Syntactic reduplication in Arabic
Utz Maas

1. Iteration and reduplication – the case of grammaticalization.

Some introductory comments

In current discussions, reduplication is generally treated as a device of word formation, i.e., as a special form of derivational patterning (cf. e.g. Wiltshire and Marantz 2000; Raimy 2000). In this paper I will take a more general perspective, which can already be found in older typological work (e.g. Pott 1862): Reduplication is seen as a formal linguistic device that can be used at all levels of linguistic structure. This theoretical move shifts the leading question to one of grammaticalization: To what extent is reduplication grammaticalized? To get a better grasp of the problem, we need a pre-theoretical (or pre-grammatical) term: In this sense I will speak of doubling or in the case of more than two forms of iteration. Investigating this phenomenon produces a problem similar to optical illusions: The way you look at it determines what you will see. Doubled forms can be the result of a lack of structural differentiation, as e.g. in early child language and in certain types of aphasic speech, or it can be the result of structure building, i.e. as the effort to keep things identical through repetition. Two perspectives on doubling can thus be distinguished:

I. Doubling can be seen as a reflex of inertia in linguistic activity,

II. It can also be seen as the result of a special type of structuring, thus as learned behavior in a linguistic community, as part of the language system.

Conflation of these two different perspectives prevails not only in lay thinking about language, where it is considered stereotypical of “primitive” languages. Even grammatical handbooks, especially when addressing a larger public, hawk this stereotype as well. It is an aim of the present paper to differentiate between perspectives I and II which must be understood as dual aspects of linguistic behavior and thus potentially present in any speech act or language. We can trace perspective I at all levels of language:
from echoed prosodies in conversation up to lexical repetitions. The need to differentiate between perspectives I and II is especially important in research on early language development: Doubling is among the earliest devices a child makes use of when trying to impose structure on its utterances, cf. Leroy and Morgenstern (this volume), using the more or less biological constraints of its linguistic activities as resources for symbolic behavior, successively imposing more differentiated structures upon its utterances in pace with its mastering of the structures of adult language. But this ambivalent relation to doubling is also present in adult behavior: Every speaker (or writer) can make an effort to use iteration as well as an effort to avoid it. This must be distinguished from grammaticalized reduplicating patterns in the structure of a language which cannot be avoided (i.e. which are not mere stylistic options). This distinction can already be found in early research: Hermann Reckendorf, a pioneer of reduplication research, distinguished bound reduplication (gebundene Paronomasie) from stylistically free reduplication (Reckendorf 1909). Bound reduplication pertains to grammaticalization. This is the subject of the present paper in which I will present some data from a linguistic family that is notorious for its extensive use of reduplication: Arabic.

On a very general level, perspective I implies the idea that language is articulated, i.e., formally differentiated in its elements. Thus where iteration is not due to imperfect or rudimentary linguistic knowledge, it is a marked form of expression (= perspective II). Markedness conveys a particular interpretation. Examples of this can probably be found in all languages, cf. the German example (1):

(1) *Es kommt selten vor, dass sich der hessische Ministerpräsident windet und windet und windet, wenn ihn jemand nach seiner Meinung fragt.*

'It seldom happens that the Prime Minister of Hessia squirms and squirms and squirms when someone asks him for his opinion'

Expressions like these can be more or less conventional or even lexicalized (as idioms). In non-lexicalized cases, the interpretation is a function of the basic meaning of the iterated elements, the markedness of the iteration adding, e.g., an element of gradation, cf. the German examples in (2), heard in injunctions, that can probably be replicated in most (all?) languages:

(2) *langsamer, langsamer!* ‘slowly, slowly!’ (i.e. calm down!)
    *schneller, schneller!* ‘quickly, quickly!’ (i.e. hurry up!)
On the formal side, doubling or iteration can be holistic or partial (analytic). Holistic iteration can be achieved by the simple repetition of utterances or their parts, e.g. words as in (1) and (2). In distinction to this, grammaticalized devices are to be expected to be analytic. This can be exploited in “poetic” uses of language, e.g. in the comic series the SMURFS (in German: Schlumpf), very popular some years ago. The Smurfs had a special Smurf language, with fully developed grammatical devices but with a lexical dummy smurf that could be used throughout. Thus, in the smurf language it would have been possible to build a sentence like:

(3) I smurfed my smurfy smurf to a smurf

Besides the emphasized functional elements, there are of course covert grammatical devices such as e.g. word order, that show English as the matrix language. Iteration in (3) is analytic: The word forms in (3) are differentiated by grammatical formatives, the iterated element smurf is a stem. A language with sentences like (3) would be undifferentiated (unarticulated) only at the level of the lexicon, not at the level of word-forms. In fact, the smurfs were rather parsimonious in using this kind of reduced differentiation. In the comics we do not find sentences like (3) but rather like (4):

(4a) Lazy Smurf, have you smurfed that play for our village fair?
(4b) And look at Smurfette! She’s much smurfer than that!

Smurf-forms are marked expressions in an otherwise lexically differentiated context. And every smurf-form can be substituted by a fully differentiated non-smurf form. Thus we can hypothesize that elements of low differentiation fulfill a special function in full-fledged languages, both in the Smurf language as well as in “natural” languages.

The grammaticalization of doubling (i.e. reduplication) implies the definition of the domains where it applies, in contrast to iteration which is unbounded and can be exploited by all kinds of parallelisms in texts, especially in poetic language. The domain of iteration in the Smurf-language is syntax (the sentence) and the form of iteration is analytic: Lexical stems are repeated, augmented by grammatical formatives that determine the syntactic function of the word forms. Analytic reduplication presupposes the parsing of forms, which can be done on a morphological basis, e.g. by parsing word forms into stems and affixes as in (3). Depending on the language type, the categories of morphological parsing can be more fine grained, e.g. roots can be iterated as in Semitic languages (see below), and there can be
phonological filters as well. The generalized pattern of this type of reduplication isolates word-internal elements that are lexically anchored. Thus the structure of a sentence like (3) can be graphically represented as (5), where Greek letters represent (grammatical) affixes and Roman letters the stems:

(5) Reduplication through the repetition of word-internal lexical elements in a sentence (reduplicands are represented by capitals):

\[
\begin{array}{cccccccc}
\# & A & -\alpha & B & -\beta & C & -\gamma & D & -\delta & E & -\epsilon & \ldots \ldots & \#
\end{array}
\]

Here, A, B, C, D, and E are derived from the same stem, while the affixes indicate the function of these words in the construction. The mirror of (5) iterates the grammatical formatives (operating again within a syntactic domain) and is graphically represented in (6).

(6) Grammatical reduplication:

\[
\begin{array}{cccccccc}
\# & A & -\alpha & B & -\beta & C & -\gamma & D & -\delta & E & -\epsilon & \ldots \ldots & \#
\end{array}
\]

Here, A, B, C, D, and E derive from different stems while the affixes are identical (in function, at least). It is evident that (6) represents what is usually called agreement, cf. the Latin example (7):

(7) tenebr-ae altissim-ae obor-t-ae sunt
darkness-N.PL.F profoundest- N.PL.F ensue-PCP-N.PL.F be:3.PL
‘deepest darkness fell’

Following the classification of Pott (1862), I will call the types of reduplication found in (5) and (6) syntactic reduplication. The main subject of this paper will be the structure (5).

The domain of reduplication can be smaller than the sentence. Grammaticalization presupposes a grammatical constituent as its domain: i.e., a construction. The limiting case of construction is the (morphologically complex) word. Thus in the limiting case of reduplication, form elements are reduplicated word-internally. In languages with morphological affixes, this leads to structures like (8):

\[
\begin{array}{cccccccc}
\# & A & -\alpha & B & -\beta & C & -\gamma & D & -\delta & E & -\epsilon & \ldots \ldots & \#
\end{array}
\]
Word-internal reduplication:

\[ \# \quad \ldots \ldots \quad \boxed{\ldots \ldots \quad \boxed{\ldots \ldots} \quad -\alpha \quad \ldots \ldots \quad \#} \]

As the delimitation between constructions and words is rather complicated from a typological point of view, we can expect these complications in reduplication research as well. This will become evident in the following, when cases of syntactic reduplication are analyzed in which the domain can perhaps be understood as “multi-worded” words.

Word-internal reduplication is the dominant subject of current research in reduplication, in general presupposing the word as the domain of reduplication, but without bothering much about defining this entity. An exception is Moravcsik (1978), who uses a broad concept of reduplication, one that provides for the analysis of syntactic reduplication as well. However, in her typological survey, she only takes word-internal reduplication into account. Recent work simply presupposes word-internal reduplication as the domain of reduplication, e.g., Marantz (1982), and Raimy (2000), where the focus is on the phonological constraints. I will take the broader perspective, and will use Moravcsik’s reduplicative construction as a cover term for the different types of reduplication shown in (5), (6) and (8), analyzing the special kind of structural unity it imposes on its constituents.

