍi-him*, contains a gnomic – and therefore tenseless – instance of suffix-conjugation, so that it qualifies as a tenseless RC (cp. end of Sect. 4). It should be observed at this point that English provides useful parallels to the phenomena which we find in OA: relying upon previous work by Chomsky and C. Lyons, Taylor (2000:110) points out that in English a RC-head that contains a possessor-denoting genitive (e.g. *John's*) cannot be generally combined with a restrictive RC (cp. **John's book that you borrowed*), except when the RC in question is reduced (cp. *John's house in the woods/the house of John's which is in the woods* and fn. 30 above).

If we recall from the end of Sect. 4 that a reduced RC is no more than a subtype of tenseless RC, a striking similarity emerges between the English relative construction *John's house in the woods* and the OA relative constructions where the relativizer *ḥayṯu* co-occurs with a partially reduced RC-head, with respect to the co-occurrence of the two structural properties indicated as (i) and (ii) immediately above. What is even more remarkable is that such properties in English are associated with the restrictive semantics of the RC, which lends

support, by extension, to the hypothesis that the same holds for OA. In this interpretive scenario, since the relativizer *ḥayṯu* that co-occurs with a partially reduced RC-head introduces a restrictive tenseless RC such as *min arḍi-nā ḥayṯu ši'tum*, and *ḥayṯu šā'ū min arḍi-him*, we are justified in interpreting it as a restrictive relativizer.

Taken as a whole, this semantic survey of the relative markers *llaḏī* and *ḥayṯu* that co-occur with a partially reduced RC-head reveals that, as anticipated above, the relativizers *llaḏī, man/mā, ḥayṯu* not only are all restrictive but also co-occur with a synthetical RC-head that can be *both* partially and totally reduced, whereas their non-restrictive counterparts *llaḏī, ḥayṯu* co-occur with an analytical RC-head.

The important corollary of this achievement is that we are now a in position to solve the problem of the anomalous phonological distribution of the restrictive relative marker *man/mā* (cp. (α) above), by simply re-conceptualizing in suprasegmental terms the phonological context in which the *whole* system of OA relativizers manifests itself. In fact, as can be gleaned from a glance at the dark grey cells of Table 8 below, this distributional anomaly disappears as soon as we shift to a suprasegmental level of analysis, which reduces it to an element of the phonological pattern of complementary distribution alluded to above (i.e. non-restrictive relativizer : analytical RC-head = restrictive relativizer : synthetical RC-head).

Table 8: OA Relative Markers – Distribution (Refined Version)

<table>
<tr><td rowspan="3">RC-head</td><td colspan="8">Phonology</td></tr>
<tr><td colspan="2">Analytical</td><td colspan="6">Synthetical</td></tr>
<tr><td colspan="2"></td><td colspan="4">Partial Reduction</td><td colspan="2">Total Reduction</td></tr>
<tr><td>Relativizer</td><td>Sem.</td><td>Syntax</td><td>Sem.</td><td>Syntax</td><td>Sem.</td><td>Syntax</td><td>Sem.</td><td>Syntax</td></tr>
<tr><td>llaḏī</td><td>non-restr.</td><td>post-nom.</td><td>restr.</td><td>pre-nom.</td><td>restr.</td><td>post-nom.</td><td>restr.</td><td>post-nom.</td></tr>
<tr><td>man/mā</td><td></td><td></td><td>restr.</td><td>pre-nom.</td><td>restr.</td><td>post-nom.</td><td>restr.</td><td>post-nom.</td></tr>
<tr><td>ḥayṯu</td><td>non-restr.</td><td>post-nom.</td><td>restr.</td><td>pre-nom.</td><td>restr.</td><td>post-nom.</td><td>restr.</td><td>post-nom.</td></tr>
<tr><td rowspan="3">Examples</td><td colspan="2">(1) Aḷḷāhu llaḏī nazzala l-kitāba</td><td colspan="4">(7a) min-a d-dīni … llaḏī awḥaynā ilay-ka /
(7b) llaḏī ǧā’a-ka min-a l-ᶜilmi</td><td colspan="2">(4) hāḏā (RAĞULUN) man aᶜrifu munṭaliqan</td></tr>
<tr><td colspan="2"></td><td colspan="4">(2a) min-a d-dīni mā waṣṣā bi-hi Nūḥan /
(2b) mā anzala-ḷḷāhu min-a l-kitābi</td><td colspan="2">(5) hāḏā (RAĞULUN) llaḏī ᶜalimtu annī lā aᶜrifu-hu munṭaliqan</td></tr>
<tr><td colspan="2">(3) bi-Ḫayfi Banī kināna ḥayṯu taqāsamū ᶜalā l-kufri</td><td colspan="4">(8a) min arḍi-nā ḥayṯu ši’tum /
(8b) ḥayṯu šā’ū min arḍi-him</td><td colspan="2">(6) ḥayṯu takūnu akūnu (= al-makānu llaḏī takūnu fī-hi akūnu)</td></tr>
</table>

It would be also convenient if a suprasegmental re-conceptualization of the system of OA relativizers along the above lines were to shed some light on the anomalous *syntactic* distribution of the restrictive relativizers *llaḏī, man/mā, ḥayṯu*, in short, their post-nominal syntax, as summed up in (β) above. Upon initial inspection, it might seem as if such a line of inquiry could not give satisfactory results, since reconceptualizing the OA relativizers according to an opposition analytical vs. synthetical RC-head, instead of covert vs. overt RC-head, does not necessarily imply a reconceptualization of the problematic post-nominal syntax of the RCs introduced by them, as illustrated in the light grey cells of Table 8 above.

However, once we examine more closely the system of OA relativizers *llaḏī, man/mā, ḥayṯu*, the matter is far less simplistic than may appear at first glance. We gather from an attentive perusal of Table 8 that the RC introduced by a restrictive relativizer increasingly deviates from the cross-linguistic tendency of having a pre-

nominal syntax as the degree of phonological reduction of the synthetical RC-head increases: on the one hand, a *partially* reduced RC-head can co-occur with *both* a typologically expected pre-nominal RC introduced by the restrictive relativizers *llaḏī, man/mā, ḥayṯu*, and with its deviant post-nominal counterpart; on the other, a *totally* reduced RC-head occurs *only* with a deviant post-nominal RC introduced by the relative markers in question.

On these grounds, it seems sensible to hypothesize that a correlation exists between the suprasegmental phenomenon of phonological reduction undergone by a RC-head, and the deviant post-nominal syntax of a RC introduced by a restrictive relativizer – a sort of direct proportion. It also seems that this correlation, because of its salience within the system of OA relativizers, is too significant to be dismissed so easily, and can be accordingly taken into consideration in order to understand the anomalous properties of the system at issue, such as (β); thus, contrary to what naïve impression suggests, the suprasegmental phenomenon of phonological reduction undergone by a RC-head *can* serve as an interpretive key to understand the deviant post-nominal syntax of the restrictive relativizers *llaḏī, man/mā, ḥayṯu*.

To this end, we can conveniently begin by discussing the directionality of the correlation under scrutiny, which in principle can be either from suprasegmental phonology to syntax – i.e. the *min* > *mi* or N > $ø_N$ change somehow triggers the RC + N > N + RC word-order shift – or the other way around. However, cross-linguistic evidence rules out the latter hypothesis in favor of the former: according to Pullum and Zwicky (1988:274-275, 278), syntax has no influence on phonology, whereas phonology can influence syntax under certain conditions, among which is "a preference for some form of expression over another". For instance, some English speakers prefer the binary syntactic structure: particle + nominal, e.g. (*the shock touched) off the explosion* over the binary syntactic structure: nominal + particle, e.g. *(the shock touched) the explosion off* when the binary *syntactic* structure: particle + nominal features a noun *phonologically* 'heavier' (= longer) than the particle (Pullum and Zwicky 1988:274, Ross 1967:48). There are four aspects to this kind of phonological influence over syntax:

Firstly, its ultimate cause is 'heaviness', a phonological constraint that in many respects overlaps with word-length and, as such, is a suprasegmental phenomenon, word-length being a multifactorial entity that involves syllable-structure etc. (cp. also Hawkins 1983: 90). In this light, binary syntactic structures such as (*the shock touched) off the explosion* arise since they obey a suprasegmental pattern of increasing word-length/heaviness – as opposed to syntactic structures such as (*the shock touched) the explosion off*, which do not.

In second place, the way syntax complies with the aforesaid phonological constraint is in essence a change in word-order, by virtue of which a phonologically-*un*conditioned syntactic structure, e.g. (*the shock touched) the explosion off*,

with decreasing word-length, is scrambled into a phonologically conditioned syntactic structure, e.g. (*the shock touched) off the explosion*, with increasing word-length(cp. also Hawkins 1983:90).

Thirdly, the phonologically-conditioned syntactic structure is a tendency, not a norm, so that its phonologically-*un*conditioned counterpart may well coexist alongside it.

Last but not least, the coexistence of two binary syntactic structures fails to come about, and just one of them occurs instead, when the nominal combined with the particle is a pronoun, since the structure: pronoun + particle is grammatical, as in *(I called) him up*, while the structure: particle + pronoun is not, as in **(I called) up him* – this is precisely as argued by (Ross 1967:48), but not cited by Pullum and Zwicky (1988) in this respect. Although Ross does not clarify the rationale behind this phenomenon, in all likelihood this has intuitively to do with the ability of the grammatical binary structure: pronoun + particle to assign the pronoun a covert realization, as opposed to the inability of its ungrammatical counterpart to do so: this is shown by syntactic paradigms such as *(I want) $Ø_{me}$/you/him to (…)* etc., where the expression *$Ø_{me}$ to* actually is a structure: pronoun + particle (cp. *(I called) him up*) with a covert pronoun *$Ø_{me}$*. The linguistic reality of the structure: pronoun + particle with a covert pronoun is diagnosed, insofar as the covert pronoun is concerned, by the presence, in the paradigm at issue, of the overt object pronouns *you/him* etc. and, concerning the particle, by the clear non-prepositional semantics of the infinitive *to* that follows the (c)overt pronouns, which qualifies it precisely as a particle[40].

Consequently, the property under scrutiny can be reformulated in a more general manner by stating that the coexistence of two binary syntactic structures exceptionally fails to take place, when either of its two constituents has a covert realization, in which case only the binary structure displaying the covert constituent occurs. This happens probably because a structure of this sort, in manifesting just *one overt* word, straightforwardly solves the problem of a possible decreasing word-length in a very economical fashion, by eliminating the factor responsible for it, namely a sequence of *two overt* words.

