ON GENERICITY AND DEFINITENESS IN MODERN STANDARD ARABIC

Aziz Jaber

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Linguistics.

Chapel Hill
2014

Approved by:
J. Michael Terry
Randall Hendrick
David Mora Marin
Katya Pertsova
Dean Pettit
ABSTRACT

Aziz Jaber: On Genericity and Definiteness in Modern Standard Arabic
(Under the direction of Mike Terry)

The vast majority of formal analyses of the semantics of generics have been developed based on data from a small number of well-studied languages, most notably English, French, Italian, Dutch, and German. The main goal of this dissertation is to take steps towards the building of a fully formal analysis of genericity phenomenon in Modern Standard Arabic (MSA), one of the least studied languages with respect to genericity, by grammatically describing and semantically analyzing manifestations of genericity in the language. Accurately accounting for the distribution of morphology within MSA generic sentences and the range of possible meanings of the sentences themselves is a crucial warrant for building reliable formal models.

This dissertation argues that similar to English and other languages, MSA characterizing sentences require semantically indefinite NP subjects. In a departure from the view of the traditional Arabic grammarians, this dissertation argues against the classification of Arabic noun phrases as definite or indefinite based entirely on the presence or absence of the definite article al ‘the’, respectively. Instead, building on the work of Lyons (1999) and others, it makes use of definiteness criteria which consist of the semantic concepts of familiarity, identifiability, and uniqueness of the intended referent. Maintaining the traditional definiteness criterion leads one to the conclusion that characterizing sentences in Arabic use semantically definite NP subjects, a conclusion rejected here.

Also addressed is the status of bare NPs in MSA characterizing sentences. It is argued that bare NPs can be used in generic sentences, although only if modified directly by an adjective or a
relative clause, or indirectly when occurring in construct state, a noun form common to many Semitic languages that is found in MSA. Evidence is presented that these bare NPs are interpreted as indefinites. Verbless sentences in MSA can express habituality, a form of genericity. This dissertation proposes an account of the asymmetric distribution of copula ya-kuun in part on the availability of habitual readings to the sentences in which it is found. It is argued that ya-kuun only occurs in verbless sentences that are modified by adverbials and carry habitual meanings.
To my beloved wife Aisheh Othman, and my beloved sons Ahmad and Yazan
ACKNOWLEDGMENT

This dissertation would not have been possible without the support, guidance, and encouragement of a number of people. I am particularly indebted to Mike Terry, my advisor, for the many hours we spent in his office shaping, tweaking, and improving the final draft of this dissertation. In his office, I learned and appreciated formal semantics in general and the semantics of generics in particular. I would like to thank Mike Terry for reading over many rough drafts of the dissertation, and answering numerous questions.

I also would like to express my gratitude to the other members of my dissertation committee, Randy Hendrick, David Mora Marin, Dean Pettit, and Katya Pertsova, whose discussion during the defense, and incisive comments after helped bring this dissertation to a better shape.

Lastly, I would like to express my deep gratitude to my family, my wife in particular. Without her patience, support, and encouragement, this project might not have been possible. I would also like to express my thanks and appreciation to Kip and Susan Gerrard. This nice couple alleviated our otherwise tough transition to Chapel Hill, especially in the first three months. I would like to give a big thank you to Kip and Susan. My family and I are blessed to have you as friends.
# TABLE OF CONTENTS

LIST OF ABBREVIATIONS IN TEXTS AND FORMATS IN GLOSSES ......................... X

CHAPTER 1: INTRODUCTION .................................................................................. 1

1. The Overall Goal of the Dissertation ............................................................... 1

2. Main Issues of the Dissertation: An Overview ............................................ 1
   2.1 Definiteness in MSA Revisited ................................................................. 2
   2.2 Genericity in the NP System in MSA ....................................................... 5
   2.3 The Verbless Sentence: Generic vs. Nongeneric Distinction .................. 7

3. Organization of the Dissertation .................................................................. 9

CHAPTER 2: GENERICITY: AN OVERVIEW .......................................................... 12

1. Introduction .................................................................................................... 12

2. What is Genericity? ...................................................................................... 13

3. The Two Basic Types of Genericity ............................................................. 15

4. Generic vs. Episodic Readings .................................................................... 22

5. Generics vs. Explicitly Quantified Sentences ............................................ 27


7. The Semantics of the Generic Operator ..................................................... 36
   7.1 A Modal Approach to the Semantics of GEN ...................................... 37
      7.1.1 Kratzer’s Modal Theory ................................................................. 37
      7.1.2 Application of Kratzer’s Approach to the Semantics of GEN ........ 40

8. Summary ....................................................................................................... 42

CHAPTER 3: DEFINITENESS IN MSA REVISITED ............................................... 43

1. Introduction .................................................................................................... 43
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>What is Definiteness?</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>2.1 Grammaticization and Definiteness</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>2.2 Grammatical Definiteness vs. Semantic Definiteness</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>2.3 Semantic Definiteness: Successful Reference</td>
<td>53</td>
</tr>
<tr>
<td>3.</td>
<td>Definiteness and Genericity</td>
<td>61</td>
</tr>
<tr>
<td>4.</td>
<td>The Traditional Grammarians’ View of Definiteness in Arabic</td>
<td>65</td>
</tr>
<tr>
<td>5.</td>
<td>Definiteness in MSA Revisited</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>5.1 The Canonical Use of the Definite Article al in MSA</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>5.1.1 Familiarity</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>5.1.2 Identifiability</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>5.1.3 Uniqueness</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>5.2 Other Uses of al</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>5.2.1 The Use of al in Indefinite Environments</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>5.2.2 Obligatory al in Indefinite Environments</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>5.2.3 Other Obligatory Uses of al</td>
<td>95</td>
</tr>
<tr>
<td>6.</td>
<td>Conclusion</td>
<td>100</td>
</tr>
<tr>
<td>...................................................................</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHAPTER 4: ON GENERICITY IN NPS IN MSA</td>
<td></td>
<td>102</td>
</tr>
<tr>
<td>1.</td>
<td>Introduction</td>
<td>102</td>
</tr>
<tr>
<td>2.</td>
<td>The Noun Phrase in MSA: A Syntactic Overview</td>
<td>104</td>
</tr>
<tr>
<td>3.</td>
<td>Genericity and the Nominal Phrase Distinction in MSA</td>
<td>110</td>
</tr>
<tr>
<td>4.</td>
<td>Genericity and the Definite NP in MSA</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>4.1 Characterizing Sentences in MSA: The Traditional View</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>4.2 A Characterizing Sentence: What Kind of Generalization is Expressed?</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>4.3 Characteristic Genericity in MSA: Is it Language Peculiar?</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>4.3.1 Characterizing Sentences and the Definite NP with al Determiner</td>
<td>120</td>
</tr>
<tr>
<td>5.</td>
<td>Reference to a Kind and the Definite NP in MSA</td>
<td>133</td>
</tr>
</tbody>
</table>
6. Genericity in the Construct State .......................................................... 139
   6.1 A Syntactic Overview ...................................................................... 140
   6.2 Generics Associating Construct State Phrases ............................... 142
   6.3 Construct State and Reference to a Kind Genericity ...................... 149
7. On Genericity in Bare NPs ................................................................. 152
   7.1 The Grammatically Indefinite NP: Syntactic Facts ......................... 153
      7.1.1 The Grammatically Indefinite NP: True or Pseudo Bare Noun .... 154
   7.2 Genericity and Bare NPs in MSA Revisited .................................. 156
8. Conclusion .......................................................................................... 162

CHAPTER 5: ON GENERICITY IN THE VERBLESS SENTENCE IN MSA ..... 165
1. Introduction ....................................................................................... 165
2. The Verbless Sentence in MSA: Asymmetrical Behavior of the Copula ... 167
3. The Semantics of the Verbless Sentence in MSA .............................. 171
   3.1 Obligatorily-Absent Copula .......................................................... 172
   3.2 Optionally-Present Copula ............................................................. 178
      3.2.1 Habituals: Two Contrasting Views ........................................ 179
   3.2.2 The Semantics of Verbless Sentences with Optionally Null Copula 184
   3.2.3 The Semantics of Verbless Sentences with Optionally Overt Copula 188
   3.3 The Semantics of the Verbless Sentence with Obligatorily Overt Copula 190
4. Manifestations of Genericity in the Negative Verbless Sentence in MSA 194
5. Conclusion ....................................................................................... 198

CHAPTER 6: CONCLUDING REMARKS .............................................. 201

BIBLIOGRAPHY ................................................................................... 204
LIST OF ABBREVIATIONS IN TEXTS AND FORMATS IN GIOSSES

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2, 3</td>
<td>1st/2nd /3rd person</td>
</tr>
<tr>
<td>Acc</td>
<td>Accusative</td>
</tr>
<tr>
<td>al-N</td>
<td>The-Noun</td>
</tr>
<tr>
<td>AP</td>
<td>Adjective Phrase</td>
</tr>
<tr>
<td>AdP</td>
<td>Adverb Phrase</td>
</tr>
<tr>
<td>BP</td>
<td>Bare Plural</td>
</tr>
<tr>
<td>CCQ</td>
<td>Closed Class Quantification</td>
</tr>
<tr>
<td>CS</td>
<td>Construct State</td>
</tr>
<tr>
<td>D-generic</td>
<td>Definite Generic</td>
</tr>
<tr>
<td>f</td>
<td>Feminine</td>
</tr>
<tr>
<td>Fut</td>
<td>Future</td>
</tr>
<tr>
<td>Gen</td>
<td>Genitive</td>
</tr>
<tr>
<td>GEN</td>
<td>Generic Operator</td>
</tr>
<tr>
<td>I-generic</td>
<td>Indefinite Generic</td>
</tr>
<tr>
<td>I-level</td>
<td>Individual Level</td>
</tr>
<tr>
<td>K-level</td>
<td>Kind Level</td>
</tr>
<tr>
<td>m</td>
<td>Masculine</td>
</tr>
<tr>
<td>MSA</td>
<td>Modern Standard Arabic</td>
</tr>
<tr>
<td>N</td>
<td>Nunation</td>
</tr>
<tr>
<td>N</td>
<td>Bare Noun</td>
</tr>
<tr>
<td>NEG</td>
<td>Negation</td>
</tr>
<tr>
<td>Nom</td>
<td>Nominativ</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>OCQ</td>
<td>open Class Quantification</td>
</tr>
<tr>
<td>Pl</td>
<td>Plural</td>
</tr>
</tbody>
</table>

x
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF</td>
<td>Phonetic Form</td>
</tr>
<tr>
<td>PP</td>
<td>Prepositional Phrase</td>
</tr>
<tr>
<td>Pres</td>
<td>Present</td>
</tr>
<tr>
<td>Q</td>
<td>Quantifier</td>
</tr>
<tr>
<td>q</td>
<td>Question</td>
</tr>
<tr>
<td>sg</td>
<td>Singular</td>
</tr>
<tr>
<td>S-level</td>
<td>Stage Level</td>
</tr>
<tr>
<td>s-select</td>
<td>Semantically Select</td>
</tr>
<tr>
<td>VP</td>
<td>Verb Phrase</td>
</tr>
<tr>
<td>?</td>
<td>Awkward</td>
</tr>
<tr>
<td>*</td>
<td>Ungrammatical</td>
</tr>
<tr>
<td>Ø</td>
<td>Null</td>
</tr>
</tbody>
</table>
CHAPTER 1: INTRODUCTION

1. The Overall Goal of the Dissertation

Most formal analyses of the semantics of generics have been developed based on data from a small number of well-studied languages, most notably English, French, German, Italian, and Spanish. The main goal of this dissertation is to take steps towards a fully formal analysis of genericity in Modern Standard Arabic (MSA, henceforth), one of the least studied languages pertaining to this linguistic phenomenon, by grammatically describing and semantically analyzing manifestations of genericity in the language. MSA manifests some properties that are shared crosslinguistically and some language-specific features pertaining to the articulation and interpretation of genericity. This study aims to contribute to the crosslinguistic study of genericity initiated by Carlson (1989, 1995) and Dahl (1995), which focuses on how the distinction between generic and non-generic sentences is formally marked in the grammars of natural languages. This type of project is a necessary first step towards achieving the crosslinguistic insights that come from building formal models. Getting the distributional facts and data right precedes building up formal models which abstract away from actual data and distributional facts to arrive at abstract principles. Some of these principles are argued to be universal like the tripartite semantic structure of characterizing sentences (see Leslie 2008). Therefore, building these principles based on data and facts taken from a wide array of languages is crucial for the reliability of these principles.


Generics like the English sentences in (1) report regularities about the world, which establish generalizations about groups of individuals, objects, events, situations, or state of affairs.
(1) a. Oranges contain vitamin C.
   b. Ducks lay eggs.
   c. Tigers are striped.
   d. Aziz drinks tea after dinner.

The interaction between genericity and other semantic phenomena like definiteness has strong bearing on the investigation of genericity in MSA. A crucial difference between the two major sub-phenomena of genericity is subsumed under the definiteness status of the NP upon which a generalization is reported. The first, D-genericity or reference to a kind, like the English example in (2), which is gleaned from the noun phrase when invested to refer to a kind or ‘genus’, restricts generic NPs to semantically definite NPs. The second, I-genericity or the characterizing sentence, like the English example in (3), however, which is established by the entire sentence in which the predicate and the other constituents collaborate compositionally to express a generalization (Carlson 1995, Cohen 1999, Greenberg 2003, Krifka et al. 1995, Pelletier 2010b, among many others), requires a semantically indefinite NP as its subject.

(2) Dodos are extinct.
(3) A mother loves her children.

In (2) the kind-referring noun ‘dodos’ is semantically definite because it refers directly to a unique kind, not to any individual or a group of individual dodos. This is a basic characteristic of D-genericity, in that NPs employed as kind-referring expressions are semantically definite, regardless of their grammatical in/definiteness; i.e., whether they are morpho-syntactically definite or indefinite in form. The sentence in (3) reports a generalization about mothers in general, not about a particular, semantically definite individual mother. This benchmark distinguishes a characterizing sentence from its non-generic counterpart. The semantic structure of a characterizing sentence limits the denotation of the NP subject upon which a generalization is reported to semantically indefinite referent; in the sense that it denotes variable individuals which satisfy the descriptive content of the noun. If the NP subject is tied to a particular individual; i.e., the NP is definite, the sentence would be non-generic, reporting an isolated fact about a particular individual at a specific time and place (Krifka et al. 1995).
The traditional Arab grammarians’ view of determining in/definites based entirely on what turn out to be the inconsistent grammaticizations of this semantic contrast poses problems for the investigation of genericity, a phenomenon that interacts with the semantic in/definiteness of a noun phrase regardless of its morphological in/definiteness status. A major goal of this dissertation is to propose more modern theoretically inspired criteria for determining in/definiteness in MSA that will, in turn, make possible a clearer view of the various manifestations of genericity in MSA. The resulting analysis is previewed in (2.1) below.

2.1 Definiteness in MSA Revisited

Noun phrases in MSA, according to the traditional grammarians’ view, are classified into definites and indefinites based entirely on the presence/absence of the so-called definite article *al* ‘the’. Noun phrases that contain *al* are definite, while those that do not are not (c.f. Al-saamirra?ii 2011; Al-ɣalaaliinii 1993; Fassi Fehri 2012; Hatoum 2007; Hoyt 2008). Moreover, indefinite NPs in subject position must be interpreted existentially. Consider (4).

(4) a. *ya-ʕab-u al-walad-u fi faariʕ*  
Pres-play-Nom the-boy-Nom in the-street  
(The boy is playing on the street).  

b. *ya-ʕab-u walad-u-n fi faariʕ*  
Pres-play-Nom a boy-Nom-N in the-street  
(A boy is playing on the street)

According to the traditional view of definiteness, each of the sentences in (4) reports an isolated state of affairs concerning a particular, existentially calculated individual boy. In (a) *al-walad* ‘the boy’ is definite by virtue of the presence of the definite article *al*. It denotes a particular definite boy known to the speaker and the hearer. The sentence says that a contextually salient unique individual boy is involved in a playing event at a particular time and place. The boy in (b), however, is indefinite; i.e., it does not refer to a familiar individual boy. This is due to the absence of the definite article *al*. Therefore, the sentence says that there is a boy who is playing on the street at the time of utterance.

The traditional view of identifying definites based on the presence of the definite article is problematic if one wants to investigate genericity, characterizing sentences in particular. If we accept
the traditional view of classifying in/definites, we end up with a problematic either or situation pertaining to characterizing sentences in MSA. Either MSA does not have characterizing sentences, or MSA’s characterizing sentences exhibit a strange property in that the subject NP upon which a generalization is made is semantically definite; i.e., it is tied to a particular individual. The first of these options is on the face of it untenable. Characterizing sentences are a widespread phenomenon that speakers of the language use all the time to report regularities true of variable individuals, events, situations, and states of affairs. Without such a device, it is unclear how speakers of a language could discuss basic facts about the world, convey the cominality of experiences, or develop science. The other possible outcome, that characterizing sentences in MSA make use of semantically definite subjects, although less startling is still problematic. A number of promising formal analyses of generic sentences assume that indefinites introduce logical variables into their semantic representations, and that these variables are either bound by a generic operator that has the structure of an adverb of quantification or are otherwise introduced into the restricor of their sentence’s quantificational structure (e.g. Cohen 1999; Greenberg 2003; Krifka et al. 1995; Pelletier 2010b). A very different type of analysis would be suggested for MSA sentences if the traditional view of definiteness is maintained.

This dissertation, however, rejects the traditional view of determining in/definites based solely on the absence/presence of the definite article al. It builds on Lyons (1999) crucial distinction between semantic definiteness, which is established on concepts like familiarity, identifiability, and uniqueness, and grammatical definiteness. Grammatical definiteness is the morpho-syntactic encoding of semantic definiteness, in the case of MSA, the definite article al. This morphological marking is neither a necessary and sufficient condition for determining definiteness in MSA. In fact, more often than not, al-Ns are found in indefinite environments, in the sense that an al-N does not denote a definite individual, but rather denotes any individual which satisfies the descriptive content of the NP. MSA definites, this dissertation asserts, must meet several criteria which incorporate
familiarity, identifiability, and uniqueness which are taken as pathways to successful reference to a particular individual on both the speaker’s part and the hearer’s part. Consider (5).

(5) a. [Two students have noticed that their classmates treat their professor very respectfully; after class, one of them says the sentence below]

\[
\text{al-ʔustaað-u} \quad \text{ya-hðʻaa} \quad \text{b-iḥtiraam-i-n} \quad \text{kabiir} \\
\text{the-professor-Nom} \quad \text{enjoy} \quad \text{with-respect-Gen-N} \quad \text{great}
\]

(The professor enjoys high respect)

b. [An employee in a university is talking to her office mate about her feelings that she does not receive enough respect from students. She expresses her wish of becoming a professor one day. She says the sentences below]

\[
\text{al-ʔustaað-u} \quad \text{ya-hðʻaa} \quad \text{b-iḥtiraam-i-n} \quad \text{kabiir} \\
\text{the-professor-Nom} \quad \text{enjoy} \quad \text{with-respect-Gen-N} \quad \text{great}
\]

(A professor enjoys high respect)

The subject NP \text{al-ʔustaað} ‘the-professor’ in (a) is semantically definite because its referent is familiar to the interlocutors; both of them have been in his class and witnessed how much respect he received from his students. It is definite not in virtue of the presence of \text{al} solely, which functions here as a real definite marker, but rather based on the definiteness criteria, the referent is familiar. In (b), however, the same subject NP used in the same sentence, but in a different context, is semantically indefinite, though it is grammatically definite. \text{al-ʔustaað} ‘the-professor’ here denotes not a particular individual professor, but rather any individual to which the descriptive content of the NP applies. It denotes variable individual professors. This is the kind of denotation that fits the semantic structure of a characterizing sentence.

The distinction between definiteness as a semantic phenomenon and its inconsistent grammaticization has strong bearing on investigating manifestations of genericity in MSA. In particular, determining definites based on the definiteness criteria is crucial to investigating genericity in the NP system in MSA. This interaction is previewed below in (2.2).
2.2 Genericity in the NP System in MSA

In MSA, noun phrases of different syntactic forms can be used in lexically characterizing sentences as long as these NPs are semantically indefinite, regardless of whether they are grammatically definite or not. Any al-N, singular, plural, countable, or uncountable, associated with Individual-level predicates, classified as denoting permanent, essential, definitional properties, can be used in a characterizing sentence if semantically indefinite. The traditional view of bare nouns restricts their denotation in subject position to existentially calculated individuals, therefore preventing them from being used in generic sentences. Contrary to this view, bare NPs, this dissertation argues, are allowed in characterizing sentences. Not only are bare NPs allowed in generic sentences, so are grammatically indefinite construct state phrases. Construct state is a type of annexation consisting of at least two members, most often nouns. The first is the possessed and the last is the possessee. In fact, a characterizing sentence incorporating a bare subject NP not only expresses a generic reading, but also excludes the other episodic reading that its counterpart with al-N subject avails. Consider (6) below.