Reduplication as an analytical device presupposes some kind of filter to define the (partial) structure that has to be reduplicated. The filtering patterns can be based in phonology (prosodic or syllabic templates) or can be purely formal (morphological, without alignment to phonological structures). These patterns are the subject of most ongoing research, which aims at identifying universal patterns (e.g., the “emergence of the unmarked” in reduplicands). The focus of this paper is complementary to this approach: I will look at reduplication devices (analytic iteration) as the germ of grammaticalized structures, where the reduplicand modifies the base, as in (8). This makes it necessary to delimitate reduplication from holistic doubling, which borders on stylistics, as well as from exclusively formal iteration without functional differentiation, be it in lexical structure or as a kind of stylistic play with form, as e.g., in alliteration, rhyme, etc.

In the following I will focus on syntactic reduplication. This has traditionally been the subject of rhetorical studies, where it is discussed as paronomasia. Paronomasias can produce tautological expressions, coming close to the Smurf language, but even then they can acquire meaning, as in Gertrude Stein’s famous a rose is a rose is a rose. But in a more trivial con-
text they run the risk of appearing rather nonsensical, as e.g. *the watery water, the hiker hikes, he screamed a scream*, etc. Usually, they will only make sense if further modified, for example:

(9) Types of syntactic reduplication
- subject – predicate: *the hiker hiked in the mountains*
- noun – attribute: *the very watery water (??)*
- verb – object: *he screamed a loud scream*

Especially problematic is the paronomastic attribute (9b); the other constructions are more acceptable, where the paronomastic element serves to support a further meaningful specification. Despite this, there is a long normative Western tradition which frowns upon paronomasias, already present in classical times\(^\text{10}\). This makes for an interesting contrast to Semitic traditions, where texts abound in them: They are favored not only in poetic works (including the *Qor\(\text{\textit{\textit{an}}\text{)}}, but are current in colloquial language as well (cf. Reckendorf 1909 for documentation from literary as well as non-literary texts). These different stylistic preferences might be a hint at differences in the “underlying” linguistic type: Semitic languages are obviously more prone to this device than Indo-European languages. This will be explored in the following, focusing on reduplication within the verb phrase (cf. 9c) and mentioning other types only cursorily.

2. Iteration and reduplication in Old (= Classical) Arabic

In this paper, I will use the term *Arabic* in the generic sense of a phylum, thus including Old Arabic as well as neo-Arabic varieties. When the reference is more specific, I will use a more specific term. It is important to differentiate between Old Arabic in the diachronic sense, for which I take *Classical Arabic* as a grammatically standardized representative, and *(Modern) Standard Arabic*. As diachronic aspects are of interest here, I will contrast Old Arabic, quoting form reference works for the Classical texts, with data from two neo-Arabic varieties, Maltese and Moroccan Arabic.

At first glance, the structure of Arabic is characterized by a high degree of grammatical differentiation, with a fusional morphology that has been retained, or more precisely reconstructed, in neo-Arabic varieties despite the very far-going changes in phonology, such as apocope of inflectional suffixes, loss of quantity opposition in the vowel system, and the complete restructuring of syllabification (especially in Western varieties). An exam-
Syntactic reduplication in Classical Arabic (or Modern Standard Arabic) is (10):

(10) \textit{katab-	ext{a}} \quad \textit{l-ka:tie-b-\text{u}} \quad \textit{l-ki:ta:b-\text{a}}

\textit{write:PF-3SM} \quad \textit{DEF-writer-NS} \quad \textit{DEF-book-AS}

‘the writer wrote the book’

(10) is different from (9) as well as from the Smurf examples in (4) as there is no perfect match between the base and the reduplicated elements: Reduplication is only partial as far as the segmental(concatenative) structure is concerned. Thus we could call the reduplication \textit{dense} in cases like (9), where the reduplicands are complete strings in the form of the base, and \textit{porous} in cases like (10), where the identifiable reduplicand is only matched by a porous structure in the base, in (10) the sequence of consonants \textit{k-t-b}. But a classification such as \textit{dense} \textit{vs. porous} would miss important generalizations of grammaticalization governing reduplication in Arabic. (10) can be represented in a more explicit (transparent) way as (11):

(11) \textit{kataba} \quad \textit{l-ka:tie-bu} \quad \textit{l-ki:ta:ba} = \textit{ktb} \{V-PF.3SM\} + \textit{ktb} \{N-Ag.NS\} + \textit{ktb} \{N-Res.AS\}

‘the writer wrote the book’

(11) shows the invariable element \textit{°ktb°} with internal grammatical markers (indicated by \{\}): The finite verb, with the functional marking of the predicate, the main actant, articulated as “agentive noun” (N-Ag) in the subject case, and the secondary actant, articulated as “resultative noun” (N-Res) in the object case.

This construction is bound to the special structure of Semitic languages, which split the stem into different melodies. Arabic has been the paradigmatic case in recent morphological work operating on a multi-layered analysis (where the layers must be defined independently), in which dense structures (\textit{tiers}) at the different layers are identified. Autosegmental approaches have generalized this model since McCarthy (1982), and most recent work in reduplication (word-internal reduplication, cf. (8)) has generally taken this model as its guide-line (since Marantz (1982)).

As this multi-layered analysis of Arabic is well-known, it will be sufficient to illustrate it with one example from the verb form \textit{kataba} “(he) wrote”. The structure is presented in a grid, separating the different dimensions:\text{\footnote{11}}.
1. Word-level, simplified as segmental (time) slots,
2. melodies, differentiated for consonantal and vocalic elements (defined by syllabificational potentials [potential nuclei vs. margins])
3. root-level: lexical invariants, traditionally called “radicals“,
4. stem-level: augmented root-structure, modified by inflectional affixes, to be differentiated between
   a. “word formation“: semantic modification of the root (stem-W),
   b. grammatical specification (stem-G)
5. inflectional level: prefixes and / or suffixes

Levels (1) and (2) define the phonological structure, level (3) defines the lexical invariant, and the function of the chosen form is indicated at levels (4) and (5). Thus levels (3) to (5) can be seen as increasing specifications of the word form: Level (3), the root skeleton, is the minimal specification, level (5) is the maximal specification:

(12) word-form (verbal) *kataba* “he has written” = {x-x-x-x-x-x} (six time slots)

<table>
<thead>
<tr>
<th>1. word</th>
<th>x</th>
<th>x</th>
<th>x</th>
<th>x</th>
<th>x</th>
<th>x</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. melodies</td>
<td>C</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. root</td>
<td>R₁ = k</td>
<td>R₂ = t</td>
<td>R₃ = b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{writing}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. a stem-W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(void)¹²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.b. stem-G</td>
<td>V₁ = a</td>
<td>V₂ = a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{ACT, PF}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. inflection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>{3.S.M.}</td>
<td>V₃ = a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Examples such as (11) show the partial reduplication of words at the syntactic level. The word forms are morphologically differentiated for syntactic functions. Holistic reduplication is possible as well, where the reduplicated forms fill the syntactic slots of words, cf.

(13) holistic reduplication in Classical Arabic:\(^{13}\):

a. *maʃa:* \(\rightarrow\) *fuwaj:a(tan)\(\rightarrow\) *fuwaj:a(tan)*
go:PF.3.S.M. little (A.S.-DEF) little (A.S.-DEF)
‘he went slowly’

b. *ʔintazər-a* \(\rightarrow\) *sanat-an* \(\rightarrow\) *sanat-an*
‘he looked out year after year’

c. *ʔistaqra:hum* \(\rightarrow\) *haj:-an* \(\rightarrow\) *haj:-an*
‘he examined them tribe for tribe’

Reduplication here is a means of semantic modification that would have an equivalent in word formation in many other languages (Arabic is especially poor in terms of word formation devices). Thus cases like (13) represent borderline cases between syntactic and word-internal reduplication, since a constructional meaning can be identified. The modification of the meaning is usually quantifying. Thus the reduplicative means is in a certain sense “iconic”\(^{14}\), often corresponding to determiners (quantifiers, adjectives etc.) in other languages. It is often expressed through equivalent syndetic (coordinated) constructions, which could be used in other languages as well, cf. the examples of syndetic reduplication in Old Arabic in (14):

(14) a. *naːda:* \(\rightarrow\) *radʒul-un* \(\rightarrow\) *wa* \(\rightarrow\) *radʒul-un*
‘every individual man called’

b. *ʔazin-at* \(\rightarrow\) *Salaj-ʰaː* \(\rightarrow\) *ʔakθara* \(\rightarrow\) *fa* \(\rightarrow\) *ʔakθara*
become.grieved:PF-3.S.F to-3.S.F more and more
‘she always becomes grieved’

Apparently the kind of reduplication in (13) and (14) is a stylistic means, and as such possible in other languages as well (such as English), which are much less tolerant of asyndetic expressions, preferring syndetic expressions instead. In the following, this kind of reduplication will not be further ana-
lyzed, and I will focus on cases of analytic syntactic reduplication. But as word-internal reduplication is often quoted as characteristic of Arabic\(^\text{15}\), these structures will also be discussed briefly.