Based on these four properties, the construction: particle + nominal, e.g. (*the shock touched) off the explosion* can be regarded as a syntactically basic structure – its basicness being diagnosed by the transitive construction VO -, and the construction: noun + particle as a structure derived from it via scrambling (‘how’),

[40] In the generative literature, the covert nature of this constituent is generally accounted for in terms of a sort of economy principle, the so-called ‘deletion under identity’: since in a hypothetical construction such as **I want me to …* the former personal pronoun is identical in content with the latter, it is felt as superfluous and can be deleted (see e.g. Ross 1967:434-435)

due to a phonological – more accurately, suprasegmental – constraint ('why'), with the caveat that the presence of a covert constituent within either of these two constructions blocks the manifestation of the other.

Having established that cross-linguistically the correlation between phonology and syntax manifests itself in terms of one's influence over the other, the hypothesis can be entertained that the correlation observed in OA relativization between these two linguistic components can be accounted for in the same manner – i.e., that the phonological reduction undergone by the OA RC-head is the factor responsible for the post-nominal syntax of its restrictive RC – if and only if the four structural properties found in the cross-linguistic manifestation of this kind of correlation are also found in its OA counterpart.

That the correlation between phonology and syntax observed in OA relativization is possessed of the first property (heaviness) is shown by the presence in this language of one morphological type and two syntactic types, which follow the Head + Dependent pattern[41]. They are, respectively, the prefix + stem type instantiated by the synthetical RC-head (#*mi(n)*-N#), the N + RC type that includes a post-nominal restrictive RC and, *outside* relativization, the ADP(osition) + N type, traditionally referred to as a P + N construction. In particular, the ADP+N type is defined here as *not* manifesting itself in OA relativization since an approach of this sort would entail that the analytical RC-head #*min*#N# out of which the prefix + stem type *#mi(n)*-N# arises is an ADP + N type, which is actually not the case, if we recall from the end of Sect. 4 that the P-analysis (or, typologically speaking: ADP-analysis) of *min* is untenable for semantic reasons (inability to receive a partitive reading). In passing, this is also indirect evidence – by means of an argument by exclusion – that the analytical RC-head #*min*#N# is rather an instance of Dependent + Head pattern, whatever the precise nature of its syntactic structure (to which we'll return in the next section).

The bearing that these three types have upon the property of heaviness is that, according to Hawkins (1983:21-22, 95-96, and cp. also Table 9 below, under II), they are defining characters of the so-called prepositional languages, which "are placing 'lighter' constituents to the left of the head, and 'heavier' ones to the right", so that OA relativization displays an increasing word-length both in the word-order N + RC, whose RC can be restrictive, and in the prefix + stem type to which its synthetical RC-head belongs, in compliance with the heaviness constraint discussed in connection with the English data.

[41] The assignment of Head and Dependent status is justified by a substitution test: a complex constituent (e.g. *poor John*) can be substituted by a Head (*John*), but not by a Dependent (**poor*). The PP complies with this generalization, considering that in languages such as English the PP *in the house* can be substituted by *in* but not by *the house* (cp. *John is in the house, John is in, *John is the house*). See Comrie (1989:94ff.) for details.

Table 9: Nominal Domain – Two Word-Order Universals

Syntax			Morphology	Implicational Universal
Types		Generalization	Type	
N + Postposition	RC + N	Dependent + Head	Stem + Suffix	I
Preposition + N	N+ RC	Head + Dependent	Prefix + Stem	II

Concerning the second and third property (shift from a phonologically-unconditioned to a phonologically-conditioned word-order, and coexistence between them), they are easily observed in the portion of syntactic paradigm that correlates with the *partial* phonological reduction of the RC-head, where the change in word-order is self-evident in the alternation between the pre-nominal and post-nominal syntax of the restrictive RC (cp. Table 8, columns 4-8). Furthermore, we know from the discussion of this instance of post-nominal and restrictive RC (N + RC) that its constituents are informed by a suprasegmental pattern of heaviness, which allows us to interpret it as a phonologically-conditioned structure, whereas its pre-nominal and restrictive counterpart (RC + N) represents a phonologically *un*conditioned structure, for the reason that it does not obey the pattern in question, as can be quickly inferred from the decreasing word-length of its constituents.

In keeping with a typological approach, we can refine this statement in two steps. The first is insisting on the fact that, within OA relativization, the RC + N type – whose RC is *only* restrictive – follows a Dependent + Head pattern just as – *prior* to phonological reduction (cp. immediately above) – its analytical RC-head does. The second step is that, outside relativization, OA bears the marks of a postpositional construction, i.e. an N + ADP type, a phenomenon that has gone largely unnoticed in the literature. In a line of poetry quoted by Hišām al-Anṣārī, in fact, the expression *ᶜalā ᶜan* occurs, whose constituents are analyzed in the Arab linguistic tradition either as ADPs or Ns, depending on their distribution: in greater detail, Arab Grammarians interpret *ᶜalā* and *ᶜan* as ADPs if they are the first member of a PP, the second being a *bona fide* N, such as a declensed N; and as Ns, if they are the second member of a PP, the first being the *bona fide* P *min* – whose purely adpositional nature is diagnosed by its inability to occur itself as the second member of a PP. The pertinent descriptions and examples of this positional test are

found in *Lisān* VI, 4281[42] and *Muġnī* I, 290[43]. The latter source also makes it possible to disclose a distributional asymmetry between *ᶜalā* and *ᶜan*: the description of such bicategorial words that Ibn Hišām al-Anṣārī makes under the lemmata *ᶜalā*, *ᶜan*, *min* (*Muġnī* I, 286, 297-300, 613-614) states that both of them are semantically synonymous to *min* in certain contexts, but among the examples he mentions to elucidate this point, those concerning the synonymity between *min* and *ᶜan* include cases in which *ᶜan* can replace *min* on a syntactic level (e.g. when *bayni* is the second member of a PP), whereas those concerning the synonymity between *min* and *ᶜalā* do not[44].

A substitution test emerges from these data, which can be of great help when the positional test of Arab Grammarians is not applicable to particular instances of *ᶜalā* and *ᶜan*. In effect, in expressions found in pre-Islamic poetry such as *ᶜalā ᶜan* (cp. *Muġnī* I, 300, and especially the passage *wa-l-ṯānī an yudḥala ᶜalay-hā [=ᶜan] ᶜalā* etc. reproduced in fn. 44 above) the position of *ᶜalā* relative to its syntactic context, to wit *ᶜan*, is not useful to determine the syntactic category *ᶜalā* belongs to, given that such a context consists itself of a word as bicategorial as the word it is meant to disambiguate, and the same carries over to *ᶜan* relative to its syntactic context, notably *ᶜalā*. The only alternative left is applying the above 'asymmetric' substitution test, by virtue of which the instances of *ᶜalā* and *ᶜan* found in the OA expression *ᶜalā ᶜan* are, respectively, a N and an ADP, so that, from a typological perspective, such an expression manifests a postpositional construction (N + ADP), which follows a Dependent + Head pattern, not unlike the pre-nominal RC and the analytical RC-head observed in OA relativization. This is illustrated in Table 9, under I.

A further similarity among these three syntactic types concerns their diachrony: the hypothesis that the analytical RC-head (stem + suffix type) is older than its

---

[42] *wa-tudḥilu min ᶜalā ᶜan wa-lā tudḥilu ᶜan ᶜalay-hā li-anna ᶜan-i smun wa-min min-a l-ḥurūf*

[43] *wa-ṯ-ṯānī min waǧhay ᶜalā an takūna-sman bi-maᶜnā fawqa wa-ḏālika iḏā daḫalta ᶜalay-hā min*

[44] *wa-la-hā [=ᶜalā] tisᶜatu maᶜānin … as-sādisu muwāfiqatu min naḥwa iḏā-ktālū ᶜalā n-nāsi yastawfūna … ᶜan ᶜalā ṯalāṯati awǧuhin aḥadu-hā an takūna ḥarfa ǧarrin wa-ǧamīᶜu mā ḏukira la-hā ᶜašratu maᶜānin … as-sābiᶜu murādifatun min naḥwa wa-huwa llaḏī yaqbalu t-tawbata ᶜan ᶜibādi-hi wa-yaᶜfū ᶜan-i s-sayyi'āti … al-ᶜāširu an takūna zā'idatan li-t-taᶜwīḍi min uḫrā maḥḏūfatun ka-qawli-hi … fa-hallā llatī ᶜan bayni ǧanbay-ka tadfaᶜu … aṯ-ṯāliṯu an takūna-sman bi-maᶜnā ǧānibin wa-ḏālika yataᶜayyanu fī ṯalāṯati mawāḍiᶜa aḥadu-hā an yudḫala ᶜalay-hā min … wa-yaḥtamilu-hu ᶜindī ṯumma la-ātiyanna-hum min bayni aydī-him … wa-ṯ-ṯānī an yudḫala ᶜalay-hā ᶜalā wa-ḏālika nādirun wa-l-maḫfūẓu min-hu baytun wāḥidun wa-huwa qawlu-hu ᶜalā ᶜan yamīni marrati ṭ-ṭayru sunnaḥan … min ta'tī ᶜalā ḫamsati ᶜašara wağhan … as-sādisu murādifatu ᶜan naḥwa fa-waylun li-l-qāsiyati qulūbu-hum min ḏikri llāhi … al-ḥādī ᶜašara murādifatun ᶜalā naḥwa naṣarnā-hu min-a l-qawmi*

synthetical counterpart (prefix + stem type) can be arguably extended to the pre-nominal RC (RC + N), with respect to its post-nominal counterpart (N + RC), as well as to the postpositional construction (N + ADP), with respect to its prepositional counterpart (ADP + N), on the grounds of productivity. More accurately, in OA the pre-nominal RC is less productive than its post-nominal counterpart, since one is exclusively restrictive in meaning, the other is also non-restrictive (cp. Table 8 above), whereas the well-known productivity of the prepositional construction in OA sharply contrasts with the unproductive nature of its postpositional counterpart, explicitly stated by Ibn Hišām (*wa-ḏālika nādir*, cp. fn. 43 above).

Remarkably, a structural and diachronic cohesion of this sort (i.e., Dependent + Head pattern and productivity) between the syntactic types found in OA relativization (i.e., analytical RC-head #*min*#N#, and RC + N) and its postpositional type (i.e., *ᶜalā ᶜan*) clearly point, in Hawkins' typological terms, to an early stage of OA which was shaped by all of these types in the form of a postpositional language, which Hawkins (1983:96) also defines as more free than a prepositional language relative to the suprasegmental pattern of heaviness: "postpositional languages have some heavier constituents to the right [of the head] with lighter ones to the left, and conversely some lighter constituents to the right with heavier ones to the left".

Hawkins' generalization unproblematically applies to the early stage of OA under discussion, where the analytical RC-head #*min*#N# obeys the suprasegmental pattern of heaviness (monosyllabic *min* vs. trisyllabic N), whereas the RC + N type and N + ADP type – as exemplified by the expression *ᶜalā ᶜan* – clearly do not, so that the above statement that the RC + N type is a phonologically *un*conditioned structure is empirically grounded in a robust cross-linguistic tendency of the word-order universal labeled as I in Table 9 above.