(6) a. fadżarat-u z-zaytuun-i tu-ʕārīi zayt-a-n sīḥhiyy
   tree-Nom the-olive-Gen Pres-give oil-Acc-N healthy
   (An olive tree gives healthy oil)

   b. fadżarat-u zaytuun-i-n tu-ʕārīi zayt-a-n sīḥhiyy
      tree-Nom olive-Gen-N Pres-give oil-Acc-N healthy
      (An olive tree gives healthy oil)

The sentence in (a) is ambiguous. It avails an episodic reading when a context that limits the denotation of the subject construct state fadżarat-u z-zaytuun-i ‘the olive tree’ to a particular definite referent is provided. In this case, the sentence reports an isolated fact about a particular olive tree producing healthy oil. The other reading is generic. The sentence expresses a generalization about olive trees in general such that they produce healthy oil. The ambiguity that (a) entertains emerges from the subject construct state being grammatically definite. In this case, grammatically definite construct state nominal phrases behave similarly to grammatically definite simple NPs. The sentence in (b) uses grammatically indefinite construct state fadżarat-u zaytuun-i-n ‘an olive tree’. The
sentence expresses a generic reading as the only available reading. This is due to the indirect modification that holds between the construct state head fadgarat-u ‘a tree’ and its complement zaytuun ‘olive’. This kind of modification changes the denotation of the nominal phrase from denoting individuals existentially to denoting variable individuals which satisfy the descriptive content of the noun phrase. An NP with this kind of denotation is allowed in a characterizing sentence. Being semantically indefinite restricts the compatibility of grammatically indefinite construct state phrases to characterizing sentences only, rather than particular sentences.

This indicates that bare NPs can be used in characterizing sentences when modified indirectly in a construct state, or directly, in a noun adjective phrase. All nominal forms, this dissertation argues, are allowed in generic sentences. This shows that MSA maps to other natural languages that put no restrictions on the noun forms used in generic sentences (Krifka et al. 1995).

Reference to a kind genericity in MSA manifests itself in semantically definite NPs that denote not particular individuals but kinds. Singular and plural nouns can be used in the language as generic NPs. However, the language under examination, as well as other natural languages, restricts kind-referring NPs to nouns which denote well-established kinds. Many languages like English entertain this semantic condition. In the literature of reference to kind genericity (e.g. Krifka et al. 1995), adverbial phrases like bifakl-i-n qaam ‘in general’ are used as a diagnostic test to distinguish generic NPs which denote well-established kinds from NPs denoting objects. Nominal phrases compatible with these adverbs are defined as object-referring, and those incompatible are kind-referring. Consider below.

(7) a. al-ʔusuud-u    mina  ʔ-ʔadiyyaat
    the-lions-Nom  from  the-mammals
    (Lions are mammals)

    b.?bi-fakl-i-n    qaam     al-ʔusuud-u    mina  ʔ-ʔadiyyaat
    in-form-Gen-N   general t the-lions-Nom  from  the-mammals
    (In general, lions are mammals)

(8) a. ʔsayyaarat-u    l-ʔudgyrat-i   ʔafraaʔ
    car-Nom  the-fare-Gen  yellow
    (A/The taxi is yellow)  (bad on kind-referring reading)
Another major manifestation of genericity in MSA is the so-called verbless sentence, a type of sentence which contains a copula, but no other verb. The syntactic distribution of the copula in these sentences is asymmetric. This dissertation provides a semantic account for this asymmetry based on a generic/nongeneric distinction. This account is previewed in (2.3) below.

### 2.3 The Verbless Sentence: Generic vs. Nongeneric Distinction

The nonmodal present tense verbless sentence in MSA shows a clear manifestation of genericity in the language. It exhibits a syntactic asymmetry in that the copula ya-kuun ‘Pres-be’ is obligatorily unexpressed in some sentences (9), obligatorily expressed in others (10), and optionally expressed in other environments (11).

1. **(9)**
   
   a. Ø/*ya-kuunu  al-djaww-u  djamil
      Pres-be  the-weather-nom  nice
      (The weather is nice)
   
   b. Ø/*ya-kuunu  alkab-u  dakiyy
      Pres-be  the-dog-nom  intelligent
      (The dog is intelligent)

   **(10)**
   
   a. ya-kuunnul/*Ø  xaalid-u-n  naa'im-a-n  min  l-Sija'a?i  ?ilaa  l-fadgr
      Pres-be/Ø  Khalid-Nom-N  asleep-Acc-N  from  the-evening-Gen  to  the-dawn
      (Khalid is asleep from evening to dawn)
   
   b. ya-kuunnul/*Ø  yazan-u  masruur-a-n  fi  nihaayaat-i  l-?usbuus
      Pre-be/Ø  Yazan-Nom  happy-Nom-N  in  ends-Gen  the-week
      (Yazan is happy in the weekends)

   **(11)**
   
   a. Ø/ya-kuunu  l-djaww-u  djamil-u/a-n  fi  r-rabi'is
      Pres-be  the-weather-nom  nice-Nom/Acc-N  in  the-spring
      (The weather is nice in spring)
   
   b. Ø/ya-kuunu  as<s>s'iyaam-u  s'asb-u/a-n  fi  s<s>s'aayf
      Pres-be  the-fasting-Nom  hard-Nom/Acc-N  in  the-summer
      (Fasting is hard in summer)

---

1 This study investigates verbless sentences where the copula is phonologically (un)expressed regardless of whether the copula is in the underlying structure or not. This issue is subject to much debate, and has no direct influence on the treatment proposed here based on a generic-non-generic meaning distinction. However, the semantic analysis proposed in this dissertation might give some insights with respect to this syntactic controversy.
This dissertation proposes that the puzzling distribution of the verbal copula *ya-kuun* in verbless sentences can be captured by the generalization in (12), which is based on a generic/nongeneric distinction.

(12) The distribution of copula *ya-kuun* in the present tense verbless sentence

a. Copula *yakuun* must be licensed by a spatio/temporal adverbial.

b. Present tense verbless sentences with such an adverbial come in two kinds:
   1. Copula *yakuun* is optional, with a generic-nongeneric meaning difference based on the choice of the predicate and the definiteness status of the copula subject.
   2. Copula *yakuun* is required based on the choice of the predicate and its interaction with the adverbial.

c. The presence of a spatio/temporal adverbial does not guarantee the presence of *yakuun*.

d. Present tense verbless sentences without adverbial licensors, and therefore without copula *yakuun* can receive varied generic/nongeneric interpretations based on the predicate used as copula complement and the nominal form used as copula subject.

The copula *ya-kuun* could be treated as a habitual marker; hence being compatible with verbless sentences that have time/place referential adverbials. These adverbials provide a set of eventualities for the habitual operator to quantify over. The lack of such adverbials block a habitual reading of the sentence, and this explains why copula *ya-kuun* cannot surface in this environment.

Sentences where *ya-kuun* can optionally surface vary in their interpretation based on the presence/absence of the copula. Sentences with expressed copula entertain a generic interpretation only. Whereas, their minimally contrasting sentences with null copula are ambiguous between generic/non-generic interpretations according to the context which helps determining the associated *al-Ns* as semantically definite or indefinite. The former is compatible with an episodic reading, and the latter with a generic reading. Sentences where an expressed copula *ya-kuun* is required entertain a generic interpretation only. The incompatibility of the predicate and the adverbial used excludes the episodic reading of such sentences. Therefore, the presence of copula *ya-kuun* marks the only generic

---

2 It is hard to tell whether the eventualities are contributed by the adverbial adjuncts or the copula itself. However, a tentative suggestion that the copula, not the adverbial adjuncts, provides these eventualities will be proposed based on some observations.
reading that such sentences avail, where phonologically absent copula renders such sentences odd, if not completely unacceptable.

3. **Organization of the Dissertation**

The reminder of this dissertation is organized as follows. Chapter 2 will present an overview of genericity as a linguistic phenomenon in natural languages. An elaborated distinction between genericity as a linguistic tool to express generalizations about the world and episodicity which covers sentences used to express accidental, isolated facts, state of affairs, events, or situations true of particular individuals or objects will be a main goal of this chapter. The long-established distinction between the two types of genericity, reference to kind and characterizing sentences will be examined. An elaborated distinction of generics as opposed to explicitly quantified sentences will be probed. The two broadly defined camps which dominate approaches to generics analyses - the inductivist view and the rule-and-regulations view, as articulated in Carlson (1995) - will be investigated. The semantics of GEN (generic operator) assumed in the semantic structure of characterizing sentences will be investigated. It will be argued that a modal interpretation similar to Kratzer’s (1981) is sufficient to interpret characterizing sentences since it fully accounts for the exception tolerance that characterizing sentences entertain as opposed to quantified sentences.

Chapter 3 will revisit definiteness in MSA. In this chapter, it will be shown that the traditional view of determining in/definites based exclusively on the presence/absence of the definite article *al* ‘the’ is problematic and unuseful to genericity investigation. Definiteness is a semantic concept. A definite nominal expression is such that its referent is a particular, identified individual to both interlocutors. It gains this privilege of being identified through a variety of linguistic or extralinguistic contexts: the referent is contextually unique; it is familiar through shared background; it can be identified and singled out; or it has been mentioned before in a previous sentence. In all these cases, both interlocutors are able to single out and successfully refer to an individual or object in the outside world without running the risk of confusing it with another individual to which the descriptive content of the linguistic expression applies. The semantic concepts - familiarity,
identifiability, or uniqueness - upon which definiteness is established will be taken as criteria to definiteness in MSA. This does not mean ignoring the grammatical marking of definiteness represented by the article * ál*, but rather approaching nouns prefixed with it with suspicion, as this article is no longer a pure, exclusive marker of definiteness.

Chapter 4 will investigate the variety of noun phrases in MSA with special focus on those which are allowed/disallowed in generic sentences. Bare, mass, definite, and referential NPs will be scrutinized in episodic and generic environments to see whether the language uses any of these constructions to formally mark generic or episodic readings. It will be shown that all nominal forms in MSA can be used in generic sentences as well as episodic sentences. This, in turn, shows that the language does not dedicate a specific nominal form to mark genericity or episodicity. The interaction between semantic definiteness and characterizing sentences will also be investigated. It will be shown that characterizing sentences in MSA are ubiquitous, and that they have a semantic structure similar to that of characterizing sentences in other natural languages. The second of these conclusions could not be reached if the analysis were to accept the traditional view of definiteness in MSA. Reference to a kind will be investigated. Both definite singular and definite plural nouns will be shown compatible with kind level predicates, and thus both constructions appear equivalent in their use as kind-referring NPs. The construct state and its compatibility with generic sentences will be probed in this chapter.

The goal of chapter 5 is to address a long-observed syntactic asymmetry within the non-modal present tense verbless sentence in MSA. Within this structure, the copula *ya-kuun* ‘Pres-be’ is obligatorily absent in some contexts, obligatorily overt in others, and optionally overt in still other contexts. This chapter aims to explain this asymmetry in terms of the generic/nongeneric interpretation of the sentence. The claim is that sentences that require a phonologically null copula vary in their interpretation based on the predicate used and the definiteness status of the subject NP. These sentences do not make use of temporal adverbials
that are necessary for such sentences to receive habitual interpretations. Sentences which optionally allow overt copula, however, can be classified according to the phonological status of the copula: sentences with expressed copula can only be interpreted generically, while those with unexpressed copula are ambiguous between both readings according to the context. Sentences in which the copula is obligatorily expressed avail a generic reading only. This generalization is extended to cover the negative verbless sentence where two negative particles, laysa and laa ‘not’, appear in complementary distribution. The latter is compatible with verbless sentences with a generic reading only. The former, however, is compatible with sentences where copula ya-kuun is unexpressed.
CHAPTER 2: GENERICITY: AN OVERVIEW

1. Introduction

Generic sentences, as the term suggests, express generalizations about the world. They are used to convey regularities that hold for an individual, group of individuals, or across members of a certain kind. They do so in an intriguing way. As pointed out by Leslie (2007) speakers of natural languages, English, for example, intuitively judge a sentence like ‘dolphins give birth to live young’ as true, and ‘dolphins are females’ as false. This judgment is puzzling because it holds despite the fact that the former reports a generalization true of less than a quarter of dolphins - only adult fertilized female dolphins can give birth to live young - whereas the latter expresses a generalization true of at least half of dolphins. What adds to the oddness of such puzzling judgments is that the set of dolphins which give birth and the set of dolphins which are female are in a subset/superset relation; thus a dolphin which gives birth to live young is necessarily a female dolphin. This and other aspects of their behavior have made generics a lively and interesting topic, intriguing enough to constantly trigger new insights, claims, arguments, and counterarguments in order to construct a full account of their intricacies.

The main purpose of this chapter is to present an overview of genericity as a linguistic phenomenon in natural language. The essence of genericity as a linguistic tool to express generalizations, and the long-established distinction between the two types of genericity, reference to kind generics and characterizing sentences, are investigated in section two. An elaborated distinction between generics as opposed to both episodics - sentences that express accidental, isolated episodes, events, or facts - and explicitly quantified sentences is investigated in sections three and four, respectively. Section five will examine and evaluate the two broadly defined camps which dominate approaches to generics analyses; namely, the inductivist view (Lawler 1973), and the rule-and-
regulations view, as articulated in Carlson (1995). Most theories of generics postulate a logical form with a silent generic operator GEN in characterizing sentences. However, they differ tremendously in the type of semantics each theory proposes to this operator. This chapter will briefly review the dominant methods introduced in the literature to capture the semantics of GEN; in particular, the modal interpretation of Kratzer (1981) will be given special focus as the most promising approach.

2. What is Genericity?

Generic sentences, as exemplified in (1), are pervasive and widely used by speakers to express generalizations for which exceptions exist. Generic propositions are distinguished from their episodic counterparts (2) in that they do not attribute an accidental property to a particular individual or a group of individuals; neither do they report isolated events, situations, or states. Rather, generics report a generalization or regularity over groups of individuals, objects, events, situations, or state of affairs. In their seminal work, Krifka et al. (1995:2) define generic sentences as statements that denote “propositions which do not express specific episodes or isolated facts, but instead report a kind of general property, that is, report a regularity which summarizes a group of particular episodes or facts.” It is said that generics often isolate themselves from any restrictions over the number of individuals, the time line, or specific situations or events (Cohen 1999). In a nutshell, generics, unlike their episodic counterparts, are almost unqualified in these respects.

(1) a. Cats meow.
   b. A vulture eats meat.
   c. Graduate students study hard for exams.

(2) a. Cats are mewing.
   b. A vulture ate meat yesterday.
   c. Graduate students studied hard for exams last spring.
The propositions in (1) abstract away from relating a certain property to a particular individual or group of individuals on a particular location and time, and report a regularity that summarizes groups of particular episodic instances or facts. For instance, (1.a) reports some kind of property attributed to cats in general. It captures a generalization based on properties of specific members of the set of cats. Similarly, (1.b) does not describe a specific individual vulture, but rather reports a property ascribed to vultures in general. The proposition in (1.c) also does not seem to describe a particular group of graduate students; it expresses a statement about graduate students in general. This generalization, however, is based on a regularity of recurrence of a plurality of events. The proposition in (1.c) can be roughly interpreted as follows: “generally speaking, situations which have graduate students and exams in them are such that graduate students study hard for exams in these situations.”

The propositions in (2), however, describe specific episodes or state of affairs. The NPs in (2) are interpreted existentially since they denote particular individuals or groups of individuals. Similarly, the predicates report accidental, isolated facts or state of affairs. The proposition in (2.a), for instance, describes an accidental event of meowing attributed to particular cat individuals at a particular place and time. Like (2.a), the proposition in (2.b) relates a specific property about a particular vulture that happened to eat meat at a specific time and location. Similarly, (2.c) does not seem to report a generalization about graduate students in general; instead, the proposition describes an isolated fact about some graduate students who were involved in a particular situation of studying hard last spring. These sentences receive what is called an existential interpretation. A helpful diagnostic for the existential interpretation of a sentence is that the sentence may be paraphrased with some or the existential operator there with little or no change in meaning; hence the terminology ‘existential reading’ (Leslie 2007). This kind of sentence can felicitously fit in an existential frame like ‘There is X Ying)’ or ‘Some X Y’, where X stands for any subject referent, and Y stands for the predicate attributed to the subject. The sentences in (2) fit in both frames, and can be rephrased with some or there with slight change in meaning (3). However, inserting any of the sentences in (1) in the
‘There frame’ renders that sentence unacceptable, whereas inserting it in the ‘Some frame’ radically changes the sentence from a generic sentence to an existentially quantified one, and hence a significant change in meaning (4).

(3) a. There is a/Some cat Ø/is mewing.
b. There is a/Some vulture that/Ø ate meat yesterday.
c. There are/Some graduate students who/Ø studied hard for exams last spring.

(4) a. *There are/Some cats mew.
b. *There is a/Some vulture eats meat.
c. *There are/Some graduate students who/Ø study hard for exams.

Genericity covers two basic sub-phenomena, which are distinct and related at the same time:

Reference to a kind (D-generics), and characterizing sentences (I-generics). They are investigated in (2.1) below.

3. The Two Basic Types of Genericity

The distinction between two types of genericity was first proposed in Krifka (1987). There he distinguished D-generic sentences and I-generic sentences. The former are statements about kinds; in such sentences, an NP is used to refer directly or indirectly to an abstract kind or genus, rather than to an individual or a group of individuals of a certain kind. Genericity here is ascribed to the noun phrase. The latter, I-generic sentences, in contrast, are statements about instances of kinds. In this type of generic sentences, genericity is represented as a feature of the sentence as a whole, where genericity is not attributable to any subpart of the sentence - the subject noun phrase or the verb phrase. This type of genericity is rather realized as an interaction between these two constituents (Pelletier 2010b). It is worth mentioning that these two types can co-occur in one generic sentence, which is used to describe a regularity that holds across members of a certain kind, usually through predicking this regularity directly to the kind. An example for this case is given in Pelletier (2010b):

‘The polar bear is a white animal’ is a characterizing sentence with a generic NP subject, in which the property ‘whiteness’ is predicated directly to the Polar Bear kind.

These two terms are treated as equivalent to the more recent terms reference to kind generics and characterizing sentences, respectively. Following Krifka et al. (1995) the more recent terms will be used invariably throughout the dissertation.
D-generic stands for ‘definite generic’, and I-generic for ‘indefinite generic’. These names are intended to suggest that usually the NP employed to refer to a kind is a definite NP; whereas indefinite NPs typically occur in generic statements about specimens of kinds. Link (1995) claims that D-generic sentences incorporate “Proper Kind Predication”; that is, a D-generic sentence serves to express a singular statement about a particular kind. I-generic sentences, however, express “Derived Kind Predication”; that is, an I-generic sentence is used to express a generalization about a certain object or group of objects of a certain kind.

This distinction between different types of genericity has been adopted by many scholars working in formal semantics and the philosophy of language (Carlson 1995, Cohen 1996, Greenberg 2003, Krifka 1987, Krifka et al. 1995, Pelletier 2010b, among many others). D-generics are sometimes called Reference to Kind generics, and I-generics, Characterizing Sentences (Krifka et al. 1995). The sentences in (5) and (6) are examples of each class of sentence.

(5) **D-Generics (Reference to Kind Generics)**

a. The Dodo is extinct.
b. Cats are widespread.
c. Water is common.
d. Dinosaurs died out long time ago.
e. The Wright brothers invented the airplane.
f. Man reached the moon in 1969.

(6) **I-Generics (Characterizing Sentences)**

a. Ahmad walks to school in the morning.
b. Water freezes at zero Celsius.
c. A tiger eats meat.
d. A copier copies 100 copies per minute.
e. Dinosaurs ate kelp.
f. Humans are mammals.