In the lexicon, the doubling of form elements is frequent at the stem level, where a distinction must be made with respect to grammaticalization. Doubling is frequently found in expressive forms, although it cannot be identified as a productive device of expressive word formation. What is decisive here is the filter of three radicals at the root level (cf. the grid in (12)) or, exceptionally, four radicals. Thus an etymological word formative consisting of two radicals would not pass this filter, unless doubled, thereby yielding a four-radical stem\(^\text{16}\). Most words with this formation have an expressive component, but not all, cf.\(^\text{17}\):

(15) doubled root formation in Arabic
- °ws°: waswasa ‘insinuate’
- °sr°: sarsar ‘cock’
- °dr°: dardar ‘oak’
- °n†°: nasnaf ‘mint’

This formal characteristic cannot qualify as reduplication in the sense defined above, as the doubled formatives cannot be identified on the lexical and / or grammatical level.

At the other extreme of the productivity scale of word formation is a device which affects the second radical. This is the most productive of the fifteen formation devices at the stem-W level (cf. 12) given by the grammarians of Classical Arabic (the so called second stem). Traditionally it is analyzed as (word-internal) partial reduplication\(^\text{18}\). With this stem-formation, two meanings can be distinguished – albeit with some overlap in their use:

(16) reduplication of second radical:

a. intensive formation (thus more or less expressive?):
   - °drb° ‘hit’: daraba ‘to hit violently’
   - °ksr° ‘cut’: kasara ‘to cut into small pieces’
   - °frq° ‘separate’: farafa ‘to disperse’

b. causative formation:
   - °frh° ‘be happy’: faraha ‘to make happy’
   - °hsn° ‘be beautiful’: hasana ‘to make beautiful’
   - °ilm° ‘know’: jalama ‘to teach’
This traditional analysis presents a number of analytical problems. From a less abstract point of view, one which is not biased by Latinized transcription habits, the second radical in these forms is forticized. The Arabic term is *tafdi.d* "strengthening", a phenomenon which is not represented in orthography. The phonetic realization is a geminate with a clear differentiation between the implosive and the explosive parts, and thus not a sequence of identical consonants. Already McCarthy (e.g. 1982) analyzed this formation as augmentation by a consonantal mora on level 4a (stem-W) with the spreading of phonetic information from the second radical – in other words, not as reduplication.

There is a clear difference between this formation and what are generally referred to as reduplicated roots: These are roots in which the second and third radicals are identical, such as *rdd°* "to reply", and which are the result from trying to fit two-radical roots into the general three-radical template of Arabic roots. In the case of the "reduplicated roots", the phonetic realization depends on syllabification: the two identical radicals can be articulated both as geminates as well as separate consonants, whereas the geminate of the forticized second radical is never separated in the paradigm.

Thus in Old Arabic, only cases such as (10) and (11) qualify as reduplication in Arabic in the sense of this argument, not cases such as (13) and (14), and even less cases such as (15) and (16). In neo-Arabic varieties, the different types of doubling / reduplication are maintained and even elaborated. As neo-Arabic varieties are used especially in oral language, doubling is common as expressive device, sometimes even lexicalized, cf.

(17) Neo-Arabic doubling

a. Maltese: 
   a1. *xorob* *naqra* *naqra*  
      drink:PF.3.S.M little little  
      'he drank little by little'
   a2. *baxx* *baxx*  
      go.out:PF-3.P low low  
      'they went out discreetly'

b. Moroccan Arabic:
   b1. *mar°:a* *mar°:a*  
      'sometimes' (literally: *time time*)
   b2. *sasa* *sasa*  
      'from time to time' (literally: *hour hour*)
   b3. *bhal* *bhal*  
      'by the same way' (literally: *like like*)

Doubling as a device of word formation is frequent, cf. the following Maltese examples, most of which are without an Old Arabic predecessor:
Doubled roots in Maltese

\begin{itemize}
  \item \textit{capcap} ‘clap’
  \item \textit{ferfer} ‘shake’
  \item \textit{gerger} ‘growl’
  \item \textit{harhar} ‘rattle’
  \item \textit{laqlaq} ‘chat’ etc.
\end{itemize}

Partial “reduplication” in the so-called second stem is especially productive with an innovative function, that of forming denominal verbs, a device used only marginally in Old Arabic, cf. in Maltese: \textit{xemx} “sun” – \textit{xemmex} “to sun”, \textit{berqa} “lightning” – \textit{berraq} “to light”, \textit{sabar} “consolation” – \textit{sabbar} “to console”.

In the following I will restrict myself to one especially productive case of syntactic reduplication in the domain of the verb phrase (cf. (9c)): Arabic has a special grammatical category, the masdar, which is used to articulate the complement of the verbal predicate. But before analyzing the construction, a clarification of the category masdar is necessary.

### 3. Masdar formation

In most European grammars, \textit{verbal noun} is given as the equivalent of masdar; while older grammars refer to it an \textit{infinitive}, which is rather confusing, as the masdar does not participate in the verbal paradigm (to be distinguished from the verbal system, cf. below). On the other hand, the masdar should be distinguished from deverbal morphological devices. In a certain sense it is a syntactically underspecified form, built directly from the radicals. It has the potential to share with verbs certain syntactic properties, especially valency: In Old Arabic it can govern its complements by the two adverbal cases: \textit{-u} “nominative” for the main actant, and \textit{-a} “accusative” for the secondary actant. However, it can also function as the head of a nominal phrase, governing an attribute marked by the adnominal case \textit{-i} (“genitive”). Owing to its potential “verbal” properties, a masdar can be the head of a proposition but it cannot be the head of a narrative sentence (“verbal sentence”), as it cannot be marked for either TAM nor for actancy (person). As the head of a secondary predication, a masdar is marked as the complement of the predicate – i.e. it is marked by case, just as other nominal forms are. Cf. (19a), with the masdar \textit{?it\textasciitilde{}a:m} as verbal object, and (19b) with the masdar \textit{darb} as prepositional object.
(19) Masdar constructions (Classical Arabic)

a. ِرِنْنَا نَنَاْسَّا كَرِهٍٰ عُّ نِتْسَأْ مُهْمَأْدَيْن أُمْرٍّ أَنْ خُبْزٍ أَنْ مَأْسُمٍ أَنْ
   ‘the people condemned Mohammed’s giving Amr poisoned bread’

b. بِدَرَبْنِ بِي سُجُفٍ رُبْعٌ أَنْ قَوْمٍ أَنْ
   by-hitting-G [-DF] by-[+DF]-sword.P-G head.P-A people-G [-DF]
   ‘by beheading with swords the heads (accusative!) of people’

So far the category is rather well defined, and has been maintained in the restructuring of the neo-Arabic varieties. As case affixes are no longer present, it is difficult to analyze the nominal forms, as there is no morphological difference between complements of a verbal form and attributes of a nominal form. Cases such as the following present the structure of regular noun phrases (by juxtaposition):

(20) ُقَيْنِ أَلْ خُبْزٍ سَأَلَ وَالَا كَنِ مِلْغَرَد ُأَلْ وَالَا لِي
   knead.MD [+DF]-bread easy but [+DF]-flatten.M REL difficult a.little
   ‘to knead the bread (i.e. the dough) is easy, but to give (it) the flat form is a bit difficult’

(20) is a case of the lexically restricted use of the *iḍāfah-*construction, in which the head (ِقَيْنِ) is definite, although it cannot be overtly marked as such. This is even more evident when the complements are attached by the regular linking elements for nominal attributes as e.g. in (21), where the head (ِتَهْدَد) is marked as definite:

(21) ُتَهْدَدِ دَجَل أَلْ حَرْرِ أَلْ سُيْب
   [+DF]-iron.MD of [+DF-silk] difficult
   ‘the ironing of silk is difficult’

What makes the analysis of the masdar somewhat difficult is both the profusion of lexicalization processes, where some masdar formations have taken on specialized meanings, and also a rather bewildering set of morphological structures that can be used as verbal nouns, and which are presented as masdars by the grammarians (for Classical Arabic, Wright 1896, I: 110–112, mentions 44 different masdar formations corresponding to the
basic ("first") verbal stem). In the following I will focus on "basic" masdars, differentiated from verbal stems merely by vocalization, i.e. by specifying stem-W only at the vocalic melody (layers 2-V and 4a in (12)). External suffixation (layer 5), which characterizes many of the grammarians' masdars, makes them more noun-like, as we will see. Some examples of basic masdars from Classical Arabic, differentiated from verbal stems only by syllabification and / or vocalization, are given in (22)²⁷:

(22) Masdars

°drb° (daraba 'he has hit'):
  ḍarb(un) 'hitting', ḍarbatt(un) 'a blow'
°srq° (saraqa 'he has stolen'):
  sariq(un) 'stealing', sariqatt(un) 'a larceny'
°fsq° (fasaqa 'he has lived immorally'):
  fisq(un) / fasiq(un) 'viciousness', fasiqatt(un) 'a vice'

The suffix -at, given with the second nominal form in these examples, makes the masdar a countable unit: The form given is singular, the plural would be e.g. ḍarbattt(un), etc.