Lastly, the fourth property (presence of a covert constituent in one word-order, blocking the manifestation of the other) can be detected directly in the portion of syntactic paradigm that correlates with the *total* phonological reduction of the RC-head (Table 8, columns 9-10): here, in fact, the RC-head coming out of the process in question, in addition to being covert, is able to be combined with a RC that follows it, but not with one that precedes it –in sharp contrast to its partially reduced counterpart[45].

The foregoing is the typological basis for the hypothesis that the direct proportion linking, in OA, the phonological reduction undergone by a RC-head and the

[45] Therefore, the process that converts an OA RC-head into a covert constituent (phonological reduction) is different from that yielding its English counterpart, e.g. *(I want) Ø$_{me}$ to (…)* (deletion under identity, cp. fn. 40 above). However, in another instance of OA RC-head, namely the N of time, a covert constituent may arise subsequent to the application of deletion under identity: cp. Bravmann (1961:391).

post-nominal position of its restrictive RC is the result of the former's influence over the latter: the restrictive RC occupying a pre-nominal position is a basic structure, which the suprasegmental phenomenon of phonological reduction undergone by the RC-head converts into a restrictive RC occupying a post-nominal position, and this by influencing syntax, and inducing scrambling accordingly.

The details of this proposal can be expounded as follows: prior to the application of phonological reduction, an analytical RC-head #*min*#N# occurs, which is structurally harmonic with the RC + N type, both of these word-orders following the Dependent + Head pattern implied by a postpositional language. After that phonological reduction takes place, the analytical RC-head #*min*#N# becomes a synthetical RC-head #*mi(n)*-N#, i.e. a prefix + stem type, which is *not* structurally harmonic with the RC + N structure, one word-order being implied by a prepositional language, the other by a postpositional language (cp. Table 9). To make things worse, the new word-order pattern is more constrained by heaviness than the previous one, so that the disharmonicity of the RC + N type with respect to the prefix + stem type also holds on a suprasegmental level. In the wake of Hawkins (1983:182) a principle of structural analogy intervenes at this point, in order to harmonize the RC + N type to the prefix + stem, and the change in word-order applies, which results in the N + RC type, or, less formally, in a restrictive RC occupying a post-nominal position[46].

Shortly put, in OA a principle of structural analogy is the 'missing link' between the phonological reduction undergone by the RC-head, and the post-nominal syntax of its restrictive RC, with the word-order universal labeled as II in Table 9, serving as an analogical pivot, because of its including the structural description of both of them, in the form of a prefix + stem type (RC-head) and N +RC type (restrictive RC).

We can draw from this typological study of the correlation between phonology and syntax, as observed in OA relativization, three important distributional generalizations, each of which corresponds to a process and related output to which such a correlation is input:

Process and Related Output#1: A suprasegmental pattern of heaviness occurs (e.g. English type N + RC; OA *partially* reduced RC-head #*min*-N#), which does

[46] Arguably, the same structural analogy operates in English, and facilitates a heaviness-driven word-order change such as *(the shock touched) the explosion off → (the shock touched) off the explosion.* This happens probably because such a principle assimilates the particle *off* to a P via their shared phonological shape (cp. *off balance*), whence the re-analysis of the structure *off the explosion* as a P + N type. In doing so, analogy makes expressions such as *off the explosion* harmonic with the N + RC type (both of them belonging to the Head + Dependent pattern), and therefore preferable to a disharmonic structure such as *the explosion off.*

not trigger syntactic scrambling (e.g. English *(the shock touched) the explosion off*; OA restrictive and pre-nominal RC).

Generalization#1: when no influence of heaviness-driven phonological reduction over syntax takes place, the *restrictive* relativizers *llaḏī*, *man/mā*, *ḥayṯu* only occur in a pre-nominal position, in complementary distribution with their non-restrictive counterparts, which only occur in a post-nominal position.

Process and Related Output#2: A suprasegmental pattern of heaviness occurs (see immediately above), which triggers syntactic scrambling (e.g. English *(the shock touched) off the explosion*; OA restrictive and post-nominal RC), through the mediation of a structural analogy concerning word-order (see fn. 46 above for English, and, in OA, the typological harmonicity between the suprasegmentally appropriate N + RC type and the prefix + stem type instantiated by a partially reduced RC-head)

Generalization#2: when heaviness-driven phonological reduction influences syntax, the *restrictive* relativizers *llaḏī*, *man/mā*, *ḥayṯu* can occur in a post-nominal position, just as their non-restrictive counterparts do, which disrupts the pattern of complementary distribution.

Process and Related Output#3: A suprasegmental pattern of heaviness occurs (e.g. English type N + RC; OA *totally* reduced RC-head), which triggers syntactic scrambling, but the latter is in turn successful only in part, since it instantiates only the binary structure that includes a covert constituent for reasons of economy (e.g. English *(I) want* $Ø_{me}$ *to (…)*; OA restrictive and *exclusively* post-nominal RC headed by a covert N)

Generalization#3: see *Generalization#2*.

These generalizations show that the anomalous *syntactic* distribution of the restrictive relativizers *llaḏī*, *man/mā*, *ḥayṯu*, which basically boils down to their post-nominal syntax (cp. (β) at end of Sect. 5), can be explained, as anticipated above, by simply invoking the same suprasegmental reconceptualization of the system of OA relativizers that explains the anomalous *semantic* distribution of the restrictive relativizer *man/mā*, and revolves around an opposition between an analytical and a synthetical RC-head, in terms of lack vs. presence of the ability to undergo phonological reduction.

It is precisely the presence of the suprasegmental feature of phonological reduction, as encoded in the synthetical RC-head, in fact, that causes such an anomalous post-nominal distribution of its restrictive RC. As a corollary, prior to the feature of phonological reduction encoded in the synthetical RC-head influencing the original pre-nominal syntax of its restrictive RC, a clear semantic and syntactic pattern of complementary distribution exists in the system of OA relativizers, as illustrated in the light grey cells of Table 10 below, which abstracts

away from this kind of influence by marking the semantic and syntactic structures involved in it in strikethrough[47].

Table 10: OA Relative Markers – Distribution (Refined Version)

RC-head	Phonology							
	Analytical		Synthetical					
			Partial Reduction				Total Reduction	
Relativizer	Sem.	Synt.	Sem.	Syntax	Sem.	Syntax	Sem.	Syntax
llaḏī	non-restr.	post-nom.	restr.	pre-nom.	~~restr.~~	~~post-nom.~~	~~restr.~~	~~post-nom.~~
man/mā			restr.	pre-nom.	~~restr.~~	~~post-nom.~~	~~restr.~~	~~post-nom.~~
ḥayṯu	non-restr.	post-nom.	restr.	pre-nom.	~~restr.~~	~~post-nom.~~	~~restr.~~	~~post-nom.~~
Symbols	X =Complementary distribution				~~X~~ =heaviness and analogy ('harmonicity')			
Examples	(1) *Aḷḷāhu llaḏī nazzala l-kitāba*				(7a) *min-a d-dīni … llaḏī awḥaynā ilay-ka /* (7b) *llaḏī ǧā'a-ka min-a l-ᶜilmi*		(4) *hāḏā (RAĞULUN) man aᶜrifu munṭaliqan*	
					(2a) *min-a d-dīni mā waṣṣā bi-hi Nūḥan /* (2b) *mā anzala-ḷḷāhu min-a l-kitābi*		(5) *hāḏā (RAĞULUN) llaḏī ᶜalimtu annī lā aᶜrifu-hu munṭaliqan*	
	(3) *bi-Ḫayfi Banī Kināna ḥayṯu taqāsamū ᶜalā l-kufri*				(8a) *min arḍi-nā ḥayṯu ši'tum /* (8b) *ḥayṯu šā'ū min arḍi-him*		(6) *ḥayṯu takūnu akūnu (= l-makānu llaḏī takūnu fī-hi akūnu)*	

In conclusion to this section, an in-depth, typological investigation of the counterexamples that militate against a semantic analysis of the system of OA relativizers *llaḏī*, *man/mā*, *ḥayṯu* in terms of complementary distribution reveals that such counterexamples disappear if we add to the picture phonological and syntactic levels of analysis (phonological reduction, heaviness, scrambling, analogy/harmonicity etc.). What's more, this investigation also reveals that the phonological and syntactic data explored by it can be felicitously integrated with the semantic ones into a pattern of complementary distribution.

[47] This table rewrites Table 8.

From a broader linguistic perspective, however, these results are flawed in two major respects, since, while describing the phonology, syntax and semantics of the system of OA relativizers, they say nothing about the latter's morphology and pragmatics.

The issue is even more prominent, if we consider that it is precisely these two level of linguistic description that have traditionally attracted the attention of Western Arabists and Semitists, albeit to a different extent (cp. Sect. 0, and especially the discussion there about the quite incomplete pragmatic definition of *la* as a 'reinforcer'): accordingly, one may wonder whether and to what extent the analysis of the OA relativizers, as developed thus far, can be reconciled with more familiar analyses of them, à la Barth, Brockelmann, Reckendorf etc.

One example that well illustrates this point is the contrast between the relative constructions **llaḏī* + RC + N (cp. Table 5) and *llaḏī* + RC + *min* + N (cp. (7b) in Table 10). While the current analysis would emphasize that the syntax of the structure **llaḏī* + RC + N significantly improves after *min*-insertion, a morphological and pragmatic analysis would rather point out that the (bound/free) morpheme *min* is able to co-occur not only with the relative markers *man/mā, ḏī*, but also with the reinforcer *la*, so that the question arises of whether these analyses are mutually exclusive or not. The next section means to bridge such an interpretive gap between the current and more traditional approaches, by focusing on the morphological and pragmatic dimensions of the relative markers *llaḏī, man/mā, ḥayṯu*.

7. Morphological and pragmatic implications of a semantic analysis of Old Arabic relative markers

Despite their dealing with the OA relative markers under the heading of morphology, traditional reference works such as Brockelmann's (1910), Barth's (1913), and Fleisch's (1961) actually tend to incorporate in their description some pragmatic considerations, at least in embryonic form.

A case in point is *llaḏī*, which Barth (1913:79, 156-157), decomposes into the smaller morphemes *l, la, ḏī*, relying upon distributional criteria such as the occurrence of these strings within other words. In particular, Barth's interpretation of *la* in terms of reinforcement, in line with Brockelmann (1910, cp. also Sect. 0) is inferred from the distributional context that the German scholar mentions in connection with this particle, in order to exemplify its occurrence outside the relativizer *llaḏī*. In fact, after characterizing such a context as the predicate of a nominal sentence introduced by *inna* (e.g. *inna rabba-ka la-yaḥkumu bayna-kum* 'Surely thy Lord will decide between them', *Koran* XVI, 124), Barth assigns to the instance of *la* found in it a reinforcer status in his German translation, where the

construction *inna* + Subject + *la* + Predicate is rendered as *siehe* + Subject +*da* + Predicate.