Each reference to kind generic in (5) contains an NP that refers directly to an abstract kind or genus, rather than to a particular individual or object. In such sentences, we are assured that the predicates apply directly to the kind, not to members of the kind. This is evident since predicates like ‘extinct’ and ‘widespread’, are not applicable to individuals; no individual can be extinct or widespread, only kinds or species can. This holds also for ‘invent’; no object item can be invented;
the type or the kind of that item can only be invented. Thus, all the NPs in the above example sentences refer to kinds, not to members of kinds. In both Carlson (1977) and Krifka et al. (1995) this kind of NPs are dubbed generic NPs as opposed to object-referring NPs, and the kind of predications associated with them generic predications, as opposed to object predications. The sentences in (6), however, express propositions thought to be generically true not in virtue of the NP only, but the sentence as a whole. Sentence (6.a), for example, expresses a regularity of situations in which Ahmad walks to school in the morning. In (6.b), however, the regularity reported expresses a fact about situations where water is present, and the temperature reaches zero degrees Celsius, water freezes.

The sentences in (5) support our intuition that grammatically definite NPs do not exhaust the range of linguistic objects that can apparently be kind-denoting. Bare plural NPs and mass nouns behave like definite NPs in this respect. Indefinite and quantified NPs can also serve to accomplish a specific type of kind reference known as “taxonomic kind reference” (Krifka et al. 1995). The data in (5) also show that kind-referring NPs are not restricted to the subject position of the sentence; they can occupy the object position as well. The data in (7) provide more examples of the range of generic constructions.

(7)  a. The lion has a bushy tail. Definite NP
    b. Flies are common. Bare Plural (BP)
    c. Water is necessary for life. Mass noun
    d. A large cat, namely the tiger, is in danger of extinction. Indefinite NP
    e. Several metals are precious. Quantified NP
    f. Gulf countries’ economies are dependent on oil. Object position.
    g. William Shockley invented the transistor. Object position

An interesting question that semanticists have been pondering over for a while relates to the kind of relation that holds between a kind and the members of that kind. Put differently, what is the semantic relation between an individual NP and a generic or kind-denoting NP? Carlson (1977) proposes two types of individuals, objects and kinds. The former realizes the latter. Individual people, for example, exemplify Mankind kind, and individual dinosaurs exemplify the Dinosaur kind. However, another related question might arise: How does a member of a kind receive a property applied to the kind as a whole? A plausible answer might be that since an individual is a member of a
certain kind, this puts it in a position where it is eligible to inherit a property shared by all members of a kind unless it has a certain anomaly that makes it an exception to that property. This way of thinking is supported by the long-observed fact that reference to a kind can be accomplished by referring to a single, particular object of a kind. Pelletier (2010b) argues that such indirect reference to a kind can be attained through two different ways.

(8) **Representative Object Interpretation**  
(taken from Pelletiere 2010b:6)

a. In Alaska we filmed *the grizzly.*
b. Look Children: *this* is the reticulated giraffe.
c. Quiet!!-- *The lion* is roaming about!

Although the reference in the sentences in (8) is made to particular objects, which suggests that the predication is relevant to that object; it is evident that the three sentences express kind-referring generics, and the NPs employed are generic NPs, rather than individual NPs. Pelletier (2010b) observes that in the sentences in (8), only one object must be relevant, and yet this seems enough to make a truth about the whole kind. Another way of referring to a kind indirectly using a specimen of that kind is given in (9).

(9) **Avant-garde Interpretation**

(b. The potato was first cultivated in South America.

Our intuition about the sentences in (9) is that although both sentences express reference to a kind genericity, particular members of the kinds were involved in the event. More precisely, it is true that in (9.a) one particular man, namely Neil Armstrong, first set foot on the moon in 1969; similarly, in (9.b) particular potatoes were first cultivated in South America. The significance of this is that these sentences are true of the kinds *Man* and *Potato*, respectively, because of actions relevant to some first particular instances of these kinds. The name ‘avant-garde’ indicates that this kind reference is essentially temporal.

Unlike characterizing sentences, whose underlying semantic structure is believed by many semanticists to incorporate a tripartite quantificational structure that contains a covert generic
operator, a restrictor, and a matrix; reference to kind generics are represented in most theories of
genericity as involving no quantification, but instead a property is directly predicated to the kind as an
entity (Greenberg 2003). Carlson (2010) argues that adopting a quantificational approach to the
semantics of reference to kind generics yields the wrong truth conditions. The subject according to
this account is treated as a predicate, rather than a referring expression. More precisely, on this
account a sentence like the one in (10a) can be mistakenly represented as (10.b)4:

(10) a. The dodo is extinct ≠ b. *Qx [x is a dodo] [x is extinct]

The formula in (10.b) gives the wrong truth conditions of the sentence in (10.a). The formula
says that there is some “Q pattern of substitutions” (Carlson 2010: 19) of individuals for the x that
will make [dodo(x) → be-extinct(x)] true. The formula fails to give the right truth conditions for
(10a.) because if something is extinct, there is intuitively no x that would satisfy this condition. This
analysis conflicts with our intuition simply because an individual dodo cannot be extinct; only the
species or the kind can be extinct. It seems more plausible to treat the subject dodo in (10.a) as a fully
referential expression denoting a kind of thing, namely Dodo. Thus, this sentence can be better
represented by another formula that treats the noun phrase subject as a kind-referring expression (See
Carlson 1977).

(11) Extinct (d) (where d represents the kind Dodo)

This proposed semantics of kind-referring NPs complies with the proposal in Krifka et al.
(1995) that kind-referring NPs can be semantically analyzed as proper names in terms of their
function. Both are definite referring terms, which pick out a contextually unique individual in the case
of the proper name, and a unique kind entity in the case of the generic NP.

The other type of genericity is characterizing sentences. As mentioned above, characterizing
sentences - unlike reference to kind generics - predicate properties of their subjects that can hold of

4 Unlike the example I cite here which is a clear reference to a kind sentence, hence completely infelicitous on
a quantificational account, the example cited in Carlson (2010) is a characterizing sentence with a generic
subject NP, ‘Cats have fur’. This example of Carlson might be analyzed quantificationally, for those adopting
the quantificational approach, but Carlson rejects analyzing it on this approach for a number of reasons
presented in the study.
particular members (6). In addition to this, characterizing sentences impose no limitation on the type of NP they take (7). This feature explains why Krifka et al. (1995) analyze this type of genericity as “sui generis type of sentence” (P.8). It is widely believed that generic sentences are stative sentences since a stative sentence expresses a property that the referent possesses; whereas a dynamic or eventive sentence, which has a predicate with a situation argument bound by existential closure, reports an event in which the referent is involved (Krifka et al. 1995).

A significant distinction in characterizing sentences is emphasized in Krifka et al. (1995). This concerns “habitual” and “lexical” sentences. The former, they argue, entertains a verbal predicate that is morphologically related to an episodic predicate commonly used to form episodic sentences, while the latter most often does not exhibit this property. While lexical characterizing sentences generalize over characteristic properties of individuals, habitual sentences generalize over episodes of events. It seems intuitive, therefore, that lexical characterizing sentences employ stative verbs like know, cost, weigh, love, hate, etc., which are “inherently independent of situations” (P.36). Habitual sentences, however, incorporate dynamic verbs that lose their dynamicity, and change into stative verbs by the effect of an operator, like the adverb usually, which is different from the existential closure operator. Canonical examples of dynamic verbs are eat, run, walk, play, and the like. Consider the sentences in (12&13)

(12) **Lexical Characterizing Sentences**

<table>
<thead>
<tr>
<th>Diagnostics:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Sugar dissolves in water.</td>
</tr>
<tr>
<td>b. Mark knows German.</td>
</tr>
<tr>
<td>c. Cats have fur.</td>
</tr>
<tr>
<td>d. Lucy likes meat.</td>
</tr>
</tbody>
</table>

(13) **Habitual Sentences**

<table>
<thead>
<tr>
<th>Diagnostics</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Lucy barks when frightened.</td>
</tr>
<tr>
<td>b. Mark speaks German at home.</td>
</tr>
<tr>
<td>c. Cats eat raw meat.</td>
</tr>
<tr>
<td>d. Jill smokes a cigar after dinner.</td>
</tr>
</tbody>
</table>

The sentences in (13) can be rephrased by sentences with quantified adverbs like typically with only a slight change in meaning resulted. (13.d), for instance, can be paraphrased as Jill
typically smokes a cigar after dinner. However, inserting typically in any of the sentences in (12) renders it unacceptable and semantically odd; this is because the verbal predicates used in these sentences are not morphologically related to predicates that can be used to express episodic events. For example, (12.d) is a lexical characterizing sentence because it contains a non-episodic verb like, which cannot occur in an episodic sentence; *Lucy is liking meat is ungrammatical.

However, both subtypes of characterizing sentences have some properties in common. In both subtypes, the insertion of used to retains genericity and the use of progressive blocks it. Consider the sentences in (14).

(14) a. Bill eats meat.  
    b. Bill used to eat meat. (genericity retained)  
    c. Bill likes meat.  
    d. Bill used to like meat. (genericity retained)  
    e. Bill is eating meat. (genericity is blocked)  
    f. *Bill is liking meat. (ungrammatical)

It is worth mentioning that reference to kind generics and characterizing sentences share a distinctive salient property. Kind-referring NPs abstract away from particular instances of the kind, and characterizing sentences abstract away from particular episodes. Thus, it seems plausible and intuitive that the two types of generics can occur together in the same sentence. In this case, the sentence is used to describe a regularity that holds across members of a certain kind. This can be usually attained by predicating this regularity directly of the kind (Pelletier 2010b). The data in (15) lay this point out.

(15) a. Snakes are reptiles.  
    b. The lion is ferocious.  
    c. Polar bears are white.  
    d. Birds fly.  
    e. Gold is a precious metal.

The sentences in (15) are true characterizing sentences. They generalize over characteristic properties of their subject referents. However, these properties do not hold of a particular individual or group of individuals; they hold of kinds. In (a), for example, the characteristic property of being reptile is attributed to the kind Snakes, not to any individual snake(s). Similarly, (b) expresses a
generalization about the kind Lion by predicating the property ferocious to the Lion. The rest of the sentences behave in the same fashion.

4. Generic vs. Episodic Readings

The line separating generic/nongeneric readings of a sentence is not always easy to draw. In fact, the majority of sentences are ambiguous with respect to the kind of interpretation they express, although the more salient interpretation can often be gleaned through context and other pragmatic machinery. The data in (16) illustrate this point.

(16) a. Bill smoked cigarettes.
    b. A dog barks.
    c. Jill will eat meat.

The sentences in (16) are ambiguous between generic/episodic readings. Each sentence exhibits both a generic interpretation and a particular interpretation. On a more salient reading, the proposition in (16.a) can be construed as describing an accidental property of smoking cigarettes attributed to Bill, which happened at a particular time prior to the time of utterance. This is the episodic reading of the sentence. However, the sentence exhibits another less salient reading which expresses a generalization based upon a group of particular events of smoking cigarettes in the past, and this habitual regularity is attributed to Bill. An accurate paraphrase of the generic reading of (16.a) can be given by incorporating the phrase used to ‘Bill used to smoke cigarettes’. The ambiguity in (16.b) does not unfold as clearly as in (16.a). It might seem that (16.b) can only be interpreted as expressing a generic reading. However, the episodic reading can be attained if the predicate barks is taken to express the so-called ‘reportive present’ (Krifka et al. 1995). Similarly, (16.c) seems to report a particular fact about Jill - that on a time after the utterance time Jill will be involved in an eating event, and meat will be eaten by Jill in that particular event. This is the more salient episodic reading. However, what if Jill is raised in a strict vegetarian family and her parents noticed that Jill does not seem to take the family’s dietary conventions and values very seriously. One of her parents could utter the sentence in (16.c) expressing a dispositional interpretation of the sentence, which can be paraphrased as ‘Jill will not object to eating meat’.
However, an interesting question related to the source of this pervasive ambiguity with respect to generic/episodic readings might arise. A plausible, simple answer might be that sentences, particularly those presented with no explicit context, present such a challenge for native speakers because genericity is hardly marked formally in natural language; and if marked, the marking is usually inconsistent. Put more precisely, most formal encodings of genericity in natural language are necessary but not sufficient. Carlson (1989:3) emphasizes this point about a specimen of the languages he surveyed in his typological endeavor pertaining to the formal markings of genericity in natural languages. In his words, “English and most other commonly accessible European and Asian languages have no consistent formal expression of genericity.”

Seeking a perspicuous differentiation between generic/episodic readings of a sentence, Krifka et al. (1995) summarized a set of diagnostic tests proposed by a number of semanticists in the literature that would help disambiguate such sentences and determine whether a sentence expresses a generic or episodic interpretation. A non-exhaustive list of five tests is introduced. Some of these tests are intended to determine whether the sentence expresses characteristic genericity, and others are supposed to determine whether the NP incorporated is generic, denoting a kind, or non-generic, denoting an individual or a group of individuals. The first test, which has been mentioned previously, uses adverbs of frequency like typically or usually. The claim is that if one of these adverbs is inserted in a sentence, and the new sentence exhibits only a slight change in meaning, the original sentence is generic, and is episodic otherwise. The data in (17-20) flesh this claim out.

(17) a. A vulture eats meat.
    b. A vulture typically eats meat.

(18) a. Yazan drinks tea for breakfast.
    b. Yazan usually drinks tea for breakfast.

(19) a. A vulture ate meat on top of a tree.
    b. A vulture typically ate meat on top of a tree.

(20) a. Yazan drank tea for breakfast yesterday.
    b.*Yazan usually drank tea for breakfast yesterday.
In (17&18) the insertion of *typically* and *usually*, respectively, brings a slight change in meaning, although as Krifka et al. claim the original sentences express strong, law-like generalizations; whereas their quantified counterparts express weak generalizations that might not sound law-like. This slight change in the strength of the generalization might be attributed to the explicit stating of exceptions to the rules given in the sentences. In the case of (19), however, the incorporation of the adverb *typically* causes a significant change in meaning. While the original sentence describes a specific event of eating meat true of a particular vulture at a particular place and time, its quantified counterpart reports a regularity of eating meat events. Put differently, the original sentence is episodic, but the resulting one is generic, precisely habitual. The following formal representations proposed by Krifka et al. (1995) are intended to capture the difference.

(21) a. $\exists [s, x, y, z;] (s \text{ contains } x, y \& z; z \text{ is top of a tree}; x \text{ is a vulture } \& x \text{ ate } y \text{ on } z \text{ in } s; y \text{ is meat})$

   a'. Generic $[s, x, y, z;] (s \text{ contains } x, y \& z; z \text{ is top of a tree}; x \text{ is a vulture } \& x \text{ ate } y \text{ on } z \text{ in } s; y \text{ is meat})$

The pair in (20), however, presents a huge change not only in terms of meaning, but also in the acceptability of the quantified sentence, which is rendered ungrammatical.

The second test is designed to clearly distinguish two kinds of referring NPs; namely, the kind-referring NP and the object-referring NP. The former is assumed to denote an abstract kind or ‘genus’. The latter, however, denotes an individual or groups of individuals of a kind. The test draws upon a number of predicates like *invent, extinct, die out*, thought to semantically-select kind-referring NPs only. The moral is that if an NP fits in the argument position of one of these kind-level predicates, this NP is proved to be kind-referring, and is object-referring otherwise. The data in (22) instantiate this claim.

(22) a. The Pied Raven died out in 1948.
   b. The Great Auk is extinct.
   c. Alexander Graham Bell invented the telephone in 1876.

---

No specific semantics of Generic operator is assumed here; what this operator is supposed to do is to function as an indicator of genericity. See section 6 for a discussion of the semantics of the generic operator.
Close scrutiny of the NPs in (22) reveals that none of them denotes an individual or an object; rather all NPs refer to an abstract kind or ‘genus’. The *Pied Raven* in (22.a), for instance, denotes the genus *Corvus corax varius morpha leucophaeus*, and similarly the *Great Auk* and the *telephone* denote the kinds *Pinguinus impennis* and *Telephonium*, respectively. This analysis can be bolstered by the fact that the above predicates do not s-select specimens of the same kinds above. Consider the data in (23).

(23) a.* This embalmed pied raven died out in 1948.
    b.* A great auk is extinct.  (*Acceptable on a taxonomic reading only*)
    c.* Alexander Graham Bell invented a telephone in 1876.

However, the propositions in (22) would remain semantically true and grammatically acceptable if the scientific kind names of the *Pied Raven*, the *Great Auk*, and the *Telephone* are substituted.

    b. *Pinguinus impennis* is extinct.
    c. Alexander Graham Bell invented the *telephonium* in 1876.

Third test is proposed in Krifka et al (1995) to differentiate between kind-referring NPs and object-referring NPs. However, this test relies heavily on the semantics of these NPs rather than on their grammatical distribution pertaining to kind-level predicates’ selection. The test shows that not any nominal constituent can form a kind-referring NP; only those which enjoy a strong semantic connection with a “well-established kind” can be taken as kind-denoting NPs. Other nominal expressions which lack this semantic privilege are rendered object-referring NPs. Carlson (1977b) cited a pair of sentences that he attributed to Barbara Partee, which show a clear contrast between the two incorporated NPs pertaining to their semantic connection to a well-established kind. Incorporating NP *Coke bottles*, which is proved an established kind, renders a sentence, like (25.a) felicitous, while associating a non-well-established NP *green bottles* renders its minimally-contrasting counterpart (25.b) semantically odd on generic reading.

(25) a. The Coke bottle has a narrow neck.
    b.? The green bottle has a narrow neck.
The fourth diagnostic test distinguishes characterizing sentences, habitual sentences in particular, from episodic sentences. According to this test, generic sentences are typically stative, in that they express regularities; episodic sentences, however, are non-stative or dynamic since they report particular events or situations. Thus if a language exhibits a certain linguistic construction that is incompatible with stative predicates (e.g. *Mike is knowing French), this construction would typically exclude the generic reading of habitual sentences and give rise to an episodic reading instead. English, for instance, has the progressive form which seems felicitous with an episodic reading only; a reading that reports a particular event. The sentences in (26) have generic readings which their minimally contrasting counterparts (27) lack. This indicates that transforming a generic sentence into the progressive entails bleaching its original genericity.

(26) a. The Jordanian eats mansaf at a wedding feast. (generic)
b. Jordanians eat mansaf at a wedding feast. (generic)
c. A Jordanian eats mansaf at a wedding feast. (generic)

(27) a. The Jordanian is eating mansaf at a wedding feast. (episodic)
b. Jordanians are eating mansaf at a wedding feast. (episodic)
c. A Jordanian is eating mansaf at a wedding feast. (episodic)

A final test is also designed to help distinguish characterizing sentences from episodic ones. According to this test, Krifka et al. claim, generic sentences are characterized by expressing essential or normative properties; however, particular sentences express inessential and accidental properties. The classification of properties as essential or accidental draws upon the referent of the NP argument. For example, being popular is an essential property of a Hollywood star, but the same property is inessential or not normative for students or a poem, for instance. Put more precisely, a sentence is classified generic if the NP and the predicate fit in the slots of the following formula: If A then B, where A stands for the NP and B stands for the property predicated. This formula states that to be an A is to have B as a property. In set-theoretic semantics where a predicate is taken to denote a set of objects or individuals to which the property denoted by the predicate applies, this formula can be stated as if A then A is a subset of B, denoted by \((A \subseteq B)\). The propositions in (28) can only be interpreted as characterizing sentences with kind-referring NPs; while their minimally contrasting
counterparts (29) seem infelicitous on a generic reading with kind-referring NPs. The reason for this asymmetry lies in the semantics of the property predicated in each sentence, whether it is normative or accidental.

(28) a. The lion is ferocious.
   b. Lions are ferocious.
   c. A lion is ferocious.

(29) a. The lion is smart.
   b. Lions are smart.
   c. ? A lion is smart.

The assumption made here is that the property *ferocious* is essential to lions. The three propositions in (28) express a normative generalization. The generalization can be placed in an explicitly normative frame: ‘To be a lion is to be ferocious’ is true. The property *smart* is assumed inessential and accidental to lions. Therefore, the generalization expressed in (29) is descriptive, and the propositions are judged bad on a generic interpretation. (29.c), in particular, which can only be interpreted as a characterizing sentence is bad, as opposed to (28.c), which appears to be acceptable as a characterizing sentence. The only difference between this pair of sentences is the connection between the predicate and the referent of the subject argument NP, whether it is a principal connection or an accidental one. An explicitly normative frame of the generalization expressed in (29) renders the sentence false: ‘To be a lion is to be smart’ does not seem to be true.

5. Generics vs. Explicitly Quantified Sentences

Most natural languages exhibit two ways of expressing a generalization or a general statement. The first employs explicit universal quantification where an overt universal quantifier like *all* or *every* is incorporated. The other way available is using characterizing sentences in which no overt quantifier is used. An example taken from Link (1995:359) lays this out.