The vocalization patterns of the (basic) masdars can be found with isolated substantives as well, but without a verbal counterpart, as e.g. ⁴klb⁵: kalb(un) "a dog", ⁴r?s⁵ ra?s(un) "a head", ⁴qmr⁵ qamar(un) "a moon", etc.

A general characteristic of Afro-Asiatic languages is the high degree of abstractness of morphological devices, which do not have a constant function (meaning) but rather only a differentiating function, creating an ALTER to the base form. The meaning of a morpheme (and thus the meaning of the construction) can often only be deduced by taking into account its combination with the thus specified morphological form (the root + formation of stem-W and / or stem-G, cf. (12)). Thus the suffix -a(t) already mentioned has this kind of ALTER-function: If the term thereby modified is sexed, its meaning is "feminine", whereas if it is generic (collective, which includes the masdar as well), its meaning is that of a single unit (countability)²⁸, cf. (23)

(23) Nominal ALTER-formation

kalb 'dog' {sexed} : kalb-a(t) 'female dog' (and by antilogism: kalb 'male dog')
hut 'fish' {non-sexed ⇒ generic} : hut -a(t) 'a single fish' (plural hutatt)}
There is a certain overlap between these groups which are defined by formal structure, and not by abstract (universal) semantic features, cf. *xubz* “bread” {generic} : *xubz-a(t)“loaf of bread” (plural *xubza:t*) which is lexically related to verbal forms: *xabaz(a)“to bake”, masdar *xabi:z(un)“baking”, noun of unity *xabi:zat(un)“(one) act of baking”.

As the examples show, masdars can serve as inputs to all kinds of nominal formations, but not to verbal formations, which is the reason why traditional grammars treat masdars as nominals, independent of the verbal paradigm. This makes sense especially for stem formation (stem-W)\(^\text{29}\): In most cases masdars do not manifest a transparent formal link to the corresponding verbal stem formation. For example the so-called second stem has an associated masdar formation formed by a prefixed *ta-*, but lacking the doubled second radical which is the distinctive marker of the verbal second stem (cf. (16)):

\[(24)\] Masdars of the second stem (Classical Arabic)
\[
\begin{align*}
\text{o}rtb° (\text*rat:aba‘he has ordered’): \text*tarti:b(un)‘ordering’,
\text{o}drs° (\text*dar:asa‘he has taught’): \text*ta\text{dri:s(un)‘teaching’,}
\text{o}krm° (\text*kar:ama‘he has honored’): \text*takri:m(un)‘honoring’,
\text*takri:mat(un)‘(one) act of honoring’
\end{align*}
\]

This again makes it clear that masdars are not deverbal formations, notwithstanding the fact that they can be associated with specific verbal stems. At the morphological level I will treat basic masdars, i.e. formations which follow a certain pattern that distinguishes them from verbal formations, as intermediate between verbal and substantival forms, schematically.

\[(25)\] nominal-verbal scale

\[
\begin{array}{c|c|c|c}
\text{SUBSTANTIVES} & \text{MASDARS} & \text{(finite) VERBS} \\
\text{kalb…} & \text{xubza(t)} & \text{xubz} & \text{xabi:zat(un)} & \text{xabi:z(un)} & \text{xabaz(a)} \\
\text{“dog” “loaf of bread” “one act of baking”} & \text{“he baked”}
\end{array}
\]

Morphologically more specified forms, e.g. masdar formations with \(-a(t)\) (singular), \(-a:(t)\) (plural), are placed nearer to the substantive pole. This will serve as the background for the following analysis.
4. Syntactic reduplication within the verb phrase in Old and New Arabic

Grammars of Arabic usually have a special chapter where reduplication by the masdar within the verb phrase is discussed as “inner object”, “cognate object”, etc. The Arab grammarians use a purely formal term: *mutlaq* – the *free* object. (26) gives some representative examples with the stylistically equivalent translations, avoiding the literal translation with a substantive as object which would correspond to the characteristic Hebraisms (Semiticisms) in Bible translations.

(26) *mutlaq*-constructions in Classical Arabic

a. °drb°  ḍarab-a  ḍarb-an  
   ‘he hit violently (he struck a blow)’

b. °dʒrr°  dʒar:-a  dʒariːrat-an  
   ‘he committed a great crime’

c. °qtl°  ʔiqtatal-na:  qītaːl-an  
   ‘we fought hard’

d. °rnn°  ja-rin:-u  r-raniːn-a  
   3-ring:IPF-P  DEF-ringing-A
   ‘he rings the bell intensively’

e. °mwt°  maːt-a  mawt-an  
   ‘he died’

f. °ʃbʃ°  fabiː/uː  fabʃ-an  
   ‘they became completely satisfied’

Reduplication operates on the phonological information of the root (on the root-level, level 3 in (12)), not on syllabified structures. Thus at least in most cases, the medial radical has a different function in syllable structure in the verbal than in the masdar form. The translation shows the function of the construction: It expresses an intensification of the content of the verb (in a certain way corresponding to the verb-internal reduplication of stem-
formation II, cf. (16)). It is even found with otherwise intransitive verbs, as in (26 d–f: °rnn° “to ring”, °mwt° “to die”, °bdi° “to become satisfied”), which shows that the construction is not an instantiation of regular verb complementation (i.e. the presence of the masdar is not valency bound). From a semantic point of view this implies that the mutlaq complement does not have a referential interpretation but is a modification of the verb. As such the masdar in these constructions is not on the substantival pole of the scale (25) and in consequence should not be marked by the nominal specifications. These include, among others, specificity / definite marking: In this construction, the masdar cannot be specific. However, this does not necessarily preclude a formal definite marking (as e.g. in 26 d: r-rani:na): This marking (in most grammars called the “article”) can correspond in Arabic (Classical as well as Neo-Arabic varieties) to a generic interpretation and does not imply definiteness in the usual semantic-pragmatic sense.

Another indicator already used in the scale (25) is the use of the the masdar augmented by -at. As there do not exist extensive frequency analyses, the distribution of the two masdar forms (i.e. with and without -at) cannot be reliably established. What must be investigated are those cases where both forms are available, which is often not found because of the extensive lexicalization of these nominal forms. For example, the root °d3rr° in (26 b) has the basic meaning “to drag” and the meaning of the basic masdar djar: is “dragging” as well – only the masdar augmented by -at djar:i:rat(un) has the meaning “crime”, which is the basis of the interpretation of the masdar construction. Where both forms are available, the basic masdar (i.e. without -at) at first sight seems to be the default form. But this distribution might conceal a more structural difference, as the augmented form appears mostly with perfective verbs. This question will be investigated more closely in the Neo-Arabic varieties below (section 5 and 6), where intuitions of native speakers are accessible.