Pursuing a pragmatic line of reasoning, it is precisely the combination of *la* with a predicate that shows, according to Testen (1998:72–73) that this instance of *la* "was a marker of 'assertion'", based on cross-linguistic evidence: in English, for instance, the assertive marker *do* (e.g. *it does snow in May*) equally combines with a predicate. Therefore, Testen's observation refines the pragmatic approach to *la* in the manner of Barth etc. by describing it as a marker of assertion, where the latter term is practically synonymous with 'new information', as highlighted by Lambrecht (1994:52, emphasis in original): "let us refer to the 'new information' expressed or conveyed by the sentence as the PRAGMATIC ASSERTION (or simply the ASSERTION)."

Lambrecht's definition of assertion finds its *raison d'être* in the fact that the predicate to which the English *do* – and, by extension OA *la* – combine typically denotes new information, and entails that what Testen labels as a 'marker of assertion' actually is a new information marker or focus-marker. Furthermore, Arab Grammarians label the *la* combined with a predicate as *lām at-tawkīd/ta'kīd*, and accordingly their notion of *tawkīd/ta'kīd* corresponds to a good extent to the Western notion of assertion (Testen 1998:72), so that Lambrecht's definition of assertion in terms of new information allows us to define *tawkīd/ta'kīd*, by transitive property, as the expression of new information.

In this regard, the instance of *la* that does not co-occur with *inna* and combines instead with a subject, thus signaling old – rather than new – information (e.g. *la-yūsufu wa-aḫū-hu aḥabbu ilā abī-nā min-nā* 'Surely Joseph and his brother are dearer to our father than we', *Koran* XII, 8) is not a serious counterexample to the focus-marker analysis advocated here, given that the combination of this kind of *inna*-less *la* with a subject in OA is far less rigid than it happens in Classical Arabic. This is shown by lines of pre-Islamic poetry such as *ummu l-ḥulaysi la-ᶜağūzun šahraba* 'U. Ḥ is an old woman, a *šahraba*' (*apud* Testen 1998:11ff), where interestingly the *inna*-less *la* combines with a predicate, as totally expected under a focus-marker analysis.

Having discussed the pragmatics of *llaḏī*, we can now turn our attention to its morphology, whose traditional analysis appears to be acceptable only in part. True, Barth's etc. analysis of *ḏī* as a relative marker is confirmed by a substitution test reported, curiously enough, by the historian Bahā' ad-Dīn al-Ğanādī (d. 732/1332), where *ḏī* replaces *llaḏī* in the context *lā budda min __ḥakama l-amīr* (*apud* Kay 1892:147)[48,49].

[48] *fa-qālat dū budda min ḏī ḥakama l-amīr wa-dū bi-d-dāli l-muhmalati fī luġati baᶜḍi l-yamāniyyīna bi-maᶜnā lā fa-ka-anna-hā qālat lā budda min ḏī ḥakama l-amīr wa-ḏī bi-ḏ-*

By contrast, once we take into due consideration the above analysis of *la* as a focus-marker, we can hardly subscribe to Barth's etc. bimorphemic analysis of the string *lla*, according to which a definite article *l* is cumulated with the particle *la*: one, in fact, by definition encodes old information, and as such clearly conflicts in pragmatics terms with the other, conveying new information instead. To this, we can add that such pragmatic considerations converge with a distributional argument found in primary sources, which rule out the possibility that the segment *l* observed at the left edge of *llaḏī* is a definite article in view of the fact that the latter in OA is generally in complementary distribution with a *tanwīn*, whereas the initial *l* in *llaḏī* is not, the pair *r-raǧulu/raǧulun* being not paralleled by the pair *llaḏī/laḏin* (see e.g. *Šarḥ al-Mufaṣṣal* II, 372-374)[50].

On these grounds, the string *lla* of *llaḏī* is better understood as a syntactically-conditioned allomorph of the focus marker *la*, notably as an instance of *la* that has undergone gemination due to its clause-initial position, considering that grammatical markers such as *anna, inna* that introduce an (object) clause are geminated in OA, and that the string *lla*, itself introduces a (relative) clause – unlike its allomorph *la*[51].

In sum, a critical review of the traditional approach to the morphology and pragmatics of *llaḏī* proves that it is made up of a relative stem *ḏī* and a geminated allomorph of the new information (= focus) marker *la*, a finding that can be easily reconciled with the analysis of *llaḏī* as a non-restrictive relativizer worked out in the previous sections, if we recall from Comrie (1989:138-139) that a non-restrictive RC conveys precisely new information. Shortly put, the focus-marker *lla* in *llaḏī* signals the presence of a post-nominal and non-restrictive RC – pending further discussion on why it also unexpectedly introduces a restrictive RC (cp. (7) in Table 10 above).

ḏāli l-muᶜǧamati bi-maᶜnā llaḏī ka-anna-hā qālat lā budda min-a llaḏī ḥakama l-amīr yaᶜnī bnu faḍl

[49] As Rabin (1951:39) remarks, the mention of the governor Ibn Faḍl relates the constructions *lā budda min ḏī ḥakama l-amīr* etc. to events of the 9th century CE, which ensures that these data belong to OA.

[50] *wa-aṣlu llaḏī laḏin ka-ᶜamin šaǧin fa-l-lāmu fā'u l-kalimati wa-ḏ-ḏālu ᶜaynu-hā wa-l-yā'u lāmu-hā hāḏā maḏhabu l-baṣriyyīna wa-qāla l-kūfiyyūna l-aṣlu fī-llaḏī ḏ-ḏālu waḥda-hā wa-mā ᶜadā-hā zā'idun fa-aṣlu llaḏī ka-aṣli hāḏā […] anna l-alifa wa-l-lāma fī l-mawṣūlāti ziyādatun lāzimatun wa-lāmu t-taᶜrīfi lā naᶜrifu-hā ǧā'at lāzimatan bal yaǧūzu isqāṭu-hā naḥwa r-raǧulu wa-l-ġulāmu wa-raǧulun wa-ġulāmun wa-lam naǧid-hum qālū laḏin ka-mā qālū ġulāmun fa-lammā ḥalafat mā ᶜalay-hi naẓā'iru-hā dalla ᶜalā anna-hā zā'idatun li-ġayri maᶜnā t-taᶜrīf*

[51] Cp. also Rabin (1951:155) for comparative arguments in favour of interpreting *lla* as a geminated allomorph of *la*.

Turning to *ḥayṯu*, the most exhaustive morphological treatment of this relativizer is to be found in Fleisch (1961:II, 61-63), who decomposes it into the smaller meaning units *ḥay, ṯ, u*, which he identifies with the demonstrative stems *ḥay, ṯ*, and the locative stem *u* on the basis of the same distributional criteria illustrated in connection with Barth. As in the case of *llaḏī*, primary sources lend further support to the Western traditional analysis of *ḥayṯu* by means of a substitution test reported in *Kitāb* IV, 233, which sets up a one-to-one correspondence between *ḥay* and *lla*; *ḏī* and *ṯ*; *u* and *fī-hi*[52]. The test at issue also allows for a more accurate characterization of *ḥay* and *ṯ* on a pragmatic, and more generally semantic level, by assigning them the same content as the morphemes they replace, with the consequence that they qualify, respectively, as a focus-marker and a relative marker.

Concentrating on the pragmatic dimension of *ḥay*, it is also captured in its essence in the morphological treatment that Fleisch (1961: II, 68-69) offers of the so-called 'formes à diphtongue', which all follow the pattern *Cay*, and include the particle in question. After observing that these morphemes (*ay-, kay-, hay-, nay-, ḥay-, ḏay-*) tend to be subject to iteration (e.g. *ḏay-ta wa-ḏay-ta* 'so', *kay-ta, wa-kay-ta* 'id.'), the French scholar states: "les formes à diphtongue s'expliquent suffisamment par la progression phonétique, d'origine affective, pour la recherche de l'expressivité".

In all likelihood, Fleisch's observations heavily draw on primary sources, where iteration is said to be a strategy to express *tawkīd/ta'kīd* (see e.g. *Šarḥ al-Mufaṣṣal* II, 219–220)[53]. The point we would like to stress here is that the ability of a given OA morpheme to undergo iteration diagnoses its conveyance of new information, given the above result that a fairly reasonable correspondence exists between *tawkīd/ta'kīd* and new information – rather than between *tawkīd/ta'kīd* and a quite vague notion of 'expressivité'. In consequence of this, *ḥay* turns out to be a focus-marker since it successfully replaces *(l)la*, and belongs to a class of forms able to undergo iteration, a morphological reflex of new information.

To summarize, as it were, this critical review of the traditional approach to the morphology and pragmatics of *ḥayṯu*, primary sources improve it by substantiating the hypothesis that *ḥay* is a new information (= focus) marker, and *ṯ* a relative stem. As in the case of *llaḏī*, such a finding can be reconciled with the analysis of *ḥayṯ(u)* as a non-restrictive relativizer presented in the previous sections, by simply recalling from Comrie (1989:138-139) that a non-restrictive RC precisely conveys new information. Briefly, the focus-marker *ḥay* in *ḥayṯ(u)* signals the presence of a post-nominal and non-restrictive RC, with the proviso that its ability to signal a restrictive RC (cp. (8) in Table 10 above) still calls for an explanation.

---

[52] *wa-ammā ḥayṯu fa-makānun bi-manzilati qawli-ka huwa fī l-makāni llaḏī fī-hi Zayd*

[53] *wa-t-ta'kīdu ᶜalā ḍarbayni lafẓiyyun wa-maᶜnawiyyun wa-l-lafẓiyyu yakūnu bi-takrīri l-lafẓi wa-ḏālika naḥwa qawli-ka ḍarabtu Zaydan Zaydan*

We are thus left with the morphology and pragmatics of the relativizer *man/mā*, and of the particle *min* found in OA relativization. Starting from morphological considerations, *man/mā* has been extensively dealt with in traditional reference works, which assign it a bimorphemic status (*ma-n*, *ma-:*) by means of the same distributional criteria used for *llaḏī* and *ḥayṯ(u)* (cp. the interrogative *man/mā*), and nothing new can be added here. By contrast, the traditional analysis of the particle *min* found in OA relativization, as carried out in such works, invites further discussion, since it awkardly identifies the particle in question with the P *min*, on the basis of the wrong assumption – not corroborated by textual evidence – that the former instance of *min* shares with the latter a partitive meaning (see e.g. Reckendorf 1895: 622-623). As already pointed out at end of Sect. 4, in fact, OA relative constructions such as e.g. *min-a d-dīni mā waṣṣā bi-hi Nūḥan* (*Koran* XLII, 13) and *mā nansaḫu min āyatin* (*Koran* II, 106), which will be referred to henceforth as *mā-min* relative constructions, are, according to Arab Grammarians, *loci probantes* for interpreting instead the particle *min* occurring in OA relativization as a *zā'ida/ṣila/laġw*, i.e. a semantically unnecessary constituent (see fn. 29 above on this terminological equivalence).