(30) a. All planets of the solar system revolve about the sun on an elliptic orbit.
   b. Man-made satellites revolve about the earth on an elliptic orbit.

(30.a) is naturally interpreted as each member of the closed set of solar system planets, which contains nine planets, is such that that planet revolves around the sun in an elliptical orbit. (30.b),
however, is not interpreted as quantifying over a closed set of man-made satellite objects similar to (30.a). Croft (1986) dubbed the former Closed Class Quantification (CCQ), and the latter Open Class Quantification (OCQ). A major difference between the two structures is that characterizing sentences seem inappropriate to express closed class quantification, while quantified sentences typically express closed class quantification, and may be used to express open class quantification too. According to Link (1995), this asymmetry can be accommodated if we observe the fact that universal quantification like (30.a) expresses an “actual universal truth,” while the generalization in (30.b) extends to every potential man-made satellite object that satisfies the property expressed in the predicate.

One of the most salient features of generic sentences that isolate them from their quantified counterparts is being exception-tolerant. In a generic sentence like ‘An orange contains vitamin C’, there is a room for some genetically-modified oranges which do not contain vitamin C. However, its universally quantified counterpart ‘Every orange contains vitamin C’, does not tolerate such an exception. In fact, if one orange in the set of oranges contains no vitamin C, the proposition will be false. This can be explicated if we observe that the latter puts a restriction on the number of individuals in the Restrictor which hold the property attributed in the Scope. In the above example the quantifier ‘every’ requires that for the sentence to be true, all members of the Restrictor - the set of oranges, have the property in the Scope - to contain vitamin C. The generic sentence does not seem to behave in the same fashion. More precisely, it does not put a limitation on the Restrictor in terms of quantity. This major discrepancy between the two structures manifests itself in the quantified sentence being more amenable to formal semantic analysis than the generic sentence which appears more complex and vague. In set-theoretic semantics, for instance, the quantified sentence above can be rigorously represented by the formula in (31), while its generic counterpart requires much more investigation to figure out the truth conditions which capture the sentence’s exception tolerance.

(31) Every orange contains vitamin C = 1(is true), iff \{x| x ia an orange\} ⊆ \{y| y contains vitamin C\}

Informally, for the sentence in (31) to be true, the set of all things that are oranges has to be a subset of the set of all objects which have Vitamin C.
Another pivotal characteristic of generic sentences that sets them apart from explicitly quantified sentences is that they are quantificationally vague. Carlson (1977) emphasizes that generics are not quantificational statements. They are not about how many or how much, in the same way as quantified sentences. A quantified sentence like ‘some/most/all trees lose their leaves seasonally’, can be an appropriate response to a question like ‘How many trees lose their leaves seasonally?’ However, it seems unsuitable to respond to the above questions as ‘trees lose their leaves seasonally’. Unlike quantified sentences which exhibit clear and explicit quantificational variability established by the quantifier, the quantificational variability of generic sentences appears to be vague and fluctuating. More precisely, generic sentences, Carlson argues, fluctuate in their truth conditions as the predicate varies. If there is a null quantifier associated with the generic sentence, its quantificational force varies with the meaning of the predicate. In order to determine the meaning of the assumed null quantifier in generic sentences, Carlson notes, we need to have knowledge of the particular predicate; “no other quantifier in English behaves even remotely in a similar fashion” (P. 44).

The vagueness of the purported quantificational meaning of generic sentences is multifaceted. The first facet touches on the issue of how many individuals holding of the predicated property are enough for the generic sentence to be true. Consider the examples below (taken from Katz and Zamparelli 2005).

(32)  
(a) Snakes are reptiles.  
(b) Telephone books are thick books.  
(c) Mammals give birth to live young.  
(d) Shoplifters are prosecuted in criminal courts.  
(e) Mosquitos carry the paramecium that causes yellow fever.  
(f) White sharks attack bathers.

All snakes  
Those of large modern cities  
Only adult fertilized females  
Most are not even caught  
Very few do  
Only a tiny minority

A quick look at the data in (32) clearly shows the fluctuation of the truth conditions of the sentences on par with the variability of the meanings of the predicates used. Sentences (32.a-f) are all true. However, what makes them true? (a) holds for all snakes, (b) for perhaps less than half of telephone books, (c) for most female mammals (less than half the total number of mammals), (d) for few
shoplifters, (e) for less than one percent of mosquitoes, and (f) for very few white sharks. The truth of explicitly quantified sentences, however, does not appear to vary according to a variation in the predicate meaning. It is the quantifier, which is always given a unique interpretation, that sets the semantic structure of the quantified sentence, and truth values are assigned accordingly.

If the semantic structure of generic sentences were similar to quantified sentences - in particular if generic sentences have a null quantifier with a unique interpretation - we would judge the sentences in (33) true because most of the individuals denoted by the NP in the restrictor hold the property predicated in the scope; nevertheless all the sentences are judged false.

(33) a.? Students in Yarmouk University are female. Most are female students.
b.? Seeds do not germinate. Most don’t (Katz and Zamparelli 2005)
c.? Lions do not have bushy tails. Most don’t; all Females& young lions
d.? Prime numbers are odd. An infinity minus two (Katz and Zamparelli 2005)
e.? Bees do not lay eggs. All don’t except the queen.
f.? People are over three years old. The majority are (Cohen 2006)

The intuition we get from the data in (33) is that generic sentences, unlike explicitly quantified sentences, do not constitute the truth of the generalizations they express based on the size of individuals holding the predicated property. This characteristic exemplifies the intricacy of generic sentences compared to their quantified counterparts pertaining to stating and calculating their truth conditions. A hypothesis that assumes a semantic structure of generic sentences similar to that of quantified sentences except for the overt quantifier must face the difficult task of reconciling our intuitions about sentences like (32) with the proposal that all of these sentences contain the same null quantifier. Consequently, it needs to account for the diversity of readings exhibited in (32) based on the same null quantifier proposal.

Another facet of the vagueness and complexity of generic sentences is related to the issue of determining the relevant set of individuals over which the generic sentence quantifies. This reflects the exception tolerance with which generic sentences are characterized. Any model that investigates the semantics of generic sentences has to account clearly for this distinctive characteristic and capture
the relevant set of individuals that the predicated property is attributed to, hence abstracting away from the exceptions. The examples in (34) lay this point out.

(34) a. Turtles lay eggs.  
    b. Lions have bushy tails.  
    c. Mosquitoes suck blood.  
    d. Mammals nurse their young.

Another side of the vagueness of generic sentences, which is taken as a major difference between these sentences and the quantified ones, is the so-called the Port-Royal Puzzle. This puzzle was first introduced in the Port-Royal Logic first published in 1662 (Arnauld 1964). Let us consider sentence (35).

(35) The Flemish are good painters. (Arnauld, 1964)

For the generic sentence (35) to be true there must be individuals in the restrictor who hold the property in the nuclear scope, and hence rendering the sentence true. The puzzle that a sentence like (35) presents can be verbalized as follows: ‘since the sentence in (35) is true in virtue of some Flemish individuals being good painters, the sentence in (36) should also be true since there have to be at least as many Flemish painters as there are Flemish good painters. In fact, it is evident that (36) is false, and this proves the discrepancy between generic sentences which are not monotonically increasing and explicitly quantified ones in which such a puzzling situation is not attained with any quantifier. Quantified sentences (37), unlike generic sentences, are monotonically increasing.

(36) a. Flemish are good painters.  
    b. ? Flemish are painters.  

(37) a. Most/All/Some Flemish are good painters.  
    b. Most/All/Some Flemish are painters.

According to Arnauld (1964) such a sentence is to be understood as ‘ The Flemish painters are good painters’, and that attributing the property to the whole class of Flemish people renders the sentence false.
In sum, although characterizing sentences and quantified sentences entertain the same tripartite quantificational structure, both phenomena are crucially different\(^6\). Quantified sentences appear to give a unique characterization to the quantifier that would in turn tell us the portion of individuals in the restrictor which holds the property in the scope. Generic sentences, however, do not behave in a similar fashion. The generic operator of a generic sentence, if any, could not be a null quantifier with a unique force; it has to be able to accommodate all the above vagueness and variability of truth conditions that generic sentences exhibit, (see section 6 for a discussion of the generic operator).

6. **Generics: Inductivist Approach vs. Realist Approach**

The vagueness of generics as opposed to explicitly quantified sentences, for example, has lead many semanticists and language philosophers to dig into the truth conditions according to which such problematic propositions obtain their truth values. The drive for this is the fact that in formal semantics meaning is construed from “an externalist truth-conditional view” (Pelletier 2010a: xiv). Two approaches have been devised to tackle the truth conditions of generics; namely, the inductive or quantificational approach, and the realist or rules-and-regulations approach (Carlson 1995). According to the former, which has been dominating the semantics of generics for a long while, the truth and falsity of a generic sentence depend on the truth values of a related group of instances of individuals possessing the same predicated property; from these instances speakers abstract such generalizations. Cohen (1996) argues that individuals related to the predicated property are the only individuals that count as satisfying that property or are exceptions to the generalization. The rules and regulations perspective, on the other hand, claims that each generic sentence denotes a rule out in the world that determines its truth and falsity. To quote Carlson (1995:227) “generic sentences depend for their truth or falsity upon whether or not there is a corresponding structure in the world, structures not being the episodic instances but rather the causal forces behind those instances.” Put differently,

\(^6\) Not all semanticists believe that the logical form of characterizing sentences entertains a tripartite quantificational structure, see Liebesman (2009) for a different view.
the inductive approach starts with concrete instances to draw an abstract generalization, while the 

rules and regulations approach posits an abstract rule that is realized by a generic proposition.

Though still triggering lively debate, both perspectives suffer serious shortcomings that 

prevent either of them from being able to fully account for genericity phenomenon\(^7\). Carlson 

(1995:224) observes this, and admits that “one must eventually allow for mixed or intermediate 

position.” It is noteworthy that not all characterizing sentences are equal; some of these generics are 

deeply rooted in science, and thus express a generalization based on a proved rule or a law in 

mathematics, physics, chemistry, biology, and other abstract sciences. It is wise to approach the truth 

conditions of such sentences from a rules- and-regulations perspective. Supporting evidence for this 

kind of thinking comes from the oddness that incorporating an adverb of frequency in such sentences 

creates. This oddness, and sometimes unacceptability, stems from the intuition that the truth of these 

sentences does not depend on inductively analyzing the truth values of relevant instances, but rather 

upon a rule or regulation in the world that these sentences reflect or represent. Consider the sentences 

in (38).

(38) a. Humans are mammals. 
    b. The earth revolves around the sun. 
    c. Air contains oxygen.

As can be intuitively predicted, these generic propositions do not allow any quantificational 

adverbs like *always*, *usually* or any other adverb of frequency. In fact, forcing one into (a) would 

render it completely unacceptable, while incorporating a frequency adverb in (b &c) renders them 

odd and counterintuitive, respectively. This indicates that such generics that take scientific rules and 

natural laws as a basis for their truth and falsity favor an account that adopts the realist approach, 

rather than the inductive approach. It is worth mentioning that Carlson (1995:231-232) expresses 

slightly the same idea when he discusses generic sentences established by rules in situations where

\(^7\)See Carlson (1995) for a detailed survey of problems faced by the inductivist approach, and see Cohen (1996) 

for a similar one for the rules and regulations approach.
no possibility to inductively analyze them exists. In his words “… for examples and situations such as these, the inductive approach appears in principle incapable of giving convincing account.”

However, it is hard to posit a rule or a regulation that causes a habitual sentence, which basically relates regularity based on a frequently recurring event, situation, or state of affair. This regularity does not seem to be relevant to any necessity: scientific, physical, biological, social, religious, moral, or anything that would require a rule or regulation. The rules and regulations approach seems to hold fine with generics gleaned from scientific and natural laws, but it does not appear to be able to account for habitual generics with the same effectiveness. This intuition can be drawn from the observation that the latter, established on a habitual recurrence of an event or episode, but not the former, well-established and consolidated according to laws, rules, or regulations, is cancellable.

(39) a.? A dog is four-legged; in fact, it has three.  
   b. Bill reads before bed; a habit he no longer keeps.

It is difficult to imagine what abstract causal structures out in the world - a law, a regulation, or a rule – that a sentence like (40) below reflects or is caused by.

(40) Aziz drinks tea for breakfast.

The generic sentence in (40) expresses a regularity that is built on frequently repeated episodic events over a period of time. It is almost equivalent to (41), its quantificational counterpart.

(41) Aziz usually drinks tea for breakfast.

Therefore, it seems plausible to establish the generic interpretation of this sentence, and the like, on quantificational or statistical characterization. The inductive approach is expected to do well on such habitual sentences by accounting for genericity through considering the truth values of the instances where Aziz drinks tea for breakfast and generalizing regularity from these instances.

A proponent of the rules and regulations approach might counter this quantificational interpretation claiming that Aziz has a disposition to drink tea for breakfast regardless of the number of exceptions or the true instances. However, one should seek a convincing law-like justification for
such disposition. What causal structure in the world would carry Aziz to have a disposition to drink tea for breakfast; genetic, religious, cultural, physical, biological, or even medical cause? There should be a causal structure in the world that would change a purely habitual event into a rule-reflectional one. If such a regulation or disposition-instigator is not established, it seems hard; if not implausible; to tackle the truth conditions that frame the truth values of such sentences from a rules-and-regulations perspective.

It should be emphasized, however, that the two competing views seek the same goal but from two different perspectives. In fact, these two views do not conflict as they would appear to; they both treat generics as sentences that express generalizations based on a rule or a regulation that causes them to appear, but both differ in the origin of this rule or regulation and in the characterizing sentence type each approach treats. For the realist approach, the rule or regulation is represented by a certain structure out in the world; it is there because the world is as it is, and generics reflect these rules and regulations and embodies them. For the inductivist approach, however, the emergence of rules and regulations that habitual generics reflect takes another route. Once a characterizing sentence is established inductively, the inductive processing stops from playing any role in determining the truth or falsity of that sentence, and thus discovered or existing exceptions will have no effect on the truth of such a sentence because it now represents a rule. In other words, the sentence is treated as a regularity that exemplifies a rule that has been established inductively. Therefore, the difference between the generalization expressed by habitual generics and that of rule-governed ones stems from the source of this generalization. While the latter emerges from a rule or regulation construed scientifically, naturally, culturally, or socially, the former is based on a rule construed through recurrence and frequency. This explains why generics gleaned from habitual induction, rather than those established from natural laws, are cancellable.

In sum, both approaches seem to target two separate kinds of characterizing sentences, and thus both are needed to fully account for the semantics of characterizing sentences. Therefore, it sounds plausible to suggest that the labor be divided between the two models in accounting for the
truth conditions of the two main types of characterizing sentences, the habitual and the basic or lexical, where the ‘realist’ approach tackles the latter and the inductivist deals with the former. This proposal echoes Carlson’s (1995:237) concluding remark after making a strong case against the inductivist approach that “it may well be possible that there is no single successful analysis of the domain of generics in toto, and that the domain must be split for the sake of arriving at a tractable semantic analysis.”

7. The Semantics of the Generic Operator

Most semanticists agree that characterizing sentences have a tripartite underlying structure that incorporates a covert GEN operator and its two arguments the restrictor and the nuclear scope. As mentioned above, this GEN operator exhibits no unique, fixed interpretation. More accurately, any characterizing sentence will generally have the form in (42).

\[(42) \text{GEN } x_1 \ldots x_n [\text{Restrictor}(x_1 \ldots x_n)] \exists y_1 \ldots y_n [\text{Scope}(x_1 \ldots x_n, y_1 \ldots y_n)]\]

Leslie (2008:6) articulates the mainstream view among semanticists as follows: “I know of no contemporary theorists that do not take this schema to underwrite the logical form of generics.”

Although semanticists agree on the logical form of characterizing sentences, and that characterizing sentences are associated with a covert generic operator which is unselective; since it can bind variables of different kinds especially individuals and situations; accurately specifying the semantics of GEN remains a widely controversial issue. Krifka et al. (1995) critically surveyed most of the proposals in the literature without being able to decide on the most acceptable and applicable one, though they adopted the Kratzerian modal approach (1981) as the model that “seems more promising” (p.49). Semanticists agree also that the GEN operator cannot be taken as the universal quantifier, simply because characterizing sentences tolerate exceptions. It is unsuitable also to interpret GEN as having the quantificational force of the quantifier most since generic sentences, unlike their regular quantificational counterparts, obtain a substantive property, namely their law-likeness or non-accidental nature. Krifka et al. (1995:44) give an elucidating example of the essential difference between generic sentences and quantified sentences pertaining to the nature of properties.
that each structure describes. The example they give (modified slightly to serve the point to be emphasized) is the following: Suppose that the last five surviving lions in the world are kept in a zoo, and because of an accident, four of these poor lions lost one leg. Then the quantified sentence Most lions are three-legged would be true, but its characterizing counterpart A lion is three-legged would nonetheless still be false. It has also been mentioned above that in all the examples in (33), the properties predicated hold of most, and sometimes 99% (33.e), of the subject referents, but the sentences are intuitively judged false.

Krifka et al. (1995) argue that the GEN operator, unlike nominal quantifiers, cannot be contextually restricted; it states a law-like regularity that does not observe any contextual restrictions. This indicates that characterizing sentences do not construct their generalizations on properties of actual individuals in the actual world per se. The tolerance of exception, law-likeness nature, and counterfactual support of characterizing generics have led many semanticists (e.g. Heim 1982; Krifka 1987, 1988; Krifka et al. 1995; Chierchia 1995; and Greenberg 2003) to claim that GEN is a modalized operator that quantifies over all accessible worlds, individuals, and situations. In a nutshell, the generic quantifier GEN is an unselective universal modalized operator. Krifka (1988:297) claims that

(I-generics) cannot be used to express facts which hold just coincidentally, but are law-like statements…. For example, if some nut were to clip the wings of every existing blackbird then the sentence A blackbird flies would nevertheless remain true. If one tries to develop semantic analysis in terms of possible world semantics, I-generic sentences cannot be statements with a truth value that can be checked at one index, e.g. the actual world. Instead, we have to take into account a set of indices. Thus, genericity is reconstructed as a modal notion – as some sort of necessity.

7.1 A Modal Approach to the Semantics of GEN

7.1.1 Kratzer’s Modal Theory

Kratzer’s modal approach can be used as an influential way of thinking about the semantics of GEN in a quantificational theory of generic sentences. In a series of writings, Kratzer (1977, 1978, 1985, 1987, 1988, 1990) has developed a modal theory of generics where GEN is a modalized operator that quantifies over all accessible worlds, individuals, and situations. In a nutshell, the generic quantifier GEN is an unselective universal modalized operator.
1981, 1986, 1991a, 1991b, 2012) developed a theory of modals and conditionals, which is accepted as the canonical model by a good number of formal semanticists working in this realm of semantics. A major assumption of this theory is that modals are not lexically ambiguous, but rather interpreted differently relative to a set of background information or assumptions, which she dubbed “conversational backgrounds.” This indicates that the different flavors of modals, epistemic, deontic, bouletic, circumstantial, etc., do not emerge from the ambiguous lexical contributions of the modals themselves, but rather are taken over from the contextually-dependent backgrounds of the particular contexts that wrap these modal propositions.

Kratzer (1977) made a point of departure from modal logic pertaining to the role assigned to context. In modal logic, context is assumed to determine the correct word out of the set of lexically ambiguous modal words like should, for example, by helping the addressee choose the correct ‘should’, among the many ‘shoulds’, which the speaker had in mind when uttering the sentence. In Kratzerian theory, however, context determines the set of worlds that a modal like should quantifies over, and hence explicating the flavor or sense of the modal. Therefore, context not only provides a set of indices of indexical features of meaning like the identity of speaker, addressee, time, or place at which the sentence is used, but plays a pivotal role in determining the accessibility relation function, and hence the sense or flavor of the modal. According to Kratzerian theory, the different modals vary in terms of the possible worlds with which they are compatible. The modal May, for example, is compatible with both epistemic and deontic accessibility relations, while Might is compatible with epistemic but not with deontic accessibility relation.