Frequently the construction is further expanded by an adjectival attribute of the masdar, thus:
(27) Masdar construction with attributes (Classical Arabic)
   a. °dțrr° dțar:-a dțari:rat-an tȧżimat-an
   ‘he committed a great crime’
   b. °žlm° žalal-tu nafsi: žulm-an
   kaḍiː:r-an
   frequency-A.S.[-DF]
   ‘I committed outrages frequently’
   c. °šbr° jaːbih-u sabr-an ḥasan-an
   ‘he endured beautifully’
   d. °xńq° xanːa-q-ahu xanːq-an
      fadiː:d-an
      violence-A.S.[-DF]
   ‘he strangled him violently’

There is in fact an equivalent to this construction in the modern Western European languages: As has been mentioned above (cf. 9), when an inner object is licensed it usually functions as the head of an attribute, cf. the English / German examples in (28), frequent e.g. in older epic texts:

(28) Inner objects with attributes in English / German:
   a. to strike a hard blow / einen heftigen Schlag tun, den Schlag eines Helden schlagen
   b. to die a beautiful death / einen schönen Tod sterben
   c. to live a good life / ein gutes Leben leben

Comparison of the Classical Arabic construction in (27) and the English / German one in (28) furnishes a key to the problem: There exist less literary stylized equivalents to (28) in English / German, but not in Arabic, cf. (29):

(29) Adverbs in English / German
   a. hit hard, hit heroically / heftig schlagen, heldenhaft schlagen
   b. die beautifully / schön sterben
   c. live well / gut leben
In fact, Arabic does not have a grammatical category of adverbs with a corresponding formative for the transposition of adjectives to adverbs as in Engl. -ly, German Ø-marking, etc\textsuperscript{32}. Adverbials in Arabic are nominal adjuncts to the predicate or to the propositional nucleus, marked by a preposition (governing the adnominal genitive case -i) or the adverbal “accusative” case (-a), which marks the masdar complement as well\textsuperscript{33}. Some of these nominal complements have become lexicalized and are as such used even in the Neo-Arabic varieties where the otherwise unexplainable case marking is retained as pseudo-adverbial marking\textsuperscript{34}, cf. (30):

(30) Morroccan “adverbs”
\begin{itemize}
\item \textit{matalan} ‘for example’ (SA \textit{madal ‘example’})
\item \textit{yaliban} ‘prevailing’ (SA \textit{ya:lib PCP of ya:lab “to prevail”})
\item \textit{haditan} ‘recently’ (SA \textit{hadi:θ ‘new’})
\item \textit{daiman} ([dima]) ‘always’ (SA \textit{da:ziim ‘during’ PCP of obsolete verb})
\item \textit{axiran} ‘lately’ (SA \textit{za:xir ‘last’})
\end{itemize}

Where the base is an adjective in Old Arabic, the masdar construction has probably served as a catalyst in its development: Adjectives in Old Arabic can easily be used in substantive function, thus a supporting masdar is not necessary in these cases. In fact, \textit{fadi:dan} (27e) is already used in Classical Arabic adverbially with the meaning “violently, intensely”. Reckendorf (1909) generalized this observation and saw in this as well as in the other cases of syntactic reduplication (“paronomasia” in his terms) the use of the masdar as a kind of pro-form.

Thus the masdar construction could be seen as a regular means of expressing adverbial modification in Arabic. For the moment this can only be a heuristic hint, based on translation equivalences with languages that have the (grammatical) category adverb. A more grounded analysis would presuppose a clarification of the category \textit{adverbial}, which at the moment is still a desideratum: Where work has been done, it has focused on adverbials that do not correspond to the masdar construction, mostly the so-called \textit{sentential adverbials}, i.e., higher predicates taking the sentence as their argument (cf. e.g. Auwera 1998). Besides the need of a general clarification, a closer analysis of the constructions competing with the masdar construction in Arabic will be necessary. In Arabic, we find as translation equivalents for adverbials in Indo-European languages the following types:
- adverbal complements in general (of which the masdar construction is, formally speaking, but a special case),
- prepositional phrases (adjuncts), i.e. higher predicates, taking the proposition (or at least the verbal phrase) as an argument,
- coverbal modification within a complex predicate, which is another peculiarity of Arabic, which does not possess infinitives\textsuperscript{35}.

5. Masdar reduplication in Neo-Arabic I: Maltese

The syntactic structure discussed above has been maintained in neo-Arabic varieties – or, more correctly, it has been reconstructed with sometimes quite radically changed phonological material. Besides the lexicalized elements, which may even maintain the otherwise apocopated old case marking $-a(n)$, cf. (30), the adverbial function in the sense of the last chapter is still expressed as in Old Arabic through adverbal complements, which comprise a more comprehensive set of functions / meanings than the usual classification as “objects” given in the grammars suggest\textsuperscript{36}, while masdar reduplication is a special type of adverbal complement\textsuperscript{37}.

Masdar formation has remained productive, but with some paradigmatic regularization, which is characteristic of all grammatical domains in the development of Neo-Arabic varieties. In the following I will focus on Maltese and Moroccan Arabic, a comparison between which is interesting because of the different cultural settings in which they have developed. The Maltese case will be presented first\textsuperscript{38}.

As masdar reduplication is in a certain sense an extension of predicate formation (in the sense of a reduplicative construction, see above), it is to be expected that it participates in the fundamental aspectual opposition of the verbal system as well. The synthetic paradigm of the verbal system in Arabic (Old Arabic as well as Neo-Arabic) is defined by the opposition of the marked perfective form and the unmarked imperfective form (with a wide range of interpretive possibilities). Temporal distinctions are made by inference or by an array of analytic modifications within complex predicate formation. This system is maintained (or reconstructed) in the Neo-Arabic varieties. Thus in Maltese, the fundamental opposition is between\textsuperscript{39}:
(31) Synthetic verbal paradigm of Maltese

- perfective (suffix conjugation): \textit{ktib-t} ‘I wrote’ (write:PF-1S), \textit{kil-t} ‘I ate’ (eat:PF-1.S)


Complex predicates are articulated by preverbal augments (only with the imperfective), e.g. progressive (PROGR) \textit{qed}: \textit{qed niekol} “I am eating”, inchoative (INCH) \textit{se}: \textit{se niekol} “I am going to eat”, or with coverbal modification for temporal situating, as e.g. \textit{kont} (\textit{kon-t} be:PF-1.S): \textit{kont niekol} “I have been eating”, \textit{kont kilt} “I have eaten”, etc.

As has been hypothesized above already for Classical Arabic, this fundamental opposition is in fact reflected in a difference in the masdar reduplication:

- the imperfective form (\textit{jiekol}) takes the bare masdar (ikel), corresponding to the unspecific / generic sense of the formation: \textit{jiekol ikel} “he eats a lot” (in a habitual sense),

- the perfective form (\textit{kiel}) takes the augmented masdar (ikla), corresponding to the concrete sense of an action (countable!): \textit{kiel ikla} “he has eaten a large meal”.

This constraint is grammatical – other combinations are generally refused, especially \textit{kiel ikel}, where the generic sense of the basic masdar (ikel) is incompatible with the perfective verb\textsuperscript{40}. The augmented form of the masdar is less restricted – in fact in many cases there is a lexicalization of a more concrete meaning: \textit{ikla} “eating” as well as “a meal” which combines easily with the imperfective and thus creates problems for testing. In other cases the specialised lexicalization is fixed: \textit{xarba} can only mean “a drink” – \textit{xorob xarba} “he drank a drink”, which can be further specified e.g. \textit{xorob xarba wiski} “he drank a (drink of) whiskey”.

\textsuperscript{40}See Tables 2-4 for examples.
Some further examples:

(32) Masdar construction in Maltese, continued

<table>
<thead>
<tr>
<th>°hrg°</th>
<th>jo-hrog</th>
<th>hrug</th>
<th>hareg</th>
<th>harga</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.S.M.-go.out:IPF</td>
<td>going.out</td>
<td>go.out:PF.3.S.M</td>
<td>going.out</td>
<td></td>
</tr>
<tr>
<td>‘he goes out extensively’</td>
<td>‘he went out for a walk’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>°hsl°</th>
<th>ta-hsel</th>
<th>hsil</th>
<th>hasl-et</th>
<th>hasla</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘she washes a lot’</td>
<td>‘she washed a lot’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>°hbt°</th>
<th>ja-hbat</th>
<th>hbit</th>
<th>habat</th>
<th>habta</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘he has a lot of accidents’</td>
<td>‘he had a big accident’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This analysis is compatible with A. Borg’s idea, who analyses the masdar as part of the verbal system of Maltese, participating in the fundamental aspectual opposition. In fact, he refers to the basic masdar as “nom imperfett” and the augmented masdar (forms with -a) as “nom perfett”\(^{42}\).

 Speakers are usually rather reluctant to give these forms in an elicitation context, especially when asked to construct them on the spot, and these forms are rather infrequent in narrative texts (not to speak of formal registers) which constitute the usual basis of linguistic corpora. There is evidently an expressive moment bound to this construction and they are thus most common in a lively spoken register. Therefore these expressions can be found in quoted speech, especially in very emotional verbal exchanges. In Maltese they are e.g. found in ballades such as in the traditional “Gharrusa tal Mosta” (The Bride from Mosta), where the young girl, kidnapped by Berber pirates, refuses all kinds of offers to fraternize with her kidnappers\(^ {43} \):

(33) \textit{ikel ma ir-rid-x n-iekol ... rqad ma}


\textit{ir-rid-x no-rqod}

\textit{1.S-wish-NEG* 1.S-sleep:IPF}

‘I will not eat at all ... I will not sleep at all’
This expressive function is accompanied by a non-narrative intonation. I checked this by asking my informants to produce sentences with masdar reduplication, and sentences with the same verb with a regular (referential) object as its complement, e.g. *haslet hasla* (M) “she washed a lot” and *haslet hwejeg tat-tfal* “she washed the clothes of the children”. The result was quite clear:

- The masdar reduplication was never produced with a falling (terminal) intonation, but usually with a kind of plateau prosody (also distinguished from question intonation);

- This intonation was never used with a referential object, where the terminal intonation was usually used.