Far from confining such an interpretation to the *min* found in the OA relative construction, Arab Grammarians, and especially the OA native speakers among them, assigned the morphemic status of semantically unnecessary constituent also to the *min* found in a particular kind of OA negative construction, generally defined in the literature as formally identical to the *mā-min* relative construction, in that the negation *mā* co-occurs with a negated NP combined precisely with *min* (*mā-min* negative construction henceforth).

Thus, in his metalinguistic description of *min*, al-Farrā' (d. 207/822) recognizes that it can denote departure from a place, a partitive expression, or a semantically unnecessary constituent (*ṣila*), and exemplifies the latter meaning through the OA *mā-min* negative construction *mā yuᶜzabu ᶜan rabbi-ka min miṯqāli ḏarratin* 'not so much as the weight of an ant […] escapes from thy Lord' (*Koran* X, 61), which he glosses as *mā yuᶜzabu ᶜan ᶜilmi-hi waznu ḏarratin*, dropping the particle *min* altogether (*Lisān* VI, 4281)[54].

A digression is in order here. If we concur with al-Farrā' that the *min* found in the OA *mā-min* negative construction is not a partitive *min*, we can reassess the traditional observation that this construction is formally identical to the *mā-min* relative construction by stating that it consists of a NP – instead of a PP – and a clause introduced by a grammatical marker *mā* that can receive – besides a negative marker interpretation – a relative marker interpretation. Note-worthily,

[54] *takūnu min ibtidā'a ġāyatin wa-takūnu baᶜḍan wa-takūnu ṣilatan qāla llāh ᶜazza wa-ǧalla mā yuᶜzabu ᶜan rabbi-ka min miṯqāli ḏarratin ay mā yuᶜzabu ᶜan ᶜilmi-hi waznu ḏarratin*

these are exactly the two syntactic properties that, according to Ouhalla (1999:342-343) diagnose the real core of a cleft-construction, as shown by the fact that two genetically unrelated languages such as English and Moroccan Arabic indeed share a NP and a relativizer in their cleft-constructions (cp. *it was the children that Nadia sent*, and *l-wlad elli huma sarrdat Nadia*, respectively), while no phonological counterpart of the Moroccan Arabic *ḍamīr al-faṣl huma* is needed in English, nor do we find in Moroccan Arabic any phonological equivalent of the English expletive-*be* complex *it is* (the lack of which Ouhalla labels as 'pruning').

With this in place, the OA *mā-min* negative construction bears the mark of a negative cleft- construction that, like its Moroccan Arabic counterpart, combines a cleft NP with the subordinate clause in the form of a relative construction, actually identical to the *mā-min* relative construction, and, still in line with Moroccan Arabic, does not realize on a *phonological* level the sememes that form the expletive-*be* complex; but, unlike its Moroccan Arabic counterpart, the OA negative cleft-construction also lacks a *ḍamīr al-faṣl* intervening between the cleft NP and the subordinate clause, as wholly expected for this variety of Arabic, where the *ḍamīr al-faṣl* cannot co-occur with an expletive-*be* complex found on a semantic level, as shown by contrasts reported by Bloch (1991:57), such as *hāḏā r-raǧulu/hāḏā huwa r-raǧulu* 'this is the man' vs. *hāḏā r-raǧulu/*hāḏā huwa r-raǧulu* 'there is the man', where the latter expression exhibits, on a *phonological* level, a deictic (*hāḏā*) that receives an expletive-*be* reading on a *semantic* level.

Concretely, it is as if the English negative construction *we're not looking for Joey* had developed out of a negative cleft-construction such as *it's not Joey we're looking for*, where *Joey we're looking for* is a relative construction (cp. *the guy we're looking for*), with the caveat that, on a phonological level, the OA negative cleft-construction 'prunes' the expletive-*be* complex and the *ḍamīr al-faṣl* found, respectively, in English and Moroccan Arabic, so that it is phonologically realized only insofar as its *mā-min* relative construction is concerned. It is exactly the covert realization of such constituents that yields the well-known overlap, on the sound-side, between the *mā-min* relative construction and its negative/cleft counterpart, which contains itself a relative construction: the corresponding syntactic structures *min* + NP $mā_{\text{RELATIVE MARKER}}$ + VP and $Ø_{\text{it is}}$ + *min* + NP + $Ø_{\text{huwa}}$ + $mā_{\text{RELATIVE/NEGATIVE MARKER}}$ + VP are effectively identical on the sound-side, where the covert constituents $Ø_{\text{it is}}$ and $Ø_{\text{huwa}}$ are not detectable.

The importance of this digression to our understanding of the *mā-min* relative construction is that typologists have long noticed a grammaticalization pattern that involves precisely the two syntactic types phonologically realized in OA through the structure: *min* + NP + *mā* + VP – namely, a negative cleft-construction that contains a relative construction, and a negative construction. According to Heine & Reh (1984:185-186) Teso, a Nilo-Saharan language, attests to the grammaticalization pattern in question, which starts with a negative cleft-construction that con-

tains a relative construction (**e-mam petero e-koto ekiŋok*, lit. 's-not Peter s-want dog', i.e. 'it is not Peter who wants a dog'), and ends up with a negative construction (*mam petero e-koto ekiŋok*, lit. 'not Peter s-want dog', i.e. 'Peter does not want a dog') through reanalysis.

This is typological evidence that in OA the *mā-min* negative construction is syntactically derived from a *mā-min* relative construction[55] contained within a negative cleft-construction, via a process of reanalysis whose details cannot be investigated here, but that affects in all likelihood an overt relativizer *man/mā* originally found in an assertive cleft-construction (cp. *it's money that I love*), by replacing it with a covert relativizer (cp. *it's Joey we're looking for*), so that *man/mā* is reanalyzed as a negative marker (cp. *it's not Joey we're looking for*), which yields a negative cleft-construction and, later, a negative construction.

Returning to primary sources and to the way they interpret the *min* that occurs in the *mā-min* negative construction; a view akin to al-Farrā'*'*s is expressed by the early grammarian al-Aḫfaš (d. 177/793), who, however, also takes a step forward. According to *Lisān* (VI, 4281) al-Aḫfaš holds that in the *mā-min* negative construction *mā ǧa^{c}ala ḷḷāhu li-raǧulin min qalbayni fī ǧawfi-hi* 'God has not assigned to any man two hearts within his breast' (*Koran* XXXIII, 4) the *min* that denotes a semantically unnecessary constituent (*laġw*) also denotes *tawkīd*[56] or, in modern terms, new information (cp. the pragmatic evidence culled from the discussion of *llaḏī*).

Therefore, the aforesaid negative construction evidences the hypothesis that the instance of *min* occurring in such a distributional context, i.e. the *min* combined with the negated NP and contained within a cleft-construction, is a focus-marker. Further proof to this effect can be brought from linguistic typology: indeed, a negated NP (e.g. *amat'ed-ile* "good reindeer") combined with a focus-marker (*-k*) has been reported for Yukagir, a Uralo-Siberian language, provided that it occurs within a negative cleft-construction, in a striking parallel to OA, e.g. *met amat'ed-ile-k el'-bun'-meng* 'it wasn't a good reindeer that I killed' (Fortescue 1996:34).

A far-reaching corollary of identifying the *min* that occurs in the OA *mā-min* negative construction with a focus-marker is that the same analysis carries over to the OA *mā-min* relative construction, given that the former is syntactically derived from the latter. A word of caution is needed here: at the present research stage, such a remark can only apply to the particle *min* that *follows* the relative stem *man*/

[55] Esseesy (2010:212-213) posits an inverse process of grammaticalization, but he admits that he has "not encountered a single study documenting a grammaticalization process" of this sort.

[56] *qāla l-Ǧawhariyyu wa-qad tadḫulu min tawkīdan laġwan […] qāla l-Aḫfašu […] mā ǧa^{c}ala ḷḷāhu li-raǧulin min qalbayni fī ǧawfi-hi innamā udḫila min tawkīdan kamā taqūlu ra'aytu Zaydan nafsa-hu*

mā in the linear order, i.e. *man*/*mā*…*min* (structure-final *min* henceforth), or, to put it differently, to the *min* that occurs within a RC-head combined with a pre-nominal RC introduced by *man*/*mā*, since the OA *mā-min* negative construction that arises out of the OA *mā-min* relative construction exhibits precisely such an ordering of the constituents *man/mā* and *min* (cp. the above Koranic examples cited by al-Farrā' and al-Aḫfaš, and Wright 1896:II, 135).

To recap, a new approach to the morphology and pragmatics of the markers *man/mā*, *min* mainly focuses on their ability to occur alongside each other both in a relative and negative construction, with the linear order *man/mā…min*, and brings textual and typological evidence that the structure-final *min* is a new-information (=focus) marker in these two distributional contexts. Such an achievement can be easily reconciled with the analysis of the synthetical RC-head #*mi(n)*-N# worked out in Sect. 6 for two reasons. Firstly, in OA this kind of RC-head precisely conveys new information, since it is combined with a restrictive, and therefore old-information RC, as pointed out by Comrie (1989:138-139). Secondly, it originally *follows* the restrictive RC in question, which implies for the constituents *man/mā*, *min* found in this instance of OA relativization precisely the linear order *man/mā*…*min*.

In brief, the focus-marker *min* that occurs in the structure-final RC-head #*mi(n)*-N# signals the presence of a pre-nominal and restrictive RC.

This analysis has the advantage of converging with the indirect evidence presented in Sect. 6 to substantiate the hypothesis that the synthetical RC-head #*mi(n)*-N# follows a pattern Dependent + Head, on the grounds that *min* falls under the category of focus-marker, which is indeed a dependent in OA, as shown by the substitution test *ḏī*/*llaḏī* discussed at the outset of this section, where the complex relativizer *llaḏī* can be substituted by the relative stem *ḏī*, but not by the focus-marker *lla* (cp. fn. 41 above).

Another advantage of an analysis of the markers *man/mā* and *min* of this sort lies in an economical and unified representation of the system of OA relativizers on a morphological and pragmatic level, which the traditional approach fails to capture. Despite their discontinuous realization, the particle *min* and the relativizer *man/mā*, in fact, parallel the composite relativizers *llaḏī* and *ḥayṯu*, in that all of them have an *isomorphic* structure that combines a focus-marker with a relative marker.