Kratzer (1977) claims that although the interpretation of a modal is often gleaned implicitly from context and conversational backgrounds, it can be fixed sometimes by explicit linguistic expressions like “in view of what I know, or in view of the rules of the secret committee”. These expressions work as a function f from the set of all possible worlds W to the set of all propositions that the speaker knows in w, in the case of the epistemic interpretation. This interpretation can be explicitly provided by an expression like the following: “in view of what I know” = for any w, f(w)=
the set of propositions which the speaker knows in $w$. Formally, in context $c$, *what I know* expresses a function $f$ such that:

(i) The domain of $f$ is that subset of $W$ in which the speaker of $c$ exists.
(ii) For any $w$ in the domain of $f$, $f(w) = \{ P: \text{the speaker knows } p \text{ in } w \}$

As standard in possible world semantics, a proposition is interpreted as a set of possible worlds in which that proposition is true, so for $p$ to be true in $w$, $w$ must be a member of $p$. Since $f(w)$ yields a set of propositions, i.e., a set of sets of worlds, and we only need a single set of worlds, we can use a known trick in logic to get that set. We resort to intersecting all of the propositions in the set, and eventually get one proposition, i.e., a single set of worlds in which all the propositions are true. Given this conversational background function $f$, the accessibility relation of a modal is defined as follows: for any worlds $w$ and $v$, $v$ is the set of worlds accessible from $w$ iff every proposition in $f(w)$ is true in $v$. So if we take $f(w)$ to be epistemic, $f(w)$ represents all the facts known by the speaker in $w$, and $v$ is accessible from $w$ iff all the facts known by the speaker in $w$ are true in $v$. More precisely, the set of worlds accessible from $w$ is $\cap f(w)$.

The second major idea of Kratzerian modal theory pertains to the simple dichotomy entertained in modal logic between the set of worlds accessible and the other set of worlds which is not accessible. In modal logic, the set of accessibility relation functions that determine the sense of the modal is determined by a semantic rule that applies to all modals with no change. Thus, pragmatics plays fewer roles than semantics in this model; it only provides the indexical features of meaning of a proposition through context. In Kratzerian theory, however, the tables are turned; pragmatics plays the main role in determining the set of worlds over which modals quantify. Kratzer (1981) argues that the simple dichotomy is insufficient and problematic, and suggests an ordering mechanism of the set of all possible worlds $W$ that modals quantify over. According to this assumption, the set of worlds is not simply dichotomized, but rather ranked through the interaction of two conversational backgrounds.
In Kratzer’s (1981) revised model, a sentence is interpreted according to two conversational backgrounds: the MODAL BASE \( f \) which provides a set of relevant worlds, those in \( \cap f(w) \), and background \( g \) which is an ORDERING SOURCE \( \leq g(w) \) which ranks these worlds as closer or distant from some ideal world. The specifications of both the modal base and ordering source are determined contextually. For example epistemic modal base combines with an ordering source related to information as ‘what the normal course of events is like’, some reports or beliefs. A stereotypical ordering source might be related to information as ‘in view of the normal course of events’. A circumstantial modal base combines with an ordering source related to laws, goals, plans, and wishes: what the law provides, what is good for you, what our goal is, what is moral, and what have you. Consider the illustrative example in (43) below:

(43) Given the state of your health, you should stay at home.

The sentence in (43) can be paraphrased as follows: “In view of your state of health (circumstantial modal base), and in view of what is best for your health (ordering source), you should stay at home”. In other words, the set of worlds in \( \cap f(w) \) which is top-ranked according to \( \leq g(w) \) is the set of accessible worlds in a simple modal sentence. The definition of ordering source in Kratzerian theory indicates that every comparable sequence among worlds in the relevant set of worlds reaches a point as we move towards the ‘ever-better’ worlds in which a proposition is true.

In sum, in Kratzerian approach the semantics of modal verbs incorporates three parameters. The MODAL FORCE parameter determines whether the modal verb is represented universally or existentially; i.e., whether it expresses necessity or possibility. The MODAL BASE or conversational background determines the set of accessible worlds. And the ORDERING SOURCE parameter ranks these worlds as closer to or distant from some ideal world. The specification of both the modal base and the ordering source is determined pragmatically.

7.1.2 Application of Kratzer’s Approach to the Semantics of GEN

Greenberg (2003) argues that applying Kratzer’s approach to the semantics of GEN operator has three main advantages. First, it captures the law-likeness nature of characterizing sentences
through the universal quantification over all possible worlds. Second, it accounts for the exception-tolerance of characterizing sentences since it allows the universal quantification to quantify over individuals in the most normal worlds only; this is attained through the ordering source parameter. Finally, it naturally accounts for the variety of flavors or kinds of rules (44) which characterizing sentences can express through the modal base which can vary as, *epistemic*, *deontic*, *instrumental*, *mathematical*, *linguistic*, etc., and give different interpretations.

(44) a. A cat meows. 
   b. A driver watches traffic laws. 
   c. A taxpayer pays state and federal income taxes. 
   d. A queen is the wife or widow of a king. 
   e. An even number can be divided evenly into groups of two. 

(Kratzer’s (1981) modal framework can be adopted to informally account for the truth conditions of generics in MSA in this project, and to build formal models in the future. This approach is suggested by Krifka et al. (1995:52) to be used to calculate the truth conditions of generic sentences like (45), as in (46):

(45) A lion has a bushy tail.

(46) \( \text{GEN} \ [x_1 \ldots x_i, y_1 \ldots y_i\ ] (\text{Restrictor, matrix})^{\prime} \text{ is true in } w \text{ relative to a modal base } B_w \text{ and ordering source } \leq w \text{ if } \text{ for every } x_1 \ldots x_i \text{ and every } w' 
\in B_w \text{ s.t. Restrictor } [x_1 \ldots x_i] \text{ is true in } w', \text{there is a world } w'' \text{ in } B_w \text{ s.t. } w'' 
\leq w \text{, and for every world } w'' 
\leq w \text{, } \exists y_1 \ldots y_i \text{ Matrix } [\ (x_1) \ldots (x_i), y_1 \ldots y_i\] \text{ is true in } w'' . 

(Where \( B_w \) is the modal base, and \( w'' \leq w \) means that \( w'' \) is closer to the ideal world determined by the ordering source than \( w' \))

According to the definition in (46) the sentence in (45) means that, “everything which is a lion in the worlds of the modal base is such that, in every world which is most normal according to the ordering source, it will have a bushy tail.” It is noteworthy, that the definition in (46) does not presuppose that there are lions in the actual world. In addition, it does not require that every single lion must have a bushy tail. Only lions in worlds which are ranked closer to the ideal world are counted, hence capturing exception tolerance of characterizing generics.
8. Summary

In this chapter, the intricacies of genericity as a pervasive linguistic phenomenon were surveyed. It was emphasized that sentences do not often receive either a generic or episodic interpretation automatically; the nature of the predicate incorporated along with other syntactic expressions seem to be substantive in deciding which interpretation the sentence is meant to express.

Another important point that was highlighted in this chapter is that genericity is not a uniform phenomenon. Two sub-phenomena; namely, reference to a kind and characterizing sentences, are distinguished by many semanticists. While the former establishes genericity based on a generic NP, where a property is directly or indirectly predicated to a kind or genus, the latter constructs its generalization by the collaboration of every constituent in the sentence, in particular the verbal predicate and the NP. The two genericity types were shown to interact; a case where the property predicated holds for the kind directly, rather than to any specimens of that kind.

A major distinction between the semantic behavior of explicit quantifiers and the covert GEN operator was highlighted. It was shown that unlike explicit quantifiers, the GEN operator could not be assigned a unique interpretation. This very attribute triggered a substantial amount of research endeavors attempting to account for the perplexingly fluctuating truth conditions of characterizing sentences. Most of the research in the literature adopts one of two theories; namely, the rules and regulations theory or the inductivist theory. It was made clear in this chapter that neither theory is sufficient to account for generic sentences, and instead, a composite approach that divides the labor between the two theories was proposed, echoing a similar proposal made by other semanticists.

The consensus in the literature is that the logical form of characterizing sentences is similar to that of explicitly quantified sentences. More precisely, characterizing sentences entertain a tripartite structure with an unexpressed generic operator. Since it seems futile to assign GEN a unique interpretation like overt quantifiers, and because of a number of characteristics that generic sentences share with modal sentences, GEN is construed as an unselective modalized operator that binds the free variables in the Restrictor.
CHAPTER 3: DEFINITENESS IN MSA REVISITED

1. Introduction

Definiteness, though hard to define, is established on one or more of the following concepts: familiarity, identifiability, and uniqueness. The use of the definite noun phrase the boy in (1a), for example, is licenced if its referenant is familiar, perhaps already introduced into the discourse, or identifiable, say, standing right in front of the speaker and the listener, or unique, perhaps the only boy in the context situation. Most often, however, definiteness is identified by its (possible) grammaticizations, mostly morphosyntactic markers (see Hawkins 1978; Lyons 1999). Languages which formally mark this semantic category employ linguistic ingredients, mostly morphemes, to encode in/definiteness. In English the determiner the is often used to mark definiteness and a(n) or the lack of any overt determiner typically marks indefiniteness. Consider (1) below.

(1) a. The boy is happy.
   b. A boy is happy.
   c. Boys are happy

Even without context, the subject NP in (a) is considered definite and the subject NPs in (b&c) are considered indefinite. This is due to the presence of the and a definite/indefinite articles in (a&b) and the lack of an overt determiner in (c). However, these markers are only the morphological representation of definiteness, which is a semantic phenomenon, and the two should not be treated as one thing. Sometimes this co-occurrence of identifying morphology and the semantics of definiteness does not hold. Therefore, in/definiteness should be determined semantically rather than morphologically. The concepts of familiarity, identifiability, and uniqueness, rather than the non-systematic presence/absence of morphological markers are used to do so. The sentences in (2) use morphologically definite NPs which are semantically indefinite.
The variety of grammaticizations of definiteness cross-linguistically and within the same language, the ongoing change and development of grammatical tools used in natural languages to encode this semantic category, and the inconsistent sense-form relation between the grammaticization of definiteness and the semantic concepts upon which this grammaticization is established (Abbot 2004; Carlson et al. 2006; De Mulder and Carlier 2011; Lyons 1999), have caused significant indeterminacy and confusion in identifying definities and indefinites. In languages which formally mark definiteness by morphological means, these markers must be approached with caution; in the sense that a thorough semantic analysis should be conducted to check whether these markers carry definiteness or simply co-occurs with it in many (but not all) circumstances. The only solid ground for checking these morphological markers is to investigate whether the nouns they introduce denote definite referents in the real sense; i.e. referents of these nouns are familiar, identifiable, or unique in the contexts of their use.

The interaction between definiteness and other semantic phenomena like genericity makes proper identification of in/definities all the more important (see section 3 below). An important insight from the formal semantic literature is that rather than refer to unique individuals, indefinite NPs may introduce variables into the semantic representations of the sentences in which they occur (Krifka et al. 1995). As noted in the previous chapter, genericity can be divided into two sub-phenomena, characterizing sentences and reference to a kind (Krifka et al. 1995). The former is compatible with semantically indefinite noun phrases denoting variable individuals, regardless of whether these NPs are morphologically marked for in/definiteness or not. More precisely, these NPs must be indefinite in sense, but not necessarily in form. The latter is compatible with semantically definite nominal expressions, in the sense that they denote unique kinds, regardless of their grammatical in/definiteness.
Definiteness, as a semantic phenomenon, is related to reference and the achievement of successful reference in conversation. Successful reference, when both the speaker and the hearer in a conversation are able to pick out a particular individual or kind, is established on the basis of one or more of the following semantic/pragmatic concepts: familiarity, identifiability, and uniqueness. This cluster of concepts composes the criteria for definiteness I will adopt in this dissertation. In this work, a noun phrase will be considered semantically definite if it can be felicitously used referentially. For this to happen, its referent has to exhibit at least one of the three properties in the definiteness criteria. If a noun phrase exhibits none properties of the definiteness criteria, it is considered indefinite.

This chapter investigates definiteness in MSA sentences, in preparation for an examination of the interaction of definiteness and genericity. This semantic analysis is motivated by the assumption that the definite article *al* ‘the’ in MSA no longer functions as a pure semantic definite marker, but rather developed other uses not related to definiteness. This makes an analysis of definites and indefinites based merely on its presence/absence non-satisfactory and confusing, particularly when this analysis is carried over to genericity investigation.

This chapter is organized as follows: after this introduction, definiteness as a semantic phenomenon is investigated in section 2. The shortcomings of identifying definites based solely on the grammaticization of definiteness are discussed. Definiteness criteria grounded in familiarity, identifiability, or uniqueness will be adopted as machinery for determining definites and indefinites, regardless of the problematic grammatical definiteness. A direct repercussion of this significant distinction between grammatical definiteness and semantic definiteness is discussed in section 3. In this section, it will be shown that genericity interacts with definiteness as a semantic phenomenon, not with its variable grammaticizations in different languages. Definiteness in MSA will be investigated

---

9 The phonological realization of the definite article *al*- varies according to the phonetic properties of the noun’s initial sound to which it is attached. More precisely, (a)- fully assimilates to the initial sound of the noun if it is one of the traditionally termed *huruf fîmsiya* ‘sun letters’ set, and remains unchanged otherwise. This set includes the following sounds: /tˤ/, /dˤ/, /t/, /d/, /θ/, /ð/, /s/, /sˤ/, /z/, /ðˤ/, /n/, /r/. For instance, *al*- surfaces as *a*- in *a-fams* ‘the sun’. In addition, in continuous speech, the ‘a’ of *al* is often dropped, and only “t” surfaces. These morphophonemic facts have no bearing on our semantic analysis of nouns prefixed with the definite article *al* in MSA.
in section 4. Here, the traditional grammarians’ view of definiteness in MSA will be probed and highlighted. It will be shown that traditional Arab grammarians, Modern Arab linguists, and Modern non-Arab grammarians share the insight that *al* encodes semantic definiteness in Arabic. It will be shown that they have a clear view of definiteness, and what it means for a noun phrase to be definite semantically; nevertheless, they stick to the idea that the presence/absence of *al* exclusively determines in/definites in Arabic. Section 5 presents a revision of definiteness in MSA. In this section, the canonical use of *al* as a pure definite article will be examined showing that in some cases the presence of *al* marks the noun it is attached to semantically definite according to the definiteness criteria. In addition, this section will examine environments where *al*-N does not denote a semantically definite referent in the real sense. Based on these arguments, *al* is placed in stage two of Greenberg (1978) life cycle of definite articles, in which the definite article is used in definite and indefinite environments, and is used for other grammatical purposes not relevant to definiteness. A conclusion is given in section 6.

2. **What is Definiteness?**

Before presenting a unified analysis of definiteness based on the semantic concepts underpinning this semantic category, and hence determining definites in MSA in accordance to these concepts, the shortcomings of identifying definiteness solely by its possible grammaticizations and determining definites based on grammatical markers are examined in (2.1) below.

2.1 **Grammaticization and Definiteness**

Noun phrases with a definite article are taken by many linguists (see Lyons 1999) as simple instantiations of definites, simply because these articles function as grammatical entities, which encode definiteness. In other words, it seems that the presence of these articles is often sufficient to classify noun phrases introduced by them as semantically definite. Lyons (1999) refers to such noun phrases as ‘simple definites’. Encoding definiteness, however, is not exclusively carried over by articles like *the* in English. Languages employ other tools to identify definiteness (Abbot 2004; De Mulder and Carlier 2011; Lyons 1999). Lyons (1999) dubbed these structures complex definites, in
the sense that definiteness of noun phrases is marked not by the presence of articles, but rather by the presence of other grammatical ingredients or syntactic structures. The variety of complex definites includes proper nouns, personal pronouns, noun phrases incorporating a demonstrative or possessive modifier, word order, case inflection, agreement inflection, and stress or intonation (De Mulder and Carlier 2011; Lyons 1999).

Lyons (1999: 36) emphasizes the syntactic and semantic significance of a definite determiner in rendering a noun phrase definite; he argues that “[a] definite noun phrase must normally contain a definite determiner, and the is the one that occurs in the absence of some other with more semantic content.” It is worth noting that definite articles used to mark simple definites vary cross-linguistically. In other words, languages identify simple definites in different ways: free-form articles, bound articles, phrasal clitics, or even a mixed system of free-form and bound articles where both article types are in complementary distribution. In addition, some articles are pre-posed articles and others are postposed articles. The more significant distinction, however, is between independent articles and bound or affixed articles. In this case, independent, free-form articles and bound morpheme articles may precede or follow the constituent they modify (Lyons 1999). This shows how complicated and varied the realization of simple definites is cross-linguistically.

Not all languages, however, encode simple in/definites formally, in the sense that not all languages have definite or indefinite articles. Among those languages which show an explicit distinction between definite and indefinite noun phrases, the distribution between the two classes varies. Some languages, MSA included, “make more use of apparently definite noun phrases than others” (Lyons 1999: 48). In most languages having one definite article, Lyons claims, the definite article, free or bound, is invariable, in the sense that it is not inflected, though it may undergo some allomorphic alternation. This is true of MSA, a highly rich inflectional language. Notice that a related demonstrative in MSA (3) shows agreement with the modified item in features like gender and number, as opposed to the definite article a- ‘the’, which carries no inflection of any sort.

(3) a. a-\textit{bint} ‘the-girl’, a-\textit{banaat} ‘the-girls’, a-\textit{walad} ‘the-boy’ a-\textit{ʔawlaad} ‘the-boys’
b. *haaḍi-*hi ₁-bint ‘this-f the-girl (this girl)’, *haaḍa* ₁-walad ‘this-m the-boy (this boy)’, *haaʔulaaʔ* ₁-banaat ‘these the-girls (these girls)’, *haaʔulaaʔ* ₁-ʔawllaad ‘these the-boys (these boys)’

The story about the grammaticization of definiteness is more complicated than this. A major characteristic of languages is the intricate sense-to-form correspondence system. This system gets even more complicated in mapping form to sense in the functional sub-domain of the lexicon (see Pérez-Leroux et al. 2004). Carlson et al. (2006) argues that in English, and other languages, there is a subclass of definite NPs which appear to behave syntactically and semantically different from regular definite phrases. They dubbed this class weak definites. Weak definites (4), Abbott (2004) argues, are inconsistent with either the uniqueness theory or the familiarity theory.

(4)  

a. Jill sat on the bank of the river.  
b. Bill wrote a poem on the wall.  
c. Jill went to the mall.  
d. I will watch the news when I come back home.

Though the NPs in (4) are syntactically definite, it seems unnatural to interpret them as denoting unique or familiar entities. In (a), for instance, the river has two banks, so no unique bank is referred to here. Similarly, in (b) no unique wall is picked since a room has four walls. The grammatically definite NP, the mall, in (c) does not naturally refer to a unique mall, unless the city has only one mall. Similarly, no unique news will be watched in (d). As Carlson et al. (2006) argue, weak definites and bare singulars can be semantically subsumed under one class based on their distribution and semantic interpretation.

Asymmetries in the grammaticization of definiteness represented by the sense-to-form correspondence system of the definite determiner are even more interesting. In Arabic, Albanian, and to a certain extent Romanian, an affixal article can be attached to an adjective for other syntactic purposes like agreement between the modifying adjective and the modified noun.¹⁰ In German and French, there is a type of noun phrase containing the singular numeral ‘one’, which is indefinite in sense, but definite in form. In German *der eine Mann* (the one man) normally means ‘one man’ or

¹⁰ Examples of languages employing the definite article in a fuzzy way or for other syntactic purposes are due to Lyons (1999), unless otherwise specified.
‘one of the men’. This form is especially used when contrasting with some other man or men. The same phenomenon is found in French; *L’un est parti, l’autre resté* ‘One left, the other stayed’. In Albanian also, *një ‘a’, ‘one’* is used with a definite in form noun accompanied by a possessive with the whole noun phrase being indefinite in sense: *një motra-a time (a/one sister-DEF my)* ‘a sister of mine’, ‘one of my sisters’.

In many languages, the definite article is not only used with nouns in non-definite environments, but also no longer expresses definiteness or any related semantic concepts like specificity. In some languages, a definite article is semantically empty, and functions only as a carrier of agreement features. In French, for instance, the definite article *le* can be used in both the usually unmarked generic constructions and in simple definites. This has motivated linguists to argue that the article has become a mere nominality particle accompanying almost all nouns in Modern French, by contrast with earlier stages of the language. Harris (1980) claims that demonstrative *se* is developing as a new definite article in French, and that the originally definite article *le* serves now as a bearer of inflectional categories for the noun phrase.