In fact, the easiest way to obtain these kinds of structures was to elicit them in a frame like (34):

(34)  
\[
\begin{array}{cccc}
\text{t-ghid-x} & \text{kemm} & \text{j-iekul} & \text{ikel} \\
2.\text{S.say.IPF-NEG:IPT} & \text{how.much} & 3.\text{S.-eat:IPF} & \text{eating} \\
\end{array}
\]

‘You can’t even say how much he eats!’

As one would expect with this kind of expressive device, it is rather frequent when body functions are mentioned in discourse, but utterances such as the following will not be found in most corpora and are not easy to elicit (they are practically excluded if one works with female speakers). I omit here the translation of the masdar, which may be freely added with the meaning “incredibly much, loudly etc.”: *bass bassa* (MD) “he farted”, *biel bewla* (MD) “he pissed”, *hara harja* (MD) “he shit”, *tfewwaq tifwiqa* (MD) “he belched”, etc. These verbs are otherwise only intransitive, i.e. they do not admit any complement other than the masdar.

A further confirmation of this analysis is the general possibility of expansion by adding a further masdar with the augmentative suffix -un, an observation which I owe to Manwel Mifsud (p.c.): *habat habta* (MD) *habtun* (MD+un) “he had a terrible accident”, *kisirmi kisra* (MD) *kisrun* (MD+un) “he gave me a terrible thrashing” (*kiser*, lit. “he broke”). This enhancement is only licensed by the augmented masdar, cf.
Masdar reduplication is a productive grammatical device, but there are constraints to its productivity which I still only partially understand. In Maltese, the construction is restricted to the Semitic part of the vocabulary. Loan verbs (whether of Italian or English origin) are otherwise quite well integrated in the grammatical patterns, where even quite original procedures have been established to characterize them, as e.g. the doubling of the initial consonant, stem augmentation by -j-, etc., which quite regularly mark a verb form as an Italian or English loan ((it)tratta “to treat” < it. trattare, dajvja “to dive” < engl. dive etc.). But even where a corresponding verbal noun exists, usually formed by the suffix -ar (< it. -are !), this will not be used in a masdar construction. E.g. corresponding to llendja “to land” (< engl. to land) there is a verbal noun llendjar “a landing”, but a construction like *ji-llendja llendjar “he makes good (lots of ?) landings (???)” is not acceptable. These verbal nouns are apparently too far on the substantival side of the scale of (25) to license the masdar reduplication.

In other cases, a masdar is not available. This is the case for all denominal verbs. Even if there is a large array of more specific nominal formations, none of these is used in this kind of construction, cf. to xemmex “to sun, to expose to the sun” (derived from xemx “sun”): xemmiex “sunbather” and the hybrid (with a suffix of Romance origin) xemxata “sunstroke”. Here a “regular” masdar, corresponding to a second stem (see above), can even be built, tixmix, but cannot be used in this construction as a masdar (the prefix ti- pushes it to the substantival pole of (25)). To get a better understanding of the particular restrictions, the lexical fields of these verbs must be analyzed individually in each case to find reasons which might block the syntactic reduplication by a masdar.
6. The parallel: Moroccan Arabic

Although the same basic structure can be found in Moroccan Arabic as in Maltese, there are nevertheless some rather unexpected differences which are not yet entirely understood. Given the relative closeness of Moroccan Arabic to the common Arabic heritage, we would expect a stronger profile of the same aspectual correlation in masdar formation and its syntactic use in Moroccan. Here too, the basic masdar (vocalized with schwa or a peripheral vowel) is preferred with the imperfective form of the verb (in (36) augmented by the habitual prefix \(ka\)-), and the masdar augmented by \(-a\) is used with the perfective form of the verb:

(36) Moroccan masdar construction

\[
\begin{align*}
\text{frk} & \quad ka-i-frk-u & \quad frik (\text{*farka}) \\
& \quad \text{HAB-3.S.M-scrub:IPF-3.S.M} & \quad \text{scrubbing} \\
& \quad \text{‘he always scrubs him thoroughly’ (e.g. in a hammam)} \\
\text{frk} & \quad fark-u & \quad farka (? frik) \\
& \quad \text{scrub:PF.3.S.M-3.S.M} & \quad \text{scrubbing} \\
& \quad \text{‘he scrubbed him thoroughly’} \\
\text{hsd} & \quad ka-i-hsd & \quad hsd (\text{*hasda}) \\
& \quad \text{HAB-3.S.M-harvest:IPF} & \quad \text{harvesting} \\
& \quad \text{‘he uses to make a rich harvest’} \\
\text{hsd} & \quad hsd & \quad hasda (? hsd) \\
& \quad \text{harvest:PF.3.S.M} & \quad \text{harvesting} \\
& \quad \text{‘he (regularly) makes a rich harvest’} \\
\end{align*}
\]

But speakers are rather unsure as to the co-occurrence restrictions with the imperfective verb. The form augmented by \(-a\) is (as in Maltese) often synonymous with a more concrete meaning, and as such can be used with the imperfective form as well, cf.

(37) \(ka-i-fr\bar{a}\) \(frib / fr\bar{ba}\)

\[
\begin{align*}
& \quad \text{HAB-3.S.M-drink:IPF} & \quad \text{drink:MD} & \quad \text{drink:MD-a} \\
& \quad \text{‘he drinks a lot’} \\
\end{align*}
\]

But the non-augmented masdar is also compatible with the perfective form:

(38) \(fr\bar{a}\) \(fr\bar{ba} / frib\)

\[
\begin{align*}
& \quad \text{drink:PF.3.S.M} & \quad \text{drink:MD-a} & \quad \text{Drink:MD} \\
& \quad \text{‘he drank a lot’} \\
\end{align*}
\]
Determination (quantification, attribution) of the masdar requires the augmented masdar: *frab *farba (*frib) djal *safani “(lit) he drank the drinking of a thirsty”, i.e. “he drank like someone who was very thirsty”, *ka-i-frab *farba (*frib) *wahda “he always drinks only one draught”. Lexical restrictions may perhaps be responsible for these distributional restrictions. In some cases, both masdars are possible, licensing even anaphoric reference, as with *bi-ha / *bi-h in (39)⁴⁷:

(39)

a. *dr̲̊b a dr̲̊b qaṣha ma *ṣem:ar-u
hit:PF.3.S.M {1} hit:MD-a {2} thick.F NEG never-3.S.M {1}

hs̲̊b bi-ha
think:about:PF-3.S.M {1} with-3.S.F {2}

b. *dr̲̊b a *dr̲̊b qaṣah ma *ṣem:ar-u
hit:PF.3.S.M {1} hit:MD {2} thick.M NEG never-3.S.M {1}

hs̲̊b bi-h
think:about:PF-3.S.M {1} with-3.S.M {2}

,he hit harder than he ever thought possible’

An important factor is the function of the suffix -*a* in the nominal system. As was mentioned above for Classical Arabic, this suffix has no single, unifying function. Rather, its function must be calculated in relation to the semantics of the augmented stem: -*a* is a differential marking, encoding the “opposite” of the meaning of the base form in the relevant semantic dimension, cf. (40)⁴⁸

(40) Functions of the suffix -*a* in Moroccan Arabic

<table>
<thead>
<tr>
<th>base form</th>
<th>augmented form</th>
<th>dimension</th>
<th>marked form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ḥut ‘fish’</td>
<td>ḥut-<em>a</em></td>
<td>specificity</td>
<td>countable (unity)</td>
</tr>
<tr>
<td>ḳ̲̊lb ‘dog’</td>
<td>ḳ̲̊lb-<em>a</em></td>
<td>sex</td>
<td>female</td>
</tr>
<tr>
<td><em>naẓ:ar</em> ‘carpenter’</td>
<td>*naẓ:ar-<em>a</em></td>
<td>quantity</td>
<td>Plural</td>
</tr>
</tbody>
</table>

Perhaps *specificity* (SPEC) can be considered the common denominator of all these augmented forms. At any rate, this is a fundamental dimension in the system of nominal marking in Moroccan Arabic, without which the
otherwise confusing distribution of the so-called article (the prefix */-*) cannot be explained\textsuperscript{49}. The marking as specific underlies the use of the augmented form of the masdar. It can thus cut across the formal aspectual opposition, cf. (41) with an augmented masdar in construction with an imperfective verb form\textsuperscript{50}:

(41) \texttt{ka-i-\textsuperscript{3}zi\textsuperscript{a}} \texttt{ka-i-\textsuperscript{3}qrb\textsuperscript{a}-u} \texttt{\textsuperscript{a}qrb\textsuperscript{a}} \texttt{u}

\begin{tabular}{ll}
\texttt{PRO-3SM-come.IPF} & \texttt{PRO-3SM-hit.IPF-3SM} & \texttt{blow-SPEC} \\
\texttt{ka-i-tih} & \texttt{PRO-3SM-fall.IPF} & \\
\end{tabular}

\begin{tabular}{l}
\texttt{'he\textsubscript{3} comes, hits him, violently and he\textsubscript{3} falls'}
\end{tabular}

Sentences such as (41) are often found in narratives, especially in introductions to a longer episode. Here the use of the imperfective form with the augment \texttt{ka-} corresponds to our “historical present”. It thus it implies a specific (referential) interpretation of the event – not a generic one, which only hypostasizes the semantic content of the verb. This interpretation is implied with a verb in the perfective, which is the unmarked form for foregrounding in a narrative, where the augmented (i.e. specific) form of the masdar is at least preferred. In contrast, we will find the unaugmented base form with the unmarked use of the imperfective, i.e. in backrounding.