This being said, the question still remains of whether such morphological and pragmatic properties of the system of OA relativizers enter the general pattern of complementary distribution that has been shown in the previous sections to govern their phonological, syntactic and semantic properties, as summarized in the dark grey cells of Table 10 above. The answer is only partially in the affirmative: on the one side, the relativizers investigated in this section, namely the non-restrictive *llaḏī*, *ḥayṯu* (cp. (1, 3) in Table 10), and the restrictive (and discontinuous) *man/mā*

…*min* (cp. (2b) in Table 10), are clearly in complementary distribution on a morphological and pragmatic level, the word-order of the focus-marker and the relative stem in *llaḏī*, *ḥayṯu* being the opposite of their word-order in *man/mā … min*; on the other, nothing certain can be stated with regard to the pre-nominal and restrictive relative markers *llaḏī*, *ḥayṯu* (cp. (7b, 8b) in Table 10), in view of their somewhat nebulous properties on a pragmatic level. The critical review of the morphology and pragmatics of *llaḏī* and *ḥayṯu* made in this section, in fact, has highlighted that such relativizers behave in a rather contradictory manner when introducing a restrictive RC, since their focus (=*new*-information) markers *lla*, *ḥay* signal a kind of RC conveying *old* information instead. Moreover, as has already been alluded to at end of Sect. 6, in the same anomalous pragmatic context the focus-markers *lla* and *ḥay* may puzzlingly co-occur with another focus-marker – *min*.

The next section addresses these and other related problems concerning the overall distributional architecture that underpins the system of OA relativizers.

8. From structural description to functional explanation

A solution to the anomalous pragmatic behavior of the focus markers *lla* and *ḥay* occurring in the pre-nominal and restrictive relativizers *llaḏī*, *ḥayṯu* (cp. (7b, 8b) in Table 10) can be devised by looking at these focus-markers through the interpretive lens of primary sources, according to which the constituents *lla* and *ḥay* thus characterized, despite their sharing with the structure-final *min* the common status of focus-marker (*tawkīd/ta'kīd*), are opposed to it in terms of lack vs. presence, respectively, of the so-called property of *zā'ida/ṣila/laġw*, as discussed at length in the previous section.

Taking such a semantic property as a departure point for our inquiry, we observe that its presence in the structure-final focus-marker *min* systematically correlates with a phonological property of this kind of morpheme, notably phonological reduction, for the reason that both the *zā'ida/ṣila/laġw* and phonological reduction manifest themselves into one and the same domain – the synthetical RC-head #*mi(n)*-N#. We could also add that the focus-marker analysis naturally extends from the structure-final *min* to the structure-initial *min*, given that the structure-initial *min* is part and parcel of a relative construction derived via scrambling from the relative construction that features the structure-final *min* (cp. Sect. 6). Practically speaking, the structure-initial *min* occurs within a RC-head combined with a post-nominal RC introduced by *man/mā*, as exemplified by (7a, 8a) in Table 10.

Bearing this in mind, we can make sense of such a correlation by resorting to Jespersen cycle, whose presence in Arabic has been extensively reported for the

verbal system (Lucas 2007). In this interpretive scenario, when phonological reduction affects the analytical RC-head #*min*#N#, thus converting it into the synthetical RC-head #*mi(n)*-N#, it triggers the post-nominal position of its RC, as discussed in Sect. 6, and – the point made here – the semantic weakening of the focus-marker *min*, whose fusion with the N that follows it into a partially reduced RC-head #*mi(n)*-N# blocks its isolability, and therefore its semantic recognizeability. This is the first stage of Jespersen cycle, which can be also exemplified by Old French *jeo ne dis*, lit. 'I not say', i.e. 'I do not say', and especially by its phonologically reduced negation *ne* (Lucas 2007:402). In the next stage, the semantic weakening undergone by the RC-head, in turn, triggers the insertion of a new constituent, which takes over it in its function as a focus-marker. As customary for Jespersen cycle, such a replacement cannot take place in the original site of the weakened focus-marker, i.e. left-adjacent to N, since this would undesirably result in a further weakening of the new focus-marker, which accordingly is required to occupy a position right-adjacent to N. The best candidates to this role are the particles *lla(ḏī)* and *ḥay(ṯu)* that typically introduce a post-nominal RC (cp. (7a, 7b) in Table 10 above), in view of the fact that, in addition to being focus-markers, they are also right-adjacent to N. It ensues that in the restrictive and post-nominal relativizers *lla(ḏī)* and *ḥay(ṯu)* that co-occur with a partially reduced RC-head #*mi(n)*-N#, the focus-markers *lla*, *ḥay* are dependent of the RC-head in question, which *precedes* them, rather than of the relative stems *ḏī*, *ṯ* that *follow* them and, in doing so, they double the weakened focus-marker *min*, in order to strengthen it, as entirely foreseeable under an analysis in terms of Jespersen cycle (cp. the co-occurrence between *ne* and *pas* in Modern Standard French *je ne dis pas*, as per Lucas 2007:402). In the next and final stage of this process, phonological reduction is fully achieved, all the other things being equal, which entails that the focus markers *lla*, *ḥay* are still dependent of the (covert) N of a totally reduced RC-head $Ø_{\#mi(n)\text{-N}\#}$ that precedes them, but they no longer double the focus-marker *min*, which has been totally reduced, i.e. deleted, along with N (cp. (4,6) in Table 10 above), and then completely superseded precisely by *lla*, *ḥay*. This behavior is in line with Jespersen cycle, since in Modern Colloquial French the negative construction *je dis pas* is found, where *pas* no longer co-occurs with *ne* (Lucas 2007:402).

The main lesson learnt from the principled application of Jespersen cycle to the OA focus-markers *lla*, *ḥay* and *min* is that *lla* and *ḥay* are partial or total replacive morphemes of *min* when they *follow* a RC-head #*mi(n)*-N#/ $Ø_{\#mi(n)\text{-N}\#}$ that has undergone partial or total phonological reduction, i.e. when they introduce a post-nominal and restrictive RC, as in (7a, 8a) (*min*/$Ø_{min}$-…*llaḏī)*, *min*/$Ø_{min}$…*ḥayṯu*). In consequence of this, the relative stems *ḏī* and *ṯ* combined with the *replacive* focus markers *lla* and *ḥay* occur in the same distributional context as the relative stem *man/mā*, which means, technically speaking, that such relative stems are

semantically conditioned allomorphs of *man/mā*, if we consider that *man/mā* takes the form *ḏī, ṯ* when semantic weakening induces its focus-marker *min*, originally left-adjacent to the RC-head, to take the form *lla, ḥay*, right-adjacent to the RC-head. In fewer words, the post-nominal and restrictive relativizers *llaḏī* and *ḥayṯ(u)* are allomorphs to the discontinuous restrictive relativizer *min…man/mā*, because of the Jespersen cycle.

It seems also safe to adopt the same analysis for the morphemes *lla, ḥay* that *precede* a RC-head #*mi(n)*-N# having undergone phonological reduction (*llaḏī… min, ḥayṯu…min*), i.e. the morphemes *lla, ḥay* introducing a pre-nominal and restrictive RC, as in (7b, 8b). This situation arises as a RC-head of this sort serves as an analogical pivot for transferring the allomorphic nature of *lla, ḥay*, and *ḏī, ṯ* from the *post-nominal* and restrictive relativizers *llaḏī* and *ḥayṯ(u)* to their *pre-nominal* and restrictive counterparts. In other words, the pre-nominal and restrictive relativizers *llaḏī, ḥayṯ(u)* are allomorphs to the discontinuous restrictive relativizer *man/mā…min*, via Jespersen cycle and analogical extension.

In this light, the problem of the anomalous pragmatic behavior of the pre-nominal and restrictive relativizers *llaḏī, ḥayṯ(u)* raised at the end of the previous section evaporates as soon as we interpret them in a principled way (Jespersen cycle plus analogy) as allomorphic to the discontinuous restrictive relativizer *man/mā*, since the *new*-information (=focus) markers *lla, ḥay* in this framework signal the *new*-information N acting as a head of a restrictive RC, as much as *min* does in the same context. Similarly, the related problem of the puzzling co-occurrence of *min* with *lla, ḥay* is straightforwardly resolved by simply taking into consideration Jespersen cycle. Furthermore, as anticipated at end of Sect. 6, the same phenomenon can be reconciled with the syntactic observation that a contrast exists between the relative constructions **llaḏī/ḥayṯu*+ RC + N (cp. Table 5) and *llaḏī/ ḥayṯu* + RC + *min* + N (cp. (7a,7b) above) with no additional theoretical machinery. It will suffice to say that the post-nominal relativizers *llaḏī/ḥayṯu* do not tolerate a pre-nominal counterpart, as totally expected under an account of them in terms of complementary distribution (cp. Table 10), and that the pre-nominal relativizers that co-occurs with *min* do not make exception to this statement, since actually they are not proper relativizers *llaḏī/ḥayṯu*, but 'disguised' instances of the relativizer *man/mā*, such an allomorphy resulting from Jespersen cycle.

Finally, an interesting diachronic implication of an allomorphic analysis of the pre-nominal and restrictive relativizers *llaḏī/ḥayṯu* along these lines is a corroboration of the hypothesis, put forward at the end of the previous section, that a pattern of complementary distribution governs the morphological-pragmatic level of the system of OA relativizers, just as it does for the latter's phonological, syntactic, and semantic levels. According to this hypothesis, the post-nominal and non-restrictive *llaḏī/ḥayṯu*, and the restrictive and discontinuous *man/mā…min*, are in complementary distribution, the word-order of the focus-marker and the relative

stem in *llaḏī/ḥayṯu* being the opposite of their word-order in *man/mā…min*. This distributional scenario gains plausibility from Jespersen cycle: prior to its application and analogical extension, in fact, there occurred in the system of OA relativizers no pre-nominal and restrictive relativizers *llaḏī/ḥayṯu* allomorphic to the restrictive and discontinuous *man/mā…min*, and only was the latter's word-order, therefore, opposed to the opposite word-order of the post-nominal and non-restrictive *llaḏī/ḥayṯu*, in complementary distribution.

However, this is not the whole of the matter, since upon closer scrutiny, it appears that such a morphological-pragmatic pattern of complementary distribution in turn encompasses a subpattern of the same sort, which involves the post-nominal and non-restrictive relativizers *llaḏī/ḥayṯu*. More accurately, this is the loose pattern of complementary distribution, traditionally known as Kuryłowicz's (1973: 79) fourth law of analogy, where the relevant distinctive feature is to a good extent a diachronic opposition between an innovative and archaic form. Somewhat simplifying, this law states that a new form supersedes an older form in its primary (e.g. unmarked, frequent) functions, whereas the older form retains its secondary (e.g. marked, infrequent) functions as much as possible, in spite of the pressure of a new form to supersede it. For instance, the English bimorphemic word *brother-s* functions as a regular plural, while its *isomorphic* ancestor *bhrethr-en* is confined to a specialized plural, denoting members of a religious order – although *brother-s* is also possible in this sense.

Likewise, the OA bimorphemic relativizer *llaḏī* appears to be a more recent word than its isomorphic counterpart *ḥayṯ(u)*, so that *ḥayṯ(u)* can solely perform the secondary function of a non-argument, i.e. 'where' (cp. the obligatory locative morpheme *u* cumulated with it), whereas *llaḏī* performs the primary function of an argument, i.e. 'who(m)' etc. – although it can also work as a non-argument (cp. *llaḏī…fī-hi*). In particular, the relativizer *llaḏī* is regarded as more recent than *ḥayṯ(u)* in both the genetic and diffusionist models.