Asymmetries in the relation holding between the grammaticization of definiteness and the semantic concepts, which grammaticization is supposed to encode, show that it is implausible to fully equate the grammatical ingredient of definiteness with its semantic ingredient. Put differently, the presence of morphological markers of definiteness is not a necessary and sufficient condition for identifying definites in natural languages. In fact, there is a crucial distinction between grammatical definiteness represented by the grammaticization of definiteness, and semantic definiteness as an instantiation of successful reference to a particular individual or kind established on concepts like familiarity, identifiability, or uniqueness. Clearly distinguishing between grammatical definiteness and semantic definiteness is very important to this dissertation because the grammaticization of semantic definiteness is not always a straightforward process. More often than not, the definite article, the grammatical instantiation of simple definiteness, does not remain the exclusive carrier of semantic definiteness. Therefore, determining definite and indefinite noun phrases based solely on the
presence/absence of the grammaticization of these two semantic phenomena is problematic. This distinction is furnished in (2.2).

2.2 Grammatical Definiteness vs. Semantic Definiteness

Linguistic asymmetries in the use of the definite article indicate a need for a serious revision of definiteness investigation. If the article is not intended to instruct the hearer to single out the referent, or if there is no unique referent in the context domain, and in this case any individual which satisfies the descriptive content of the noun can be picked, the article cannot be taken to mark definiteness. In this case, there is a possibility that its appearance serves other grammatical functions or restrictions entertained by the grammar of the language in which such constructions exist. This and other similar observations led Lyons (1999) to propose that definiteness should not be defined completely as a semantic/pragmatic category, but rather as a grammatical category just like tense, number, or gender. This grammatical category, which he dubbed grammatical definiteness, is usually paired with a category of meaning, semantic definiteness. This explains why there is no one-to-one correspondence between definiteness as a grammatical category and the category of meaning it is based on. In addition, it explains why this grammatical category, like other grammatical categories, is not present in all world languages though the semantic/pragmatic concept of definiteness is attested in a wide range of languages. Dryer (1989) claims that articles are attested in only one third of languages, and for those which have articles, only 8% have both definite and indefinite articles. This supports Lyons’ (1999:267) observation that “[i]t is generally the case that grammatical categories are not direct expressions of the semantic/pragmatic concepts which they can be said to be the grammaticalizations of.”

Lyons (1999) argues that “definiteness is the grammaticalizations of what I have informally termed “semantic/pragmatic definiteness”” (p. 278). However, he assumes that definiteness as a grammatical category is only present in languages which encode an explicit definiteness marker, a kind of definite article. He draws upon studies which tackle the diachronic evolution of definite articles to argue for his proposed dichotomy of grammatical-semantic definiteness. In particular, he
referred to Greenberg (1978). Greenberg claims that the normal end of a definite article life cycle is to be a mere marker of nominality, noun class, gender, or case. He posits three stages of development in the definite article’s life cycle:

- A definite article expressing definiteness
- A definite article not expressing reference, and instead either expressing the combined uses of definite and indefinite articles, or accompanying the noun as a grammatical requirement in certain syntactic environments
- A definite article being a gender morpheme or a mere nominality marker

On this diachronic proposal, the definite article usually starts its life with a specific use related to a high level of individuation of the entity referred to, and when individuation of the entity starts to go low, the use of the article starts to fluctuate between definite and indefinite spaces. Finally, its presence is generalized to all noun types, and it develops into a noun marker. Partee (2006:275) agrees with the observation expressed in Lyons (1999) that “… once definiteness is grammaticized, it does not always stay tied directly to the pragmatic/semantic properties that most centrally motivated it.” Therefore, the definite article starts its life as a pure semantic/pragmatic ingredient, and might end up as a pure syntactic element.

De Mulder and Carlier (2011), also drawing on Greenberg’s (1978) diachronic presentation of the normal cycle of the definite article, note that even though this grammatical category is far from being universal, its grammaticalizational development exhibits some cross-linguistic regularities; in the sense that typically the definite article develops from a weakened demonstrative. They claim that “[w]hereas the semantic dimension is predominant in the first stage of the grammaticalization process, it can progressively fade out, which is reflected in a spread to new contexts where the articles convey neither definiteness nor specificity.” (P. 3) On Greenberg’s (1978) stage two, they argue that “[f]or … stage II, no frame of accessible knowledge is supposed to allow the identification of the referent so that the referent need not be pragmatically or semantically definite, but can be conceived as discursively new.” (P.7) However, they noted that the definite article does not inevitably go into
these three stages described in Greenberg (1987). These transitions are general tendencies, rather than ineluctable evolutions in the life cycle of the definite article.

The above argumentation about the evolution of grammatical markers from encoding definiteness exclusively to being more or less semantically bleached grammatical entities, serving other grammatical requirements, is bolstered by facts about the interaction between these grammatical entities and other grammatical phenomena in a good number of languages. Grammatical definiteness interacts with a number of grammatical processes in an interesting way. Some grammatical features are sensitive to definiteness category. An example of this is object marking which does not take place if the object is not grammatically definite in some languages. Accusative case, in particular, is restricted to grammatically definite objects in some languages, like Turkish, Persian, and Modern Hebrew; whereas grammatically indefinite objects receive either a different case in some languages, or are totally unmarked for case in others. In Punjabi, the reverse is correct, where accusative case appears only on grammatically indefinite objects, and grammatically definite objects either receive an oblique case or take a marker expressing some other function, typically a function of indirect object. Spanish, as other Romance languages, does not admit bare plural NPs in canonical argument positions (Pérez-Leroux et al. 2004). This indicates that the definite article in these languages functions, among other uses, as a licensor for nouns to occur in certain syntactic positions, a grammatical use not essentially related to definiteness.

For the purposes of studying its interaction with genericity, it seems that the distinction between grammatical definiteness and semantic definiteness is crucial as it makes the presence of the many asymmetries attested in world languages pertaining to this linguistic phenomenon less puzzling. Definiteness, as a semantic/pragmatic phenomenon is better analyzed based on its semantic concepts.
2.3 Semantic Definiteness: Successful Reference

In the literature, two approaches have been devised to capture the essence of definiteness: uniqueness, a semantic property, and familiarity, a pragmatic function (Abbott 2004; Carlson et al. 2006; Lyons 1999). Familiarity implies that both the speaker and the hearer are familiar with the referent of the definite noun; it was introduced previously, or is pragmatically familiar. Lyons (1999) capitalizes on familiarity hypothesis as a crucial distinction between the semantics of definites and indefinites. He argues that with definite NPs the speaker signals to the hearer that the referent of the NP is familiar to both interlocutors. With indefinite NPs, however, the speaker does not signal this shared familiarity of the object denoted by the NP. Familiarity, Lyons argues, could be attained through three uses of the definite NP: situational, anaphoric, and bridging cross-reference or associative use – a combination of background knowledge and anaphoric use. Consider the examples below.

(5) a. [In a car wash station, a man getting off his car told a worker there]

Please! Wash the car.

b. I saw a woman sitting alone in the park. The woman looked suspicious to me.

c. I have seen a terrible car accident on Franklin Street. The driver must have been injured badly.

The example in (a) shows the situational use of the-N. The referent of the NP ‘the car’ is familiar to both interlocutors through the physical situation, which contributes this familiarity. The hearer can easily pick out the intended referent because it is familiar from an immediate, visible situation. In (b), however, we have an example of an anaphoric use of the-N. The referent of ‘the

---

11 This section draws heavily upon Lyons’ (1999) seminal work on definiteness, but unlike Lyons’, this section is restricted to investigating definiteness in simple definites; definiteness in complex definites falls out of the scope of this section. See Lyons (1999) for a thorough investigation of both simple and complex definites.

12 Although there might be some correlations between definite in form NPs and semantically definite NPs, the morphological markers are not taken as the basis for determining definites, but rather as a trigger for checking them against the semantic concepts they are supposed to co-occur with. The solid ground for determining definites is the definiteness criteria. More precisely, grammatically definite NPs are not taken as semantically definite until their referents are shown to obtain at least one property of the definiteness criteria.
woman’ is not familiar to the hearer from a visible situation, but rather from a linguistic context. The NP denoting the referent is mentioned before. Interestingly, the referential use of *the-N* is distinct from the other uses in that successful reference here is dependent on a linguistic context, in the sense that the same NP is introduced either earlier in the sentence or in a previous sentence; the other two uses, however, provide access to their referents extralinguistically. This dichotomy, Lyons (1999) argues, is bolstered by the fact that some languages have a definite article which is used exclusively anaphorically such as Hausa and Hidatsa. Other languages like Lakhota distinguish an exclusively anaphoric article from a more general definite article.

In the example in (c), the referent of the NP ‘the driver’ is not familiar from a physical situation or a linguistic context per se. It is familiar from bridging cross-reference knowledge. This is the associative use of *the-N*. In this case, the hearer employs both the shared general knowledge and the linguistic context. The driver has not been mentioned before, but a car accident is mentioned. The hearer knows that traffic accidents involve cars, and cars on wheels require drivers. Thus, the hearer is able to make the inference that reference is to the driver at the particular accident just mentioned. The inference is attained through association with the antecedent ‘car’ and the general knowledge which triggers all things associated with a car, and among these things ‘the driver’ can be singled out as a familiar individual.

Familiarity concept is taken as a solid ground for differentiating definite and indefinite NPs, but it is not fully reliable. In (5.c), the hearer might be able to infer the referent of the definite NP ‘the driver’ from the linguistic context and commonsense knowledge, but this does not necessarily indicate that the hearer knows, or is familiar with that driver in the real sense. More interestingly, if the definite NP is replaced by its indefinite counterpart ‘a driver’ in this sentence, the familiarity status of the ‘driver’ would not be crucially different. This indicates that either there must be another semantic ingredient which legitimizes successful reference in sentences like this, or that familiarity concept itself needs to be modified to accommodate such uses of successful reference.
Similar problems with familiarity hypothesis led many linguists sympathizing with this hypothesis to argue that definiteness is about identifiability. On identifiability account, Lyons (1999) notes, the use of a definite NP, for instance, signals to the hearer to exploit linguistic and extralinguistic tools available to identify the referent of the NP. This view does not explicitly reject familiarity, but rather uses it as a medium to attain identifiability. Familiarity, when available, helps the hearer to identify the referent of the definite NP. With familiarity account, the hearer is told that he or she knows which, but with identifiability account, the hearer is directed to work out which.

Consider (6) below.

(6) [A patient struggling with walking is trying to fetch a book his wife brought him ten minutes ago as a gift for his birthday. She had put it on the TV and left. A moment after he gave up trying to reach it, a nurse stepped in; the patient asks the nurse]

Pass me the book, please!

It is impossible to account for the successful reference of ‘the book’ on familiarity account; nothing in the sentence or in the context provided indicates that the hearer is familiar with the referent of the NP ‘the book’. However, the use of the definite NP triggers to the hearer that there is one particular book the speaker is referring to, and that he/she needs to employ his/her extralinguistic tools to identify the referent, guided by the descriptive content of the noun ‘book’. The use of the verb ‘pass’ also helps the hearer to identify the referent; it indicates that the referent of the ‘book’ is within reach, somewhere in the physical situation.

Lyons (1999) notes that descriptive grammarians’ works on definiteness have been dominated by an ideas still echoing nowadays in pragmatics, with versions of identifiability and familiarity. This tradition traces back to second century AD, with the work of Apollonius Dyscolus who pairs out the presence and absence of the definite article in Greek with whether or not the referent of the noun has already been mentioned or is known in a way or another. For Partee (2006:257) “[t]hese observations thus add further evidence in support of the conclusion expressed by Farkas (2002), Lyons (1999), and others that identifiability is the concept most central to definiteness… .”
Although identifiability might be more explicative than familiarity, in the sense that familiarity is subsumed under identifiability, some cases of definiteness pose themselves as problematic to this account, and at least resist a full explanation from this perspective on definiteness. The classical example of these cases, Lyons (1999) argues, is the associative use of the definite NP. Here, it is assumed that the hearer can associate a definite NP with an entity which he or she finds in the given situation. Consider the example in (7).

(7) I have just arrived at a dentistry clinic. It is still closed, but I will wait until the dentist comes. In the above example, it is clear that neither the speaker, nor the hearer can identify the dentist. This is indicated by the use of the indefinite NP ‘a dentistry clinic’. Definite reference here is successful not because the hearer can work out the identification of the referent using linguistic and extralinguistic clues. It is even impossible for both interlocutors to know whether that dentist is a man or a woman. It is successful because the hearer, through association and general knowledge, knows that every dentistry clinic has a dentist. Nevertheless, this inference has no bearing on the identification of that dentist assumed to be in charge of that clinic.

Lyons (1999) identifies other areas of definiteness where identifiability concept does need seem to be a fully satisfying explanation. Definiteness using relative clauses (8), cases where the referent is hypothetical, potential, or in the future (9-11), resist identifiability account.

(8) Ali is in the other room playing with the toy I bought him yesterday.
(9) The woman who marries me will be lucky.
(10) The winner of this lottery will be rich.
(11) The first women to reach Mars will be famous.

These examples demonstrate the idea of uniqueness, in the sense that there is a unique entity salient in the context domain that the definite NP denotes. The definite NP signals to the hearer the idea that there is one entity which satisfies the descriptive content of the NP used. In (7), for example, the dentist is not identified pragmatically, but it is singled out context-dependently as the unique dentist salient in the context, and the hearer can even talk about this dentist referentially, ‘I will leave; he/she is too late’.

56
In order to account for mass and plural definites, Lyons (1999) argues that uniqueness criteria must be supplemented with another concept for which Hawkins (1978) is credited for introducing, **inclusiveness**. On inclusiveness account, reference is not made to a unique entity relative to a particular context as the case with singular count nouns, but rather “… to the totality of the objects or mass in the context which satisfies the description” (Lyons 1999: 11).

The above semantic analysis of definiteness gives an impression that definiteness is not a unified phenomenon. There is “definiteness” established on familiarity of the referent, “definiteness” based on identifiability of the referent, and “definiteness” phenomena based on uniqueness or inclusiveness. Regardless of whether definiteness is a unified phenomenon or not, this dissertation will stick to the definiteness criteria established upon concepts like familiarity, identifiability, and uniqueness in determining *al*-Ns in MSA as semantically definite or indefinite. The impulse for treating definiteness as a unified phenomenon, however, comes from its gramaticization. The same or simpler clusters of definiteness criteria are attested in languages as licensors of the so-called definite morphology. Definiteness criteria underscore the role that definiteness and definite morphology play in establishing successful reference. For successful reference to be attained, the referent of a noun phrase has to exhibit at least one characteristic property which enables interlocutors to single it out from all other alternative individuals satisfying the descriptive content of the nominal expression. These properties are what semanticists identified as concepts of definiteness: familiarity, identifiability, uniqueness, and inclusiveness. These properties are essential pathways for successful reference, and they are not heterogeneous or conflicting. They all lead to successfully identifying and referring to an unequivocally particular individual or kind in the outside world without running the risk of confusing it with another individual to which the descriptive content of the linguistic expression applies. This is the essence of semantic definiteness; it is a conceptual designation of a referent of some linguistic expression as being particular and identifiable in a relevant context, and hence can be singled out by both interlocutors. Therefore, one way to explain what can appear to be
inconsistent use of definite morphology, while maintaining the view that definiteness is a semantic phenomenon, is recognizing steps in the grammaticization process.

Another interesting question is about the relation between definiteness as a semantic phenomenon and its inconsistent grammaticizations in different languages. More precisely, since the grammaticization of definiteness is shown to be inconsistent and problematic, should it be ignored altogether from definiteness investigation? The answer to this question is grounded in another relevant question: Why does grammaticization of semantic categories like definiteness occur in natural languages? Definiteness is a semantic concept in the sense that it stands for something in the outside world, not within the sentence. Speakers only see the grammaticization of a semantic category like definiteness, not the category itself. For speakers to be able to indicate to hearers that they are referring to definite entities in the outside world, rather than to indefinite entities, they need linguistic tools to articulate this. These tools vary cross-linguistically, from explicit definite articles to implicit markings like word order. Languages use these grammatical ingredients to carry this semantic notion of definiteness. However, these grammatical tools are not necessary and sufficient condition to determine definites in natural languages. As shown above, more often than not, these markers start their life as pure definite markers, and because of the ongoing grammaticalization process, they end their life serving other grammatical functions, not necessarily related to definiteness.

Grammaticalization is an ongoing process; it in principle never stops from developing new functions of a certain grammatical category, eliminating some functions or meanings of a certain grammatical tool, or creating other tools to carry meanings previously carried by other grammatical ingredients. However, definiteness as a semantic phenomenon is not susceptible to change, but its linguistic representation is. That said, a semantic analysis based on concepts underlying definiteness like familiarity, identifiability, uniqueness, and inclusiveness, is necessary to accurately determine definites in a language. This does not mean ignoring grammatical markings of definiteness in a language, but rather approaching them with caution, as they might not be faithful representatives of definiteness. Put differently, a definite in form noun is determined definite in the semantic sense if it
successfully refers to a particular individual or kind satisfying at least one property of the definiteness criteria, and indefinite otherwise. In a nutshell, investigating definiteness in languages which formally mark this semantic category requires checking whether these markers are still functioning as pure definite markers by devising a semantic analysis based on the definiteness criteria.

It seems, therefore, that the essence of definiteness is successful reference to a particular individual or kind, and that successful reference is attained based on two grounds, pragmatic and semantic, which a referent of NP exhibits. The former is explicated by familiarity or identifiability and the latter by uniqueness or inclusiveness.

Before I conclude this section, I would like to make a clear distinction between definiteness and specificity since these two concepts overlap quite often. In addition, articles encoding specificity exclusively, rather than definiteness, are widespread (see Lyons 1999). This distinction is important to this thesis because genericity is sensitive to definiteness, not to specificity. Therefore, it is useful to set them apart. Specificity is expressed when the speaker has a specific entity which satisfies the descriptive content of a noun phrase. Definite and indefinite NPs can be specific and non-specific, though definites tend to be specific, and indefinites tend to be non-specific. Consider the examples below.

(12) a. Did you see the book I bought from the book fair? I left it on this shelf.
   b. The president of the United States runs up to two terms.
   c. I bought a car this morning.
   d. I want to buy a car.

In (a) both NPs ‘the book’ and ‘the book fair’ are definite and specific. The speaker refers to specific book and book fair. In (b) however, the NP is definite, but non-specific; reference is not made to a specific USA president, or to the current president. Reference is rather made to any contextually unique individual which satisfies the descriptive content of the noun. In (c) the NP is indefinite, but

\[ \quad \]

\[ ^{13} \text{I believe that uniqueness is more comprehensive than inclusiveness in that even mass and plural count nouns can be captured under uniqueness. The maximality of a mass noun or a plural count noun yields in speaking of one whole unique unit, which consists of integral parts or members. This wholeness can be referred to as one particular entity salient in the context domain. That said, I will use uniqueness as a unified concept for both uniqueness and inclusiveness.} \]
specific. The speaker is not referring to any car; he or she has a particular car in mind that he or she is referring to, and yet the NP ‘a car’ is semantically indefinite. Example (d) supports the canonical indefinite NP as being non-specific. Therefore, as Lyons (1999) argues, the difference in meaning between a definite NP and an indefinite NP cannot be simply that the former denotes a specific individual and the latter does not.

However, it is worth mentioning that although in (12c) ‘a car’ is particular and specific to the speaker, it is not necessarily so to the hearer. However, using a definite NP ‘I bought the car this morning’ changes the semantics of the sentence. The sentence with a definite NP, Lyons argues, clearly presupposes familiarity and specificity of the referent for both interlocutors. Therefore, the sentence with the indefinite NP can be felicitously used in a situation where the car in question does not fall within the hearer’s experience, hence encoding novelty of knowledge on the hearer’s part. However, the other sentence, with a definite NP, would be infelicitous in such situation, and would be appropriately uttered in a situation where the hearer is familiar with the referent of ‘the car’ in question in a way or another. This supports Lyons’ (1999) claim that an indefinite singular noun phrase might be used to speak of an arbitrary object of the class described by the noun phrase, or it can be used to denote a specific entity. In the former, the indefinite is said to be non-specific, but the latter is specific. Unlike definites, indefinites might be specific but not referential. This is because a specific indefinite is only identifiable by the speaker but not the hearer. The hearer cannot pick out the referent of a specific indefinite in a sentence like ‘I bought a bike’. Therefore, successful reference to a particular individual or kind is the crucial line that separates definites from indefiniteness, not specificity of the referent.