To understand the complex distribution of the masdar forms in Moroccan Arabic, their formal aspect must be taken into account. Moroccan Arabic can make the opposition between the (finite) verbal form and the masdar (noun) by epenthetic syllabification: i.e., in word forms without a vocalic melody on the lexical layer (level (4a) in (12), cf. (42):

(42) Noun – verb opposition in Moroccan Arabic

\begin{tabular}{llllll}
\texttt{\textsuperscript{o}`d\textsuperscript{b}rb\textsuperscript{o}} & \texttt{\textsuperscript{o}`d\textsuperscript{b}rb} & \texttt{\textsuperscript{o}`d\textsuperscript{b}hk\textsuperscript{o}} & \texttt{\textsuperscript{o}`d\textsuperscript{b}hk} & \texttt{\textsuperscript{o}`fr\textsuperscript{h}r\textsuperscript{h}} & \texttt{\textsuperscript{o}`fr\textsuperscript{h}r\textsuperscript{h}} \\
\end{tabular}

\begin{tabular}{llll}
\texttt{“he has hit”} & \texttt{“hitting”} & \texttt{“he has laughed”} & \texttt{“laughing”} & \texttt{“he has rejoiced”} & \texttt{“joy”} \\
\end{tabular}

Interestingly, this minimal phonological contrast does not seem to be exploited productively in the masdar construction: All forms I have found in the construction (cf. (36)) are more strongly differentiated\textsuperscript{51}, either by a suffix or a peripheral vowel ([dBiz] > [Dobz], and / or by a suffix: [D\textsuperscript{b}hka] > [D\textsuperscript{b}hk]). This corresponds to the results of my investigation of this morphoprosodic contrast\textsuperscript{52}: In most cases the masdar pattern [K\textsubscript{a}KK] is judged as archaic, as a classicised form (of the so-called “middle language”, the
lugha wusta\textsuperscript{53}). The most frequently used forms tend to have a high degree of lexicalization or are idiomatic, as the masdar form has taken on a transposed meaning, and as a lexicalized form no longer participates in the aspectual opposition (cf. 43a). Thus the expressive idioms in the field of bodily functions and / or activities license only one masdar form (usually the augmented form), independent of the aspectual form of the verb (43b). There are even some masdar formations following the pattern of the second stem (43c, cf. above remark to (24)):

(43) masdar construction with expressions for bodily functions

a. transposed meaning
\begin{itemize}
\item \textit{\textdegree{}dr\textdegree{} ka-i-d\textdegree{}rs dris (MD)} “he makes a good thrashing” –
\item \textit{d\textdegree{}rs d\textdegree{}rsa (MD)} “he made a good thrashing”, BUT
\item \textit{ka-i-d\textdegree{}rs-u d\textdegree{}rsa (MD)} “he gave him a thorough thrashing”
\item \textit{\textdegree{}ys\textdegree{}l\textdegree{} ka-i-\textdegree{}ysl ysil (MD)} “he washed a lot” –
\item \textit{ysl yasl\textdegree{}a (MD)} “he washed a lot”, BUT
\item \textit{ka-i-\textdegree{}ysl-u yasl\textdegree{}a (MD)} “he scolded him”
\end{itemize}

b. no basic masdar
\begin{itemize}
\item \textit{\textdegree{}hzq\textdegree{} hzaq hazqa (MD)} “he farted”
\item \textit{\textdegree{}xr\textdegree{}j\textdegree{} xa\textdegree{}r\textdegree{}a xarfa (MD)} “he shit”
\item \textit{\textdegree{}bw\textdegree{}l\textdegree{} bal bula (MD)} “he pissed”
\end{itemize}

c. with the “second stem”:
\begin{itemize}
\item \textit{\textdegree{}t-gr\textdegree{}r\textdegree{} t\textdegree{}grr\textdegree{}e tgr\textdegree{}i (MD)} “he belched”
\end{itemize}

More extensive research is necessary, based not on elicited forms but on a larger conversational corpus.

7. By way of conclusion – some general comments

In Arabic, syntactic reduplication is a regular grammatical device used to modify the verbal predicate in focusing its semantic content. As such it is characteristic of Arabic: It can be traced back to Old Arabic (Classical Arabic) and has been reconstructed in the neo-Arabic languages analyzed in this paper (Maltese and Moroccan Arabic). Yet this construction is only weakly grammaticalized, as it is used as a marked option in competition with unmarked expressions: In modern spoken varieties there is a tendency to restrict this construction to an evaluative, non-narrative register and in a regular expanded proposition other semantically more-or-less equivalent
devices will be preferred (e.g. coverbal modification or prepositional phrases).

As a regular construction, syntactic (masdar) reduplication participates in the fundamental dimensions of the verbal system, where a specificity dimension can be identified as the basis for the aspectual differentiation both in the verbal conjugation and the masdar forms. Even if stylistically restricted in its use, the construction remains productive in Arabic. Contrary to (Indo-)European languages, marked by a long school tradition where this construction has been frowned upon since antiquity, Arabic languages (and probably other Semitic languages as well) use this device rather freely, where grammarians’ treatises on stylitics even cherish this construction. These different stylistic predilections might reflect a fundamental typological difference between these language families: At our present state of knowledge we can only guess what this might be. A candidate made plausible by the argumentation in this paper is the need to possess a grammatical device corresponding to adverbial formations in e.g. Indo-European languages, which do not have a formal equivalent in Semitic languages.

This typological difference between Semitic and Indo-European languages might explain the restrictive stylistic prescriptions in the European school tradition, further analysis pending.

The restrictions imposed upon the use of paronomastic constructions in modern Indo-European languages are instructive in this regard:

- Paronomasia (figura etymologica) is generally avoided, as the “inner object” often has only a semantic, but not a formal similarity to the verb (einen grausamen Tod sterben, “to die a cruel death”, but Maltese miet mewta krudili “he died a cruel death”),

- It is generally allowed only as a support for an attribute, cf. *einen Schlag schlagen. In most cases it thereby compensates for the fact that the particular attribute does not permit an adverbial formation, cf. German den Schlag eines Helden schlagen, literally “to beat the beating of a hero (= to beat heroically)” and ?heldisch schlagen “to beat heroically” (cf. (28–29)).

By definition, all languages are dominated by the polarizing force of matching formal and semantic articulation, thus exploiting Smurf-like linguistic structures only as a special-purpose linguistic masquerade. Even in the Arabic languages, which seem to be more tolerant in this regard than Indo-European languages (as grammaticalized syntactic reduplication
shows), this construction is licensed only as a marked form, transporting specific expressive connotations.

Symbols and abbreviations

Arabic numerals, other than indices: 1,2,3 = the three grammatical persons

- A. accusative
- C. onsonant
- DF. initeness
  ([+DF] = definite, [-DF] = indefinite)
- G. enitive
- HAB. itative
- IPF = Imperfective
- IPT = Imperative
- M. asculine
- MD = masdar
- N. nominative
- N-Ag = agentive noun
- N-Res = resultative noun
- NEG. ation (NEG* = neg. circumfix)
- P. lural
- PCP = Participle
- PF = Perfective
- PRES. entative
- R. adical
- REL. ative particle
- S. ingular
- SA = Standard Arabic
- SPEC. ificity
- stem-W = stem formation in regard to valency and / or semantics
- stem-G = stem formation in regard to aspect and / or voice
- V. owel

Notes

1. This paper was presented at the reduplication conference, Graz, 3.-6. November, 2002. I am grateful to K. Ebert, R. Fabri, B. Hurch, J. Peterson and to an anonymous reviewer for comments on an earlier version of this paper.

2. Grammaticalization is understood here in the sense of the older typological tradition, where e.g. Boas defined grammaticalized structures as filters for “regular” utterances in specific languages, thus without the focus on diachronic processes as is generally the case in current discussions.