The genetic model, in fact, would consider as diachronically relevant the different morphological shape of the focus-markers *lla* and *ḥay* that occur in a *nominal* domain such as the relativizers *llaḏī* and *ḥayṯ(u)*, and would account for them in terms of a suprasegmental allomorphy that opposes gemination to diphthongation, based on similar cases from the Hebrew nominal domain, where geminated collectives such as *pequdda(h*) 'group of attendants' coexist with diphtonged collectives such as *gōḇay* 'swarm of locusts' (Corriente 1971b:21-23). The same model would also relate such an opposition to two different linguistic strata, nominal gemination going back to North-West Semitic, nominal diphtongation even back, to Early Semitic (1971b:60). On these grounds, the geminated *llaḏī* would be more recent than the diphtonged *ḥayṯ(u)*, not unlike the geminated *pequdda(h)* would be more recent than the diphtonged *gōḇay*.

The diffusionist model, meanwhile, would rather focus on the (pro)*nominal* allomorphy between *ḏī* and *ṯ*, which opposes a vowelled to a vowelless C, and also occurs in another instance of *nominal* morphology, the Arabic biradical nouns. Here, in fact, a vowelled stem such as *bin* 'son' coexists with a vowelless stem such as *bn*. To Garbini and Durand (1994:87-90, 115) as well as Testen (1998:208-209), in this kind of allomorphy the vowelled form goes back to North-West Semitic, the vowelless one to Early Semitic, so that, in this view, the vowelled *ḏī* would be more recent than the vowelless *ṯ*, as much as the vowelled *bin* 'son' would be more recent than the vowelless *bn*.

Interpreting the coexistence between *llaḏī* and *ḥayṯ(u)* in terms of such a diachronic opposition is a result particularly welcome here, for the reason that it enhances the symmetrical architecture of the multi-layered pattern of complementary distribution that this paper has plausibly identified as underlying the phonological, morphological, syntactic, semantic, and pragmatic dimensions of the OA relative markers *llaḏī, ḥayṯ(u)* and *man/mā…min.* This is illustrated in Table 11 below, which rewrites Table 10.

Table 11: OA Relative Markers – Distribution (Final Version)

<table>
<tr><td rowspan="2">RC-head
Relativizer</td><td colspan="10">Phonology</td></tr>
<tr><td colspan="2">Analytical</td><td colspan="8">Synthetical</td></tr>
<tr><td colspan="2"></td><td colspan="6">Partial Reduction</td><td colspan="2">Total Reduction</td></tr>
<tr><td colspan="2">Morph.-Prag.</td><td>Sem.</td><td>Syntax</td><td>Sem.</td><td>Syntax</td><td>Sem.</td><td>Syntax</td><td>Sem.</td><td>Syntax</td></tr>
<tr><td>lla</td><td>ḏī</td><td>non-restr.</td><td>post-nom.</td><td>~~restr.~~</td><td>~~post-nom.~~</td><td>~~restr.~~</td><td>~~post-nom.~~</td><td>~~restr.~~</td><td>~~post-nom.~~</td></tr>
<tr><td>man/mā</td><td>min</td><td colspan="2"></td><td>restr.</td><td>pre-nom.</td><td>~~restr.~~</td><td>~~post-nom.~~</td><td>~~restr.~~</td><td>~~post-nom.~~</td></tr>
<tr><td>ḥay</td><td>ṯ(u)</td><td>non-restr.</td><td>post-nom.</td><td>~~restr.~~</td><td>~~post-nom.~~</td><td>~~restr.~~</td><td>~~post-nom.~~</td><td>~~restr.~~</td><td>~~post-nom.~~</td></tr>
<tr><td colspan="2">Symbols</td><td colspan="2">X=complem. distribution, cp. Table 11bis for details</td><td colspan="2">X=complementary distribution</td><td colspan="2">~~X~~ =Jespersen cycle and analogy</td><td colspan="2">~~X~~ =heaviness and analogy (‘harmonicity’)</td></tr>
<tr><td colspan="2" rowspan="3">Examples</td><td colspan="2">(1) Aḷḷāhu llaḏī nazzala l-kitāba</td><td colspan="4">(7a) min-a l-dīni … llaḏī awḥaynā ilay-ka /
(7b) llaḏī ǧā’a-ka min-a l-ᶜilmi</td><td colspan="2">(4) hāḏā (RAǦULUN) man aᶜrifu munṭaliqan</td></tr>
<tr><td colspan="2"></td><td colspan="4">(2a) min-a l-dīni mā waṣṣā bi-hi Nūḥan /
(2b) mā anzala-ḷḷāhu min-a l-kitābi</td><td colspan="2">(5) hāḏā (RAǦULUN) llaḏī ᶜalimtu annī lā aᶜrifu-hu munṭaliqan</td></tr>
<tr><td colspan="2">(3) bi-Ḫayfi Banī Kināna ḥayṯu taqāsamū ᶜalā l-kufri</td><td colspan="4">(8a) min arḍi-nā ḥayṯu ši’tum /
(8b) ḥayṯu šā’ū min arḍi-him</td><td colspan="2">(6) ḥayṯu takūnu akūnu (= l-makānu llaḏī takūnu fī-hi akūnu)</td></tr>
<tr><td colspan="2">Morphology-Pragmatics (Table11bis)</td><td colspan="2">focus-m.</td><td colspan="2">rel. stem</td><td>focus-m.</td><td>rel. stem</td><td colspan="2">sem. role</td></tr>
<tr><td>man/mā</td><td>min</td><td colspan="2">final</td><td colspan="2">initial</td><td></td><td></td><td colspan="2"></td></tr>
<tr><td>lla</td><td>ḏī</td><td colspan="4" rowspan="2"></td><td>initial</td><td>final</td><td colspan="2">argument</td></tr>
<tr><td>ḥay</td><td>ṯ(u)</td><td>initial</td><td>final</td><td colspan="2">non-argument</td></tr>
</table>

Having clarified throughout this paper how such a generalized pattern of complementary distribution arises out of several patterns of complementary

distribution that manifest themselves in the system of OA relativizers, we may still wonder *why* such patterns manifest themselves in the latter. While we can invoke structural factors of analogy (cp. the general drift from analytical to synthetical language) to account for the phonological pattern of complementary distribution (analytical vs. synthetical RC-head), in the case of the syntactic and semantic patterns of complementary distribution, cross-linguistic evidence rather points to a functional explanation, as already discussed at end of Sect. 4. In this framework, a linguistic opposition that one way or the other comes through on the sound-side, such as the different syntactic position that pre-nominal and post-nominal RCs occupy in the linear order, is the audible manifestation of a semantic opposition, such as the distinction between restrictive and non-restrictive RCs, which would be otherwise more difficult to retrieve for the speakers. The lack of such a syntactic pattern of complementary distribution, in fact, would force the speakers to infer the related semantic pattern of complementary distribution from an empirical domain more complex than a relative construction alone, such as contextual or extra-linguistic information.

A functional explanation of this sort, however, has traditionally faced OA and, more generally, Arabic with two interpretive problems. On the one hand, the world's languages frequently signal the aforesaid semantic pattern of complementary distribution by means of a strategy alternative to the syntactic pattern of complementary distribution, namely a suprasegmental pattern of complementary distribution, which associates the presence of a pause or – in graphemic rendering – a comma to a non-restrictive relativizer, and its lack to a restrictive relativizer (cp. English (,)*which*), but OA lacks it, and the ultimate reason of this behavior remains obscure. On the other hand, as far as is known, the morphological-pragmatic pattern of complementary distribution typical of OA, i.e. *llaḏī, ḥayṯ(u)* vs. *man/mā…min* finds no room in the functional explanations of Arabic relativization – for instance, Gensler (2004) countenances a semantic pattern of complementary distribution for the system of OA relativizers, and yet such an analysis does not expound *why* an OA relativizer such as *llaḏī* is morphologically more complex than its English counterpart ,*which*, due to its incorporating the 'reinforcer' *la*.

A plausible solution to both problems lies in the oral-poetic nature of OA, as highlighted in Monroe's (1972) and Zwettler's (1978) work. The argument is built as follows.

Firstly, the presence of the focus-markers *lla, ḥay, min* in the system of OA relativizers systemically correlates with the lack, in the same system, of an opposition between them in terms of presence or absence of the suprasegmental device of pause – as is inferred from the fact that no contrast between the presence or absence of comma is found in the writing of such relativizers.

In second place, we can take such a correlation to be too pervasive to be accidental, and can interpret accordingly the focus-markers *lla* and *ḥay*, which actually signal a *non*-restrictive RC (see Sect. 7), to be the functional equivalents of the presence of pause/comma in languages such as English, Italian etc., and the focus-marker *min*, which (indirectly) signals a restrictive RC (via its RC-head: cp. Sect. 7), to be the functional equivalent of the absence of pause/comma in these languages.

Last but not least, such an inability of OA to avail itself of pause in its relativization, and the related usage of the focus-markers *lla, ḥay, min* instead, is deeply rooted in the oral-poetic nature of this language, which was in origin a *Kunstsprache* that privileged poetic contents, and expressed them in oral form, as attested by the formulaic nature of pre-Islamic poetry (Monroe 1972, Zwettler 1978) and, to a lesser extent of the Koranic *sağ*^{c} prose (Rippin 2005:121, and cp. also the formulaic expression *bi-Ḫayfi Banī Kināna ḥayṯu taqāsamū* ^{c}*alā l-kufri* studied at beginning of Sect. 5). Remarkably, such a diachronic and sociolinguistic scenario entails for the suprasegmental dimension of OA that pause was unavailable as a device to distinguishing a non-restrictive relativizer from its restrictive counterpart, since it was already used to signal a different kind of information, notably a rhythmic unit, as shown by the widespread usage of pausal forms in pre-Islamic poetry, the Koran, and so on. Accordingly, the focus-markers *lla, ḥay, min* were deployed instead of the pause, as a repair-strategy.

One key-point stands out from these remarks on the suprasegmental dimension of an oral-poetic language such as OA: the rationale behind this situation is essentially functional, and amounts to a sort of adaptive mechanism, which converts the suprasegmental opposition between presence or absence of pause into a morphological-pragmatic opposition between the focus-marker *lla, ḥay* and *min*, in order to adapt the suprasegmental opposition in question to the metrical conditions that the oral-poetic nature of OA imposes onto its relativization.

To conclude, a multi-layered pattern of complementary distribution governs the system of OA relativizers, which has at its core a restrictiveness-based opposition, with a morphological-pragmatic distinction as its audible counterpart. From a cross-linguistic perspective, the latter opposes the focus-markers *lla, ḥay* to the focus-marker *min*, which are very likely to be metrically-conditioned replacive morphemes, respectively, of the presence or absence of pause/comma, a suprasegmental device deployed for marking the restrictiveness-based opposition, familiar from English, Italian etc. relativization.