Definiteness as an instantiation of successful reference to a particular individual or kind on speaker-hearer’s part is what is crucial when dealing with the interaction between definiteness and other semantic phenomena like genericity. This interaction is examined below.
3 Definiteness and Genericity

Although the semantic distinction of noun phrases into generic and non-generic is independent of the distinction of definite and indefinite noun phrases, these two categories interact, and sometimes overlap, in world languages. There is a strong connection between semantic definiteness and genericity. This intrinsic interaction between the two phenomena manifests itself in a slightly shared ground of concepts. This interaction supports the claim that grammatical markings of definiteness in natural languages are not necessary and sufficient condition to determine definites. The reference to a kind NP, for instance, is necessarily definite as it denotes a unique species or kind. However, reference to a kind can be manifested by different grammatically in/definite nominal expressions. Uniqueness is a concept related to semantic definiteness; hence need to be explored semantically, regardless of the grammatical definiteness status of the kind-referring NPs. English, for instance, admits singular definite noun phrases, bare plural noun phrases, and bare mass nouns to express reference to a kind genericity.

(13) a. The dinosaur is extinct.
    b. Dinosaurs are extinct.
    c. Gold is a precious metal.

If indefinite in form bare nouns are taken as indefinite in sense, they cannot be used as kind-referring expressions simply because kinds are semantically definite by default due to their uniqueness. By using bare nouns to denote kinds, English speakers do not acknowledge that reference to kind generics are semantically indefinite; i.e., denoting variable kind individuals, but rather that the grammaticization of definiteness in English is not a straightforward process. Indefinite in form NPs can denote definite in sense referents. Equating definiteness as a semantic phenomenon with its variable, changing grammaticalizational manifestations runs into challenges like these, which are abundant in natural languages where grammatically marked definite/indefinite nouns are found in indefinite/definite environments, respectively (Lyons 1999).

However, using the N seems better than using Ns, which incline to express the totality of members of a class or species, as kind-referring nominal expressions. In English, Lyons (1999) notes
that unlike grammatically indefinite noun phrases, grammatically definite noun phrases denoting a class or kind admit exceptions more readily than bare NPs denoting a class or species in the same environments. He also notes that the definite plural in English also accepts Kind-level predicates more easily than does BP. These observations, however, do not contradict the fact that all these heterogeneous NPs in terms of grammatical in/definiteness are semantically definite, and can be used as kind-denoting nominal expressions. Consider the examples in (14&15), taken from Lyons (1999:184):

(14) a. **The Italian** drinks rather a lot, though I must say Luigi is very abstemious.
    b. *An Italian drinks rather a lot, though I must say Luigi is very abstemious.
    c. **The Italians** drink rather a lot, though I must say Luigi is very abstemious.
    d. ?Italians drink rather a lot, though I must say Luigi is very abstemious.

(15) a. **The Brazilians** are twice as numerous as thirty years ago.
    b. ?Brazilians are twice as numerous as thirty years ago.

The realization of generic NPs varies cross-linguistically in terms of grammatical in/definiteness. In some languages, generic NPs are typically definite in form, and in others, they are indefinite. In Spanish, for instance, BP can occur, but with a non-generic reading only. French, however, does not have non-singular indefinite generic NPs (Lyons 1999). Although generic NPs are encoded in languages with grammatically definites, and in others with indefinites, there are languages which use both definite and indefinite NPs to express genericity. In such case, it is not clear if these heterogeneous noun phrases in terms of grammatical definiteness, which are capable of expressing genericity, form a unified class or not, and whether they might get their generic value in different ways. This becomes clear if we notice that alternatives are not interchangeable, and that these NPs can occur in non-generic sentences (Lyons 1999).

Borik and Espinal (2012:124) argue that “[a]ll languages that have Determiners … have **definite kinds**, a possibility which does not prevent languages from using other means to refer to kinds (e.g. bare plurals in E).” Krifka et al. (1995) treat the definite singular generic as the central generic nominals, because unlike BP and other nominal types, it is compatible with well-established kinds only. The assumption is that only definite singular NPs can be generic by themselves, not
deriving their generic values, as BPs are, from the context or the predicate associated, where the noun phrase itself might not be generic. This claim is bolstered by Lyons’ (1999) observation that unlike grammatically definite and indefinite non-singular generic NPs, which are rarely attested to be both permitted in one language, definite and indefinite singular generic NPs can co-occur in the same language. These asymmetries in the realizations of generic NPs in world languages indicate that reference to kind generics do not observe the grammatical definiteness status of NPs employed to denote kinds, but rather their semantic definiteness in that they successfully refer to a particular kind or species. More precisely, NPs can be used as generic nominal expressions as long as they are semantically definite, and fail to do so otherwise.

Interestingly, generic NPs behave semantically like definites to a certain degree. The singular generic NP, for example, denotes a unique kind or species; it patterns with unique definites like the sun, the moon on the individual level. Lyons (1999) claims that generic noun phrases, definite or indefinite in form, behave like definites in that they meet both criteria for definite reference - uniqueness and identifiability. This does not mean, he asserts, that languages must realize them grammatically definite. This claim holds to the distinction he proposes between grammatical definiteness and semantic definiteness, and is adopted in this study. Generic noun phrases are a good example of this distinction, in the sense that though they may appear grammatically indefinite in certain languages like English, they behave, in certain ways, as definite expressions because they are semantically definite.

Hawkins (1978) notes that in/definite articles – when both function as pure, semantically true in/definite markers - can be used in generic and non-generic reference. He argues that in both cases the nature of the referential function of each is the same. The indefinite article still refers exclusively, and the definite article refers inclusively within a scope of pragmatically defined contextual parameters. A singular non-generic indefinite might denote a specific individual known to the speaker but not to the hearer, or a non-specific entity not known to both interlocutors. In its generic use,

14 Reference here means denotation.
however, reference is made to an individual as opposed to the species, but in a random fashion. This randomness of the choice of the individual makes it possible to any entity in the species meeting the descriptive content of the noun to be a representative of the species itself. “And the fact that the particular referent chosen may be any member of the total class is responsible for the class idea inherent in singular indefinite generics” (P.214). Whenever an indefinite singular NP is used, the context tells which is more salient an individual or a whole species interpretation. As for the definite noun, inclusiveness or uniqueness is subsumed under both uses, generic and non-generic. Instead of instructing the hearer to identify a unique referent salient in the context domain, the definite article in its generic use instructs the hearer to identify a unique species or kind. “Thus, the universal idea is present in all uses of the, both generic and non-generic. (P. 217).

The distinction between definiteness as a grammatical category, which represents some semantic/pragmatic categories of meaning, and semantic definiteness explicates the variable encodings of generics using grammatically definite and indefinite NPs. Generics which are related to identifiability or uniqueness can be realized with grammatically definite NPs in languages in which these two concepts are represented grammatically definite, or with indefinite in form NPs in languages which do not represent them as grammatically definite. Therefore, as Lyons (1999:2870) puts it, “[t]he effect of these possibilities is that … there will be considerable variation between languages in the use of the category. Thus some languages will require generics to be definite while others do not; in some languages definiteness will be optional even in noun phrases clearly interpreted as identifiable….”

It seems that semantic definiteness is closely related to genericity, though the grammaticization of this meaning category might cause asymmetries in the realization of genericity in world languages; a fact that should have bearing on our analysis of definiteness and its interaction with genericity in the NP system in MSA, (see chapter 4). However, despite the inconsistencies in use owed to its lifecycle, morphology is a clear place to begin investigating definiteness. Therefore, it is useful to present the received, traditional grammarians’ view of definiteness in MSA and its
problematic shortcomings before moving ahead and proposing another view which takes semantic definiteness, the definiteness criteria in particular, as a basis for determining definites in MSA.

4. The Traditional Grammarians’ View of Definiteness in Arabic

As many world languages, MSA entertains both simple definites and complex definites. Definites in MSA include\(^{15}\) (Al-saamirraʔii 2011; Al-ɣalaaliinii 1993; Hatoum 2007):

- Pronouns like ʔanaa (I), nahnu (we), ʔanta (you-m)
- Relative words like allaḍi (who/which.m), allati (who/which.f), man (who)
- Demonstratives like haaḍa (this.m), haaḍi-hi (this-f), haaʔulaʔ (these)
- A noun introduced with al (the)
- Proper nouns
- A noun in a definite construct state kitaabu l-walaldi (book the-boy, the boy’s book)
- A noun preceded by the vocative particle yaa (O), yaa walad tạatal (O, you boy come over)

Simple definiteness in Arabic seems to be a straightforward phenomenon, where nouns prefixed with the definite marker al ‘the’ are considered definite (16a), and those lacking this marker are indefinite (16b). It is worth noting that the definite article al is invariably used with all noun types: animate/inanimate, singular/plural, count/mass, and feminine/masculine.

\[
\begin{align*}
(16) \quad \text{a. al-bint} & \quad \text{‘the-girl} \\ 
\text{al-walad} & \quad \text{‘the-boy} \\ 
\text{al-banaat} & \quad \text{‘the-girls} \\ 
\text{al-qalam} & \quad \text{‘the-pen} \\ 
\text{al-ʔaqlaam} & \quad \text{‘the pens} \\ 
\text{al-maaʔ} & \quad \text{‘the-water}
\end{align*}
\]

Traditional Arab grammarians (e.g. Al-haramii 2005, d. 1302; Al-wardii 2008, d.1348; Sibawayh 1988, d. 796)), many modern Arab grammarians (e.g. Al-saamirraʔii 2011; Al-ɣalaaliinii 1993; Hatoum 2007), and many non-Arab scholars (e.g. Buckley 2004; Haywood and Nahmad 1962; Wright 1898) studied definiteness in Arabic extensively. They all share two important insights: (1)

\(^{15}\) This section restricts itself to simple definites, nouns introduced with the definite article al ‘the’ and its allomorphic alternants, in MSA; other complex definites fall out of the scope of this section.
definiteness in Arabic is realized by one of the seven categories mentioned above\textsuperscript{16}, and (2) \textit{al} is a definite article that renders a noun it is added to definite. Many modern grammarians have inherited key ideas about the distinction and determination of definites from traditional grammarians. Brustard (2000:18) summarized the perceived dichotomy of nouns in Arabic as definite and indefinite based on the morphology only. In her words “[p]erspective and descriptive grammars alike describe the system of definiteness and indefiniteness as dichotomous: nouns are either definite or indefinite … [n]ouns can be made definite with the addition of the definite article ….” Al-wardii (2008) investigates the grammatical and semantic function of the definite article \textit{al} in Arabic. He states that a noun in Arabic could be either definite or indefinite. An indefinite noun, which is the unmarked noun in Arabic, is the noun which can be made definite by adding \textit{al} to it. Sibawaih (1988), the great Arab grammarian, describes the role of the definite article \textit{al} in semantically differentiating between nouns having it and their minimally contrasting bare counterparts in terms of its bearing on definiteness. For him, \textit{al} renders a noun definite, in the sense that it signals to the hearer that the referent of the noun is familiar to him/her, and that he/she is supposed to pick out a particular individual out of all individuals in the class to which the descriptive content of the noun applies. That referent, he asserts, is marked known to the hearer, and the hearer is assumed to recall the referent intended by the speaker, in a way or another. In his words (Vol.1, p.5)


(With regard to a and l [\textit{al}] such as in the man, the mare, the camel, and the like, it [the noun] became definite because you [the speaker] intend [to refer to] a certain entity out of all members of a genus [class]; this is because if you say: I passed by a man, then you claim that you passed by an individual to which this noun applies, not referring to a particular individual known to the addressee. And if you insert \textit{al} [to the noun], you remind him [the hearer] of a man he already knows; thus you say: the man who has such and such attributes; so that he [the hearer] recalls the one [the man] he [the hearer] knows such and such attributes about.)[Clarification added].

\textsuperscript{16} There is a disagreement among Arab grammarians concerning definites in Arabic whether they are seven or five, but all agree that nouns introduced with \textit{al} are definite.
Al-harami (2005), another prominent traditional Arab grammarian, explains other uses of *al* as a grammatical and semantic definite marker. He talks about two uses of *al*: anaphoric and genus referring. In the former, the reference of a noun prefixed with *al* is successful because that noun has been mentioned previously, and so becomes identifiable from the linguistic context. In the latter use, *al* is used with nouns uniquely referring to a whole genus, rather than to any members of a genus.

This is roughly equivalent to reference to a kind genericity introduced in late twentieth century in Krifka (1988). In discussing definite nouns introduced with *al*, Al-harami (2005:245) says that such nouns

\[\text{nahwa: ar-radżul wa-l-yulāam, fa-haadīhi 'l-laam' takuun li-l-Sahd, wa-ma'snaa l-Sahd: } \text{'an ya-kuun qad taqaddam dikr ra'džul Sāhīdah wa-Sārafa, fay'aqul lak muxe'ub: ra'ayta dāalika r-ra'džul, yasni: al-la'dīhi Sārafa, wa-yuwa'di'ihhu lak qawlhu tu'zala: " kamaa 'arsalnaa 'ilaa fir'aawnar rasuulaa. fa'sā'a fir'aawnu r-rasuul, " 'alaa taara'an 'annaa 'rasuulaa' al-'awwal dāla'ā nakira, falamma ra'dīdā dikruhu bil-'alīf' wa-l-laam, Sālimnaa 'annahu dāalika r-rasuul al-ma'shuud la'dīti taqaddama dikruhu nakira ... wa-qa'd takuunu haadīhi 'l-laam' lil-djīns, ka-qawlihi tu'zala: " 'īnna l-'insaan la-fīi xusr" lam yurid 'īnnaan ba'saynihi, wa-'īnamaa 'araada biqawlihi: 'al-īnnaan dāli'i bani'ī 'aadam, fa-dalla 'alaa 'annaa 'al-'alīf wa-l-yu'laam' takuunaan li-dāli'i b-l-djīns.\]

(as: the man, and the boy, where this ‘I’ marks familiarity, and familiarity means that the hearer knows and is familiar with the man who has been previously mentioned; in this case the speaker says to you (the hearer): Did you see that man, which means the one you knew, and this use is explicable in Quran (73: 15&16) “As We sent to Pharaoh a messenger. But Pharaoh disobeyed the messenger,” don’t you see that the first use of ‘messenger’ is indefinite, and when re-used with *al*, we knew that that was the familiar messenger previously mentioned as indefinite … and this ‘I’ could be used with a genus, as in Quran (103:2) “Indeed, the human is in loss,” where no specific human is referred to, but He meant when saying “the human” all mankind, so this shows that ‘al’ is used to signal a genus as a whole.)

It is clear from these two quotations, which represent the mainstream view of the definite article *al* in Arabic traditional grammar, that traditional Arab grammarians had a clear understanding of what definiteness is, and what it semantically means for a noun to be definite, in particular. This indicates that for traditional Arab grammarians, a noun prefixed with *al* is a definite in the sense that it is either identifiable in a way or another, or unique. The definite article *al* contributes this semantic identifiability/uniqueness to nouns to which it is attached. In other words, a noun prefixed with *al* is semantically definite, and a noun lacking it is semantically indefinite. On this perspective, *al*-N can only be semantically definite, and cannot be used in indefinite environments.
Most modern Arab grammarians agree with traditional ones with respect to the semantic function of the definite article *al*, and its sole role in rendering a noun to which it attaches semantically definite. Al-γalaaliinii (1993: 147) explicitly states that adding *al* to a bare noun in Arabic renders that noun definite; in his words, “*al*-muqtarin bi-*al*: ʔismun sabaqathu (*al*) faʔafaadathu t-taʕriif, fas'aara maʕrifatan baʔda ʔan kaana nakira. Ka-r-radγul wa-l kitaab wa-l-faras.” (The [definite noun] prefixed with *al* [is] a noun preceded by (*al*) which contributes definiteness; thus, [it] becomes definite though it was indefinite [before adding *al* to it], such as: the man, the book, and the mare). However, he talks about uses of *al* as not a definite marker as with some proper nouns like *al*-haariθ ‘the-Harith (personal noun), *al*-ʕiraaq ‘Iraq’, and with relative and adverbial words as *al*-laði ‘who/which.m’, and *al*-lati ‘who/which.f’. However, he mentions that the received interpretation among Arab grammarians is that even in these cases definiteness is expressed, and that *al* contributes to the semantic definiteness of these words.

After listing the seven definite categories in Arabic, Hatoum (2007) argues that four of these definites can be made indefinite. Most importantly to our study, he mentions the definite noun prefixed with *al* as one of these four. He explains that among these four definite categories which can be rendered grammatically and semantically indefinite is “*al*-isma l-maʕrif al-mutakawwin bi ((*al*)): naqul: ((*al*-kitaab)), faya-kwuun haada l-isma maʕrif; nuʤarriduhu min ((*al*)), biqawlinaa: ((kitaab)), fayunakkar, ʔay: yusˤbih nakira” (the definite noun with ((*al*)) attached to it: if we say ((the-book)), the noun is definite, and if we remove the al as in ((a book)), the noun becomes indefinite.

Al-saamirraʔii (2011:100)¹⁷ is exceptional in this regard. He argues that the definite article can be used with nouns denoting not a definite referent, but rather any referent that meets the descriptive content of the noun, in the sense that the descriptive content, not the referent, is known to the hearer. He argues that this use of *al* is intended to “*al*-ʔifaara ʔilaa waahid mimma ʕurifat

¹⁷ No wonder that Al-saamirraʔii digs that deep in the semantics of *al*, as his seminal book is entitled ‘the meanings of syntax’. He thoroughly investigates the semantic basis and ramifications of most of the syntactic phenomena in Arabic.
Referring to an individual whose characteristic attributes are known in mind, without intending to specify a certain individual such as in saying (go to the market, and buy us such and such) to a person who had not entered the city before, and had never seen its market. In this case, you do not refer to a particular market). This is an insightful point raised by Al-saamiraa because it explicitly indicates that al, the definite article, can be used with nouns in indefinite environments. However, the way he interprets this use does not reflect that he refers to this developed use of al in Arabic. For him, this use of al is another instantiation of using al with nouns denoting a genus. It is not used to denote a particular member of a class, in the sense that al in this case does not contribute familiarity, but rather used with a noun with an indefinite referent, but still contributes definiteness in terms of uniqueness; the noun refers to one unique genus. He compares two contrasting perspectives on the interpretation of this use of al among Arab grammarians, though neither view disputes the claim that al here contributes definiteness, but rather disagree about what type of definiteness it expresses, uniqueness or familiarity. He concludes (P. 109), agreeing with their interpretations, that both views are correct, and that the disagreement is rhetorical. In his words:


(… this al, in fact, indicates a genus [reference] since its incorporation is not intended [to denote] a particular entity, but rather an entity of the familiar genus, as the genus is familiar and known, and the [noun] to which al is attached is [refers to] a non-specific individual of this genus… so it [al] is not to express familiarity which we explained above, but indicates that the whole genus is familiar. Though some of them (Arab grammarians) see that al in all its uses is for expressing familiarity, and familiarity is divided into two kinds: a familiar individual and a familiar genus… and their argument for this [interpretation] is that genuses are familiar in the minds [of interlocutors], [and] known to the addressees, [and genuses] are distinguishable. And it seems to me that the controversy between the two camps is a difference in wording).
It appears that although Al-saamirra?ii (2001) identifies an interesting use of *al* in an indefinite environment where the noun to which it is prefixed naturally denotes an indefinite referent. The seminal, heavy traditional literature blurred his characterization of this use of *al*, and caused him to refrain from proposing something that would disagree with the received canonical status of *al* as a definite article which contributes either a familiar or unique individual or genus. In his long and rather detailed discussion of this interesting use of *al*, he never uses the term *nakira* ‘indefinite’ to characterize the referent of this type of NPs. However, such use of *al* will be capitalized, examined, and interpreted differently in my analysis of the definite article *al* in section (3) below.

Non-Arab scholars who wrote descriptive grammar books of Arabic do not seem to propose a different view from that of traditional Arab grammarians. Wright (1898) and Haywood & Nahmad (1962) state that adding the definite article *al* to a noun renders it definite. Buckley (2004:19) states that “[a] word with the definite article becomes defined.” He identifies many uses of the definite article *al* without specifying whether in all these uses the definite article contributes semantic definiteness, uniqueness and familiarity, or not. This indirectly indicates that he adopts the received traditional view of the status of *al* in Arabic. He summarizes (pp. 19-31) the variety of uses of the definite article *al* as follows:

- a. The definite article is used with a noun which becomes definite because it has been previously mentioned.
- b. The definite article is used with a noun that has not been previously mentioned, but assumed familiar to the hearer through shared knowledge.
- c. The definite article is used with nouns with unique referents, like *the sun, seasons, weekdays, meals*, etc.
- d. The definite article appears with some words indicating place, often used as adverbials like *min l-?asfal* (from the-below ‘from below’); *min d-daaxil* (from the-inside ‘from inside’); *min l-xalf* (from the-back ‘from the back’).
- e. The definite article is used with nouns denoting a species, class, or kind.
- f. The definite article is used with nouns governed by the preposition *min* ‘from’.
- g. Abstract nouns are often used with the definite article.
h. The definite article often occurs in verbal nouns.

i. The definite article is used with adjectives and active and passive participles when functioning as nouns.

j. The definite article is used with adjectives and participles governed by the preposition *min*.

k. The definite article must be used with adjectives modifying definite nouns.

l. The definite article is used with adjectives and participles governed by the preposition *min*.

m. The definite article is used with adjectives modifying definite nouns.

n. The definite article is randomly, but frequently, used with place names.

o. The definite article is used with titles followed by names referring to persons.

p. The definite article is used with cardinal and ordinal numbers.

q. The definite article is used with some words referring to time; in this case it functions as the demonstrative ‘this’.

r. The definite article is used instead of a possessive pronoun suffix with names of parts of the body, and family relations and companions.