3. To give one example: “im Deutschen finden sie (i.e. Silben- und Wortdoppe-
lung, UM) sich in Form der einfachen Doppelung am ehesten in der Kinderspra-
che ... und in der Umgangssprache ...; gelegentlich auch in expressiven Ausdrucksbildungen” [in German reduplication is found in the form of simple doubling mostly in child language and in colloquial language; sometimes even in expressive formations], Drosdowski 1995: 427. However, the text mentions the fact that reduplication may be a “productive device of word formation” in other languages.

4. Reckendorf used the term paronomasia.
5. Arabic examples already figure prominently in one of the first typological studies on reduplication, Pott 1862.
7. If there were only sentences like (3), the language would be like playing cards only with jokers.
8. This problem could be dealt with in this way within the frame-work of constructional grammar.
9. Thus I follow Moravcsik’s (1978) definition, in which reduplication has to be defined on the formal as well as on the semantic / functional side of the reduplicands.
10. In Latin it is rather frequent in Plautus, but this only proves its connotation as “vulgar” (or rather, archaic). “Educated” writers tend to avoid at least paronomasias in the strict sense, substituting e.g. proelia pugnare for the earlier pugnam pugnare “to fight a battle”.
11. Note that these grids are not identical with those in McCarthy’s analysis.
12. An example of a modified stem of the same root °ktb° would be the (?i)ktataba “he subscribed” with the “infix” -t- after the first radical of the root.
14. Iteration as a quantitative augmentation encodes modification in the semantic dimension of quantity – be it augmentative or diminutive (cf. (2) above), a point already made by Moravcsik (1978).
15. To mention just the older tradition: Moravcsik (1978: 309) includes these cases under “internal reduplication“; Pott (1862: 17–18) treats them as “Reduplication von (consonantischen) Einzellauten” (with Arabic-Hebrew parallels).
16. This has been well known ever since Greenberg (1950), cf. e.g. Frisch & Zawaydeh (2001).
17. The radicals are marked by °°.
18. Wright 1896 (I: 31), Brockelmann (1913, I: 508) and others speak of the “doubling of the second radical”.
19. Cf. radda “he has replied” but radadtu “I have replied”, where the two radicals °dd° fulfill different syllabic functions.
20. There is no need of illustrating the type of (6), as the rich agreement systems of Arabic are well known.
21. Borg & Azzopardi-Alexander (1997) and the dictionary of Aquilina (1987–90). Maltese examples are given in conventional orthography, which isolates these augments as words. Pronunciation should be more or less transparent, with the exceptions: <ɛ> represents [, <q> represents œ], <g> represents [dʒ], <c> represents [tʃ]; in some cases I add non-orthographic morpheme segmentation to make the forms more transparent. Other neo-Arabic Varieties (Moroccan Arabic) are given in IPA-notation, with morphological segmentation through hyphenation.
23. Observe the distinction between *adverbal* and *adverbial*: *adverbal* means added to the verb (governed by the verb), *adverbial* means a free element (element of the propositional periphery).

24. Examples from Classical Arabic from Wright (1896) and Reckendorf (1895), respectively.

25. From Caubet 1993, II: 249. In (19) and (20) I give Caubet’s notation.

26. I.e., the genitive construction.

27. In the following, I quote forms from Classical Arabic, based on Wright 1896 and Reckendorf 1895. Some of these forms are not in current use in Modern Standard Arabic and not acceptable to my Arabic informants. This drives home the difference between Classical and Modern Standard Arabic, often ignored in the literature.

28. There are further functions of –a(t), not dealt with here.

29. Grammars present a corresponding set of masdar formations for every one of the fifteen verbal stems (stem-W).

30. Cf. Reckendorf (1895: 96 ff.)

31. But cf. section (6), especially example (37).

32. A word of caution is in order here. There are perhaps traces of a more differentiated case system in Old Semitic, with “adverbial” cases, which might have left traces in lexicalized sentential adverbs, cf. Brockelmann 1913: I: 459.

33. In fact, the suffix –an in the forms quoted above is complex: -a+n, where –n is the indefinite marking.

34. Quite plausibly they can be seen as traces of the specific diglossic situation in the Arabic (Muslim) world, being remnants of the ever-present “Classical Arabic” and a certain amount of rote learning to which grammatical structures remain opaque. An interesting confirmation for this is that these elements are not found in Maltese, which developed in a non-Muslim cultural context. In frequent forms, there exist in Muslim varieties more adapted alternatives as well, e.g. *dima* “always” (cf. (30)).

35. Coverbal modification is in most European grammars of Arabic treated as auxiliarization, which ignores the typologically significant difference between a complex predicate where all constituents are finite, as e.g. in

\[
\text{sā 'd-at} \quad \text{ta-kā:das-u}
\]

“it (the wheel, F) is always turning”

\[
\text{kā:na} \quad \text{na-hlik-u}
\]

“we almost perished”

The Arab grammarians discuss these verbal modifiers in a paradigmatic fashion, with the coverb *ka:na* “to be” as model under the title of “*ka:na* and her sisters”.

36. Cf. e.g. verbs of movement which are in a formal sense transitive, e.g. in Maltese *mar* “to go” (he went): *mar Malta* “he went to Malta”, *mar va:canza* “he went on vacation” etc., which have the same structure as e.g. *xtara hubz* “he bought bread”.

Bereitgestellt von | Vienna University Library
Angemeldet
Heruntergeladen am | 04.05.18 18:41
37. There are other translation equivalents, especially prepositional phrases (adjuncts), and coverbal modification, which has been considerably extended. For a study of coverbal modifications in Neo-Arabic, cf. Maas (1995) and Maas et al. (2000).

38. I am grateful to my main informants to whom I owe not only the examples given here but also most of the insights (if there are any) – even when these are due to a simple refusal of the examples fabricated by myself as a result of over-generalisation. For Maltese I had the help of Flavia Chetcuti and Ray Fabri, for Moroccan Arabic that of Redoine Hasbane and Mina Zinify.

39. The English translation by temporal forms is both conventional and misleading, as it suggests tense and not aspect.

40. A. Borg (p.c.) would prefer a “softer” formulation here: in an appropriate context (specifications added) such combinations could be possible, but would sound “forced”. Another problem are idioms which sometimes show other combinations such as \textit{ikel tas-sema} “a meal of heaven”, independent of the aspectual form of the verb, thus \textit{kiel ikel tas-sema}, is also possible.

41. Not acceptable for all speakers.

42. In fact A. Borg (p.c.) agrees with this analysis. There are of course cases of neutralisation, both through syntactic context (cf. note 40), as well as in the lexicon, where not all lexemes license contrastive forms.

43. Cf. the study of this ballade by G.C. Pullicino (1987-88).

44. The basis of 20 such sentences and two speakers is clearly not sufficient to warrant a definite statement. But it at least suggests that this is plausible. The prosodic pattern was independent of the order of production (the same result was obtained when masdar reduplication came first as when it came second). There where also a few examples of question intonation, possible with both structures.

45. Etymologically Italian –\textit{one}. This is one of the rare cases of hybrid Semitic-Romance formation and apparently is used to add expressive power. It is also found in other combinations: \textit{ghandu dar darun} “he has a huge house”.

46. Corresponding to Maltese –\textit{at}.

47. The augmented form of the masdar is treated in agreement relations as feminine, the basic form as masculine. Referential relations are indicated by indexes.

48. Examples of similar oppositions can be found in Maltese as well. My guess is that they are more lexicalized in Maltese, with the exception of the gender marking: \textit{kelb} “dog” – \textit{kelba} “bitch”.

49. There are other significant differences to the Maltese nominal system which, however, I cannot deal with here.

50. I owe this example (and other similar ones) to Mina Zinify.

51. Stronger differentiation is indicated by ‘ > ‘.

53. This term is used to refer to the registers of Arabic which make extensive use of classicized forms.

54. For a survey, cf. e.g. Paul (1920).

References

Aquilina, Joseph

Auwera, Johann von der (ed.)

Borg, Albert and Marie Azzopardi-Alexander

Brockelmann, Carl

Caubet, Dominique

Drosdowski, Günther (ed.)

Frisch, Stefan A. and Bushra A. Zawaydeh

Greenberg, Joseph

Leroy, Marie and Aliyah Morgenstern
this vol. Reduplication before age two.

Maas, Utz

Maas, Utz et al.

Marantz, Alec
McCarthy, John

Moravcsik, Edith A.

Paul, Hermann

Pott, August Friedrich
1862 *Doppelung (Reduplikation, Gemination) als eines der wichtigsten Bildungsmittel der Sprache, beleuchtet aus Sprachen aller Welttheile*. Lemgo: Meyer.

Pullicino, G.C.

Raimy, Eric

Reckendorf, Hermann

Wiltshire, Caroline and Alec Marantz

Wright, William