Such a multi-layered pattern of complementary distribution is obscured, in diachrony, by syntactic and morphological factors, which consist, respectively, of a word-order change affecting a Dependent RC with respect to a Head *min*-N ('heaviness'), and in the rise of an allomorphy concerning the relativizer *man/mā; min* ('Jespersen cycle'). Both these phenomena, in turn, are triggered by a

phonological factor that can be identified with a phonological reduction of the particle *min* reported by primary sources, and that, from a modern typological perspective, is part and parcel of a more general shift from analytical to synthetical language which scholars such as Petrácek and Owens have found out to be at work in (Old) Arabic (cp. fn. 36 above).

Analyzing along these lines the apparent deviations from the multi-layered pattern of complementary distribution that informs the system of OA relativizers *llaḏī/ḥayṯu*, *man/mā…min* has two important typological implications. Firstly, the study of such anomalies gathers direct evidence in favor of the existence of a postpositional construction in early stages of OA (cp. the relic form *ᶜalā ᶜan* in Sect. 6), thus lending empirical credence to the hypothesis entertained in the literature that Arabic, and more generally Semitic, was originally a postpositional language. Secondly, a study of this kind brings to light the 'missing link' between the formally identical *mā-min* relative and negative constructions, in the form of a negative cleft-construction, a result that receives indirect confirmation from Teso, the Nilo-Saharan language where a similar grammaticalization pattern occurs (cp. Sect. 7).

REFERENCES

A. Primary sources

al-Buḫārī, *Ṣaḥīḥ* = Abū ᶜAbd Aḷḷāh Muḥammad b. Ismāᶜīl al-Buḫārī, *al-Ǧāmiᶜ aṣ-Ṣaḥīḥ*. Edited by Muḥammad Zuhayr Ibn al-Nāṣir an-Nāṣir, 9. vols., Beirut: Dār Ṭūq an-Naǧāt, 2001.

Sībawayhi, *Kitāb* = Abū Bišr ᶜAmr b. ᶜUṯmān b. Qanbar Sībawayhi, *al-Kitāb*. Edited by ᶜAbd as-Salām Muḥammad Hārūn, 5 vols., Cairo: Maktabat al-Ḫānǧī, 1982.

Ibn Manẓūr, *Lisān* = Muḥammad b. Mukarram b. ᶜAlī Ibn Aḥmad Ibn Manẓūr. *Lisān al-ᶜarab*. Edited by ᶜAbd Aḷḷāh ᶜAlī l-Kabīr & Muḥammad Aḥmad Ḥasab Aḷḷāh Hāšim al-Šāḏilī, 6 vols., Cairo: Dār al-Maᶜārif, 1981.

Ibn Hišām, *Muġnī* = Ibn Hišām al-Anṣārī. *Muġnī labīb ᶜan kutub al-aᶜārīb*. Edited by Ḥasan Ḥamad, 3 vols. Beirut: Dār al-Kutub al-ᶜIlmiyya, 2005.

al-Astarābāḏī, *Šarḥ al-Kāfiya* = Raḍī d-Dīn Muḥammad b. al-Ḥasan al-Astarābāḏī, *Šarḥ al-Kāfiya*. Edited by Yūsuf Ḥasan ᶜUmar, 5 vols., Benghazi: Manšūrāt Ǧāmiᶜat Qāryūnis, 1996.

Ibn Yaᶜīš, *Šarḥ al-Mufaṣṣal* = Muwaffaq ad-Dīn Abī l-Baqā' Yaᶜīš b. ᶜAlī Ibn Yaᶜīš, *Šarḥ al-Mufaṣṣal*. Edited by Imīl Badīᶜ Yaᶜqūb, 6 vols., Beirut: Dār al-Kutub al-ᶜIlmiyya, 2001.

B. Secondary sources

Badawi, E., Carter, M. & A. Gully. 2004. *Modern Written Arabic: A Comprehensive Grammar*. London & New York: Routledge.

Barth, J. 1913. *Die Pronominalbildung in den Semitischen Sprachen*. Leipzig: Hinrichs.

Benmamoun, E. 2005. "Construct State." in: *Encyclopedia of Arabic Language and Linguistics, Volume I*, ed. by K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski, 477-482. Leiden: Brill.

Bloch, A. 1991. *Studies in Arabic Syntax and Semantics*. Wiesbaden: Harrassowitz.

Bravmann, M. 1961. "Genetic Aspects of the Genitive in the Semitic Languages". *Journal of the American Oriental Society* 81.386-394.

Brockelmann, C. 1910. *Précis de linguistique sémitique*. Paris: Geuthner.

Comrie, B. 1981. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

Corriente, F. C. 1971a. "On the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology". *The Jewish Quarterly Review* 62.20-50.

____.1971b. *Problematica de la pluralidad en semitico. El plural fracto*. Madrid: Consejo Superior de Investigaciones Cientificas.

Dévényi, K. 1988. "The Treatment of Conditional Sentences by the Mediaeval Arabic Grammarians". *The Arabist* 1.11-42.

Diem, W. 2007. "Arabic *allaḏī* as a Conjunction: An Old Problem and A New Approach". *Approaches to Arabic Linguistics*, ed. by E. Ditters & H. Motzki, 67–112. Leiden: Brill.

Esseesy, M. 2010. *Grammaticalization of Arabic Prepositions and Subordinators: A Corpus-Based Study*. Leiden: Brill.

Fleisch, H. 1959. "L'aspect lexical de la phrase arabe classique". *Studia Biblica et Orientalia* 12.78-94

____.1961. *Traité de philologie arabe*. Beyrouth: Imprimerie catholique.

Fortescue, M. 1996. "Grammaticalized Focus in Yukagir: Is It really grammaticalized and is it really Focus?". in: *Content, Expression and Structure: Studies in Danish Functional Grammar*, ed. by E. Engberg-Pedersen, M. Fortescue, P. Harder, L. Heltoft, L. Falster Jakobseng, 17-38. Amsterdam: Benjamins.

Garbini G. & O. Durand. 1994. *Introduzione alle lingue semitiche*. Brescia: Paideia.

Gensler, O. 2004. "The Arabic *mā – min* Construction: Rethinking the Priority of Restricted Relatives". Paper presented at *Syntax of the World's Languages* (SWL 1), Leipzig, August 5-8.

Grande, F. 2013. *Copulae in the Arabic Noun Phrase. A Unified Analysis of Arabic Adnominal Markers*. Leiden: Brill.

Halliday, M. K. 2002. *On Grammar*. New York: Continuum

Hawkins, J. A. 1983. *Word Order Universals*. New York & London: Academic Press.
Heine, B. & M. Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske Verlag.
Kay, H. C. trl. 1892. *Yaman. Its Early Mediaeval History by Najm ad-Din ᶜOmārah al-Ḥakam. Also The Abridged History of Its Dinasties by Ibn Khaldūn, and An Account of The Karmathians of Yaman by Abu Abd Allah Baha ad-Din al-Janadi*. London: Edward Arnold.
Kuryłowicz, J. 1973. *Esquisses linguistiques*. Munich: Wilhelm Fink Verlag.
Lambrecht, K. 1994. *Information Structure and Sentence Form*. Cambridge: CUP.
Lehmann, Ch. 1984. *Der Relativsatz: Typologie seiner Strukturen; Theorie seiner Funktionen; Kompendium seiner Grammatik*. Tübingen: Gunter Narr Verlag
Levin, A. 1985. "The Distinction between Nominal and Verbal Sentences according to the Arab Grammarians". *Zeitschrift für arabische Linguistik* 15.118-127.
Loock, R. 2010. *Appositive Relative Clauses in English. Discourse Functions and Competing Structures*. Amsterdam: Benjamins.
Lucas, Ch. 2007. "Jespersen cycle in Arabic and Berber". *Transactions of the Philological Society* 105.398-431.
Monroe, J. T. 1972. "Oral Composition in Pre-Islamic Poetry". *Journal of Arabic Literature* 3.1-53
Moscati, S., Spitaler, A., Ullendorf, E. & W. von Soden. 1964. *An Introduction to the Comparative Grammar of the Semitic Languages: Phonology and Morphology*. Wiesbaden: Harrassowitz.
Ostle, R. C. 1982. "Review of *The Oral Tradition of Classical Arabic Poetry: Its Character and Implications* by M. Zwettler". *Bulletin of the School of Oriental and African Studies* 45.159-161.
Ouhalla, J. 1999. "Focus and Arabic Clefts". in: *The Grammar of Focus*, ed. by G. Rebuschi & L. Tuller, 335-359. Amsterdam: Benjamins.
____. 2000. "Possession in Sentences and Noun Phrases". in: *Research in Afroasiatic Grammar*, ed. by J. Lecarme, J. Lowenstamm & U. Shlonsky, 221-242. Amsterdam: Benjamins.
Owens, J. 2006. *A Linguistic History of Arabic*. Oxford: OUP.
Petrácek, K. 1981. "Le système de l'arabe dans une perspective diachronique". *Arabica* 28. 162-177.
Pullum, G. K. & A. M. Zwicky. 1988. "The Syntax-Phonology Interface". in: *Linguistics: The Cambridge Survey; Volume I, Linguistic Theory: Foundations*, ed. Frederick J. Newmeyer, 255-80. Cambridge: CUP
Rabin, Ch. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.
Reckendorf, H. 1895. *Die Syntaktischen Verhältnisse des Arabischen*. Leiden: Brill.
Rippin, A. 2005. "Review of *Fables of the Ancients? Folklore in the Qur'ān* by A. Dundes". *Bulletin of the School of Oriental and African Studies* 68.120-122.

Ross, J. 1967. "Constraints on Variables in Syntax." PhD diss., MIT.
Sadan, A. 2012. *The Subjunctive Mood in Arabic Grammatical Thought*. Leiden: Brill.
Sara, S. 2006. "Sībawayhi's Kitāb Chapters 21-25: Translation and Notes". *Journal of Arabic Linguistic Tradition* 4.30-45.
Saussure, F. 1916. *Cours de linguistique générale*. Paris: Payot.
Taylor, J. R. 2000. *Possessives in English: An Exploration in Cognitive Grammar*. Oxford: OUP.
Testen, D. 1998. *Parallels in Semitic linguistics: The Development of Arabic la- and Related Semitic Particles*. Leiden: Brill.
Versteegh, K. 1993. *Arabic Grammar and Qur'ānic Exegesis in Early Islam*. Leiden: Brill.
____. 1997. *The Arabic Language*. Edinburgh: Edinburgh University Press.
Zwettler, M. 1978. *The Oral Tradition of Classical Arabic Poetry: Its Character and Implications*. Columbia: Ohio State University Press.
Wright, W. 1896. *A Grammar of the Arabic Language*. Cambridge: Cambridge University Press.