Astute readers can identify some uses of *al* in which it contributes semantic definiteness, and others in which it does not. Some of these cases will be revised, investigated, and given a more plausible interpretation in the following section.

5. Definiteness in MSA Revisited

Taking into consideration that definiteness is a semantic category established on concepts of familiarity, identifiability, or uniqueness of the referent; this set of criteria will be used to identify simple definites in MSA. Put differently, if a noun introduced with *al* is shown to denote an individual satisfying one or more of the definiteness criteria, this noun will be considered definite. However, a noun introduced with *al*, denoting an individual which fails to express either familiarity, identifiability, or uniqueness, will be considered semantically not definite. In this case, these asymmetrical uses of *al* where nouns introduced with it are not shown to satisfy the definiteness criteria, and still the presence of *al* is acceptable, or even necessary, will be investigated to show the developed grammatical functions of the definite article in MSA, and its grammatical status on Greenberg’s (1978) cycle of definite articles.

This section is divided into two major sub-sections. The original use of the definite article *al* as a real definite article introducing a noun whose referent is familiar, identifiable, or unique is
discussed in (5.1). Other uses of al in indefinite environments and for grammatical functions not related to definiteness are investigated in (5.2).

5.1 The Canonical Use of the Definite Article al in MSA

The definite article al, at least on its canonical use, contributes definiteness to the noun it is attached to; in the sense that the referent of that noun is either familiar, identifiable, or unique - referring to a context-dependent unique individual, or to a unique genus. In the following sections, the interaction between the definite article al and each of these meaning ingredients is examined.

5.1.1 Familiarity

The definite article al marks familiarity of the referent of the noun prefixed with it by signaling to the hearer that he/she is familiar with the denotation of the noun through three uses of al:

A. Anaphoric Familiarity:

On this use of al, the noun prefixed with al has been previously mentioned, hence becomes familiar to the hearer through the linguistic context. Consider the examples in (17):

(17) a. [Two men are talking about good deeds each of them has done today. One of them says:]

\[
\text{marar-tu bi-radʒul-i-n yatljub-u musaaSadah, fa-saaSad-tu r-radʒul} \\
\text{passed-1 by-a man-Gen-N 3.ask-Nom help, so-helped-1 the-man}
\]

(I passed by a man asking for help, so I helped the man)

b. [A student talking to her friend on the phone]

\[
\text{ʔams, iltaqay-tu t\'aalib-a-n djadiid-a-n. at\'t\'aalib-u} \\
\text{yesterday, met-1 a student-Acc-N new-Acc-N. The-student-Nom}
\]

\[
yaskun-u bidjaanib-i baytii \\
\text{3-live-Nom next to-Gen house-my}
\]

(Yesterday, I met a new student. The student lives next to my house)

The reference in (17.a) is successful because the speaker signals to the hearer that the referent of r-radʒul ‘the-man’ is the same as that of the radʒul ‘a man’, mentioned earlier in the sentence. This is attained through the linguistic context. The first mention of radʒul ‘a man’ is indefinite, which indicates that novel information is delivered, and that the hearer is not assumed to be familiar with the referent of that man. However, on its second appearance in the sentence, r-radʒul ‘the-man’ is
definite, which is intended to signal to the hearer that this is the same man that you are now familiar with, and even can refer to successfully in other sentences. Therefore, *r-radjul* ‘the-man’ refers not to any individual in the class of men, but rather to a definite man asking for help that the speaker passed by. This is different from saying *marar-tu bi-radgul-i-n yatlub-u musaadat-a-n, fa-saadid-tu radjul-a-n* ‘I passed by a man asking for help, so I helped a man’. The natural reading of this sentence is that both instances of ‘man’ are not co-referential, and the use of indefinite *radjul* ‘a man’ signals to the hearer that the speaker is talking about two different men. Therefore, the speaker of (17.a) intentionally chooses the second mention of *r-radjul* to be definite to make it explicit to the hearer that both instances of ‘man’ are co-referential.

This type of familiarity does not guarantee or indicate that the hearer knows the referent, and can identify him or her. It means that the hearer can single out the referent and refer to him/her using the same linguistic context. In (17.b), for instance, the second definite instance of *at-t‘aalib* ‘the-student’ which appeared indefinite *t‘aalib* ‘a student’ in the previous sentence is now familiar to the hearer. This familiarity does not indicate that she can identify him; in fact, the situation provided denies such an idea, since both interlocutors are talking on the phone and probably living in distant places from each other. It is familiar in the sense that the noun introduced with *al* in this case can be singled out by the hearer through linguistic context. She can later refer to *at-t‘aalib* ‘the-student’ by saying ‘the student who my friend met yesterday and lives next to her house….’ Therefore, this kind of familiarity does not indicate identifiability, but enables the hearer to single out and refer to the referent. In (17.b), it is natural for the hearer to use *al-N at-t‘aalib* ‘the-student’ if she wants to get more details about the student: *maa huwa taxas‘s-u t‘aalib?* ‘What is the student’s major?’ And if the hearer’s mother comes in and asks her about which student she is talking about, she can felicitously answers, ‘the student who my friend met yesterday and lives next to her house’. Other examples of anaphoric familiarity are given in (18):

(18) a. [A girl telling her friend about her sister’s wedding, and that they need a big place to have the wedding in]
b. [A man is talking on the phone with a friend he has met in another country while doing his masters, and has not seen him for more than two years. He tells him about his experience in moving from a city to a village]

\[
\text{qabla } \\text{jaam, } \text{inthal-naa } \text{ilaa } \text{qarya. } \text{ana } \text{wa-saalaat-ii } \text{asbah-naa}
\]

before a year, moved-3.pl to a village. I and-family-my became-3.pl

\[
\text{nuhib-u } \text{l-gariat-a } \text{kaaliran}
\]

3.Pl.like-Nom the-village-Acc much

(A year ago, we moved to a village. (Now) My family and I like the village very much)

c. [A mother encouraging her son to do as Yazan, her friend’s son]

\[
?ahdaa \text{Yazan-u } \text{umma-hu } \text{saasah. } \text{maa-zala-t } \text{tahtafid-u } \text{bi-s-saasat-i}
\]
gave Yazan-Nom mother-his a watch. Still-3.f keep-Nom of the-watch-Gen

\[
\text{?ilaa } \text{l- lyawm}
\]

until the-day

(Yazan gave his mother a watch as a gift. She still keeps the watch until today)

B. Shared Knowledge Familiarity:

Familiarity here is not attained through the linguistic context. In other words, the noun needs not to be mentioned previously for its referent to be familiar; in fact, on this use of al, the noun to which al is attached appears for the first time, though reference is successful. Familiarity, on this use of al, is established through assumed shared knowledge of the referent of the noun between interlocutors. This shared knowledge guarantees successful reference of the noun, and assists the hearer pick out the exact referent intended by the speaker. The examples in (19) explicate this use of al:

(19) a. ?axiiran, iftaray-tu \text{\textit{l-his’aan}}

Finally, bought-1 the-horse

(Finally, I bought the horse)

b. \text{hal } faahad-ta \text{\textit{l-musalsal-a } } \text{?ams?}

\text{q} watched-2.sg the-series-Acc yesterday?

(Did you watch the series yesterday?)

The speaker of (19.a) makes his’aan ‘horse’ definite by adding al ‘the’ to it, l-his’aan ‘the-horse’. In doing so, the speaker intends to signal to the hearer that he/she is not talking about any horse, but rather referring to a particular horse both interlocutors are familiar with, and can pick out.
Reference is successful here because the speaker assumes shared knowledge about this particular horse with the hearer. This shared knowledge can be obtained in many different ways: both have been talking about the speaker’s willingness of buying that horse; the hearer knows that the speaker has been saving money for buying that horse; the hearer accompanied the speaker more than once to buy that horse, but the owner refused to sell it, etc. Interestingly, if this shared knowledge exists between these two interlocutors, and the speakers instead utter ʔaxiiran, iflaray-tu hisʕaan-an ‘Finally, I bought a horse’, the hearer will unambiguously get the message that the speaker bought a different horse from that both of them are familiar with. In this case, it is natural for the hearer to either infer that the speaker could not manage to buy that horse he/she likes, and decided to buy any horse instead, or to ask the speaker about the reason that made him/her give up the idea of buying that particular horse.

The use of the definite al-musalsal ‘the-series’ in (19.b), rather than its indefinite counterpart musalsal ‘a series’, is intended to indirectly convey to the hearer that he/she knows the particular series the speaker refers to. This familiarity could be established in different ways: both interlocutors talked about the plot of that series and how entertaining it is; the speaker hears the hearer talking about that series to his friends often, telling them how much he/she likes it, and cannot wait to see the new episode, etc. If there were more than one series that the hearer follows and seems very excited about, reference would be unsuccessful. In this case, it is natural for the hearer to ask which series the speaker is referring to. This indicates that familiarity per se cannot be taken as a solid ground for definiteness in all cases. More examples of familiarity based on shared knowledge between interlocutors are given in (20):

(20) a. [Two students attended a lecture by a new professor. One of them tells the other two hours later]

\[
i\-tamt\-tu\ bi\-\text{hadii}\-\text{t}\-\text{i}\quad l\-\text{justaad}\-\text{i}
\]

enjoyed-1 with-talk-Gen the-professor-Gen much
(I enjoyed the professor’s talk very much)

b. [A girl thanking her mother for taking her on a nice trip]
(The trip was awesome)

(c. [A man consoling his friend who was informed that his son had a car accident])

I am sorry for the saddening news!

C. Situational Familiarity:

In this use of al, familiarity is attained through a physical situation; in the sense that the referent of the definite noun is present in the visible situation in which the interlocutors are. Therefore, the speaker intends to use a noun with al because he/she wants to signal to the hearer that reference is made to that particular individual or entity visible to both of them. This familiarity type presupposes uniqueness of the referent in the physical situation; otherwise, reference might not be successful if more than one individual to which the descriptive content of the noun applies exist. In this case, the speaker might use a demonstrative to clear up an expected ambiguity. Consider the examples in (21):

(21) a. [A father helping his son improve his spelling]

O my son, hold the pen, and write what I dictate to you

b. [Two men are trying to move a heavy desk and saw a friend of them stepping in the room]

Help us carry the desk

Reference to al-qalam ‘the pen’ in (21.a) is successful because the referent of the definite noun phrase is familiar to the hearer through the physical situation. He can see it, and pick it up with no room of confusing it with any other pen from that intended by the speaker. The use of a noun with al here, presupposes the existence of only one pen in the physical situation. If there were more than one pen in the physical situation, the speaker might either use a body gesture to direct the hearer to the particular pen, use a demonstrative, or use an indefinite noun qalam ‘a pen’, instead. In the latter
case, the speaker signals to the hearer that he/she does not have a particular pen in mind that he/she wants him to pick up, but rather that the hearer can pick any pen from the set of pens available in the physical situation. By using *al-qalam* ‘the pen’, however, visible situation familiarity on the part of the hearer is expressed.

In (21.b), the hearer obtained familiarity of *al-maktab* ‘the desk’ by physically seeing both men holding the intended desk trying to carry it. If the speaker used the indefinite counterpart *maktab* ‘a desk’ instead, the hearer might get a message that there are more than one desk available and that the two men are asking him to move any desk out of the ones available. In this case, it is natural for the hearer to look around the room to see how many desks are there, and if he sees more than one, he might ask about which one they want to move. This is true even if the two men were holding a certain desk when one of them uttered the sentence with an indefinite noun phrase *maktab* ‘a desk’. The hearer in this case might get the impression that both men are not intending to move a particular desk, but chose the one they are holding because it is lighter than the others, and might want to change their minds about the one they are holding after getting more hands, for example. Familiarity established on physical situation is the clearest of all familiarity types, and *al* on this use functions roughly as a demonstrative. More examples of this type of familiarity encoded by *al* are given in (22):

(22) a. [A woman visiting her friend who has bought new sofas]

\[
\text{maa faa? Al-laah, al-kanab-u dsamiil-u-n djiidan}
\]

what will God, the sofas-Nom nice-Nom-N very

(What God wills, the sofas are very nice)\(^\text{18}\)

b. [A host reminding his guests to drink the tea he served them before it cools down]

\[
\text{fal-nafrab f-jaay-a qabl a 'an yabrad}
\]

let-3.Pl.drink the-tea-Acc before to 3.cool down

(Let us drink the tea before it cools down)

c. [A husband eating his dinner, which was prepared by his wife]

\[
\text{at'-t'ašaam-u lađiđ, 'ant-i t'abaaxat-u-n maahirat-u-n yaa ʃaa?ʃa}
\]

\(^\text{18}\) The introductory phrase is used by Arabs to expresses appreciation, joy, praise for something, event, or person. It is used as an expression of respect, while at the same time serving as a reminder that all accomplishments are so achieved by the will of Allah.
the-food-Nom delicious, you-f cook-Nom-N skillful-Nom-N O Aisheh
(The food is delicious; you are a good cook, Aisheh)

However, as hinted above, in MSA, as other languages, not all definite references can be analyzed as instances of familiarity. The following section investigates cases where familiarity fails to explain simple definiteness in MSA, and instead identifiability is used to analyze definiteness in these cases.

5.1.2 Identifiability

Familiarity, though used often as a semantic-pragmatic ground on which definiteness is established, fails in certain cases to explicate successful reference. Therefore, as mentioned in section (2), semanticists sympathetic with familiarity concept propose that a more general term, identifiability, is the basis of definiteness. The major distinction between familiarity and identifiability is that on identifiability account, reference can be successful, though familiarity of the referent on the hearer’s part is not obtained. Put differently, the hearer does not need to be familiar with the referent of a noun phrase to be able to identify and single it out in a certain context. In a situation where the speaker is trying to fix his car’s flat tire, that speaker might say to his wife who just showed up, even without needing to turn around, the sentence in (23).

(23) naawil-ii-nii l-mifak radzaaʔan
pass-2.f-me the-screwdriver please
(Pass me the screwdriver, please)

In this example, the reference of l-mifak ‘the screwdriver’ is impossible to be successful on familiarity account. There is nothing in the physical situation that necessarily assures that the hearer even saw that screwdriver or knew about it. The hearer upon hearing l-mifak ‘the screwdriver’ tends to infer that there must be one screwdriver in this physical situation, and she needs to employ her extralinguistic apparatus to identify it, such as looking around and see the screwdriver somewhere, or looking for it and locate it. So the use of al in this case does not presuppose that l-mifak ‘the screwdriver’ is familiar to the hearer in the real sense, but rather functions as a set of indirect instructions which triggers the hearer to work out the referent of the NP indirectly.
Moreover, identifiability account explains the use of *al-N* successfully to refer to a definite individual not only unfamiliar to the hearer, but also does not exist in the visible situation. In a context where the speaker has just bought a new car, and drove back home, he might felicitously utter the sentence in (24), asking his son, who does not even know that his dad bought a new car, to wash that new car.

(24) *iðhab* w-*aysil* s-*sayyarat-a* l-*dgadiida*
   2.m.go and-wash the-car-Acc the-new

(24) *Go and wash the new car*

Reference would not be successful in (24) on familiarity account; in fact as stated above, the hearer has no idea about his father buying a new car. What makes reference felicitous here is the message conveyed indirectly to the hearer by the use of *al-N*, which gives him two things: there is a particular new car his dad is referring to - this is the linguistic part, and that he needs to identify that particular car – this is the extralinguistic part of the message. Upon hearing *s-sayyara ‘the car’,* the hearer infers that his father is not referring to any car, but rather to a particular car. Following this inference, the hearer tends to exploit extralinguistic tools to identify that car. He might go directly to the garage since he knows that his father often parks his car there, or get out and try to locate it in front of the house if he knows that the old car is still in the garage. Interestingly, the hearer would be able to identify the referent of a noun phrase that he is not familiar with, and does not even exist in the visible situation. More interestingly, *s-sayyara ‘the car’* represents completely novel information to the hearer, and is still definite in form and sense. This case weakens the widespread claim in discourse analysis that definiteness represents old knowledge, and indefiniteness represents novel knowledge. Other examples of *al-N* which tends to be analyzed on identifiability account are given in (25); the analysis framework provided for (23&24) can be applied to those in (25).

(25) a. A: *limaʔanta murtabik?* why 2.sg.m confused?
   B: *ʔada‘imm-u ʔanna-nii ʔd̩a‘-tu l-miḥfaḍ’at-a* 1.sg.think-Nom that-I lost-1.sg the-wallet-Acc *hunaa fi l-yurfa* here in the-room
   A: *ħasanan, sa-ʔabhaθ-u ʔsan-haa* well, Fut-1.sg.look-Nom for-it
(A: Why are you confused? B: I think I have lost the wallet here in this room. A: Well, I will look for it)

b. [In a soccer field where boys are playing soccer, one boy kicked the ball behind some bushes. A boy just showed up and expressed his desire to join them. One of the boys says to him]

\[
\text{ibhaθ ʕan l-kurat-i ʔin ʔaradt-a ʔan talaʕab-a maʕa-naa}
\]
2.m.look for the-ball-Gen if want.2.m-Acc to play-Acc with-us

(Look for the ball, if you want to play with us)

c. [In a two-story house, a man who is visiting his friend asks the host where he can sleep. The host answers]

\[
as-sariir-u fii tʕaabiq-i l-ʔawwal
\]
the-bed-Nom in the-floor-Gen the-first

(The bed is in the first floor)

In certain cases, however, both familiarity and identifiability fail to explain \textit{al-N} reference in MSA. These cases entertain an account based on uniqueness criterion of the referent of the \textit{al-N} used. This aspect of definiteness is investigated in (5.1.3) below.

5.1.3 Uniqueness

Among the major uses of \textit{al} as a definite article is that it signals to the hearer that the referent of the noun it is attached to is a unique individual or genus in a certain situation. Put differently, in these cases, the speaker by using \textit{al-N} intends to convey to the hearer that in the situation he/she is referring to there is only one entity that satisfies the descriptive content of the noun used, and based on this, the speaker can successfully use \textit{al-N} in such contexts. In these cases, it is worth noting, there is no guarantee that the hearer can identify that unique individual in the real sense. Consider the example in (26):

(26) [On first day of classes, one student asks his classmates]

\[
\text{hal sa-yəʔəi l-ʔustaad-u l-ʔaywm?}
\]
q Fut-3.come the-professor-Nom the-day?

(Will the professor come today?)

The use of the definite \textit{al-N} in (26) is successful, and rather necessary, though there is no guarantee that the hearers can identify who the professor of that class is. There is a strong possibility that the name of the professor teaching that class was anonymous, as often happens in some schools.
It is hard to account for the use of *l-ʔustaað* ‘the-professor’ based on familiarity or identifiability. If the professor did not show up on that class, and some of the students ran into that professor few hours later, there is no guarantee, based on (27), that they would identify him. This use of the definite can be more felicitously accounted for based on uniqueness criterion. The use of *al*-*N* in this example is successful because it is part of our common-sense knowledge that every face-to-face class must have a professor. In other words, the use of *al*-*N* here signals to the hearer that there is a unique individual, salient in the domain of the class context, which satisfies the descriptive content of the noun.

Uniqueness, rather than familiarity or identifiability, legitimizes the use of *al*-*N* in this example.

As stated in Lyons (1999), uniqueness criterion seems more appealing to account for cases where the referent is potential or in the future. In a context where two interlocutors are discussing a future semi-final soccer match, one would utter (27) which uses an *al*-*N* noun phrase:

(27) *al-faaʔiz-u*  
*the-winner*-Nom in this the-match-Gen *faaʔiz*  
the-final *niaaʔi* 

(The winner of this match will move on to the final match)

Successful reference of *al-faaʔiz* ‘the winner’ cannot be established on familiarity or identifiability account; simply because neither the speaker nor the hearer knows which team will win. If you ask either the speaker or the hearer before the match about who the winner is, the natural answer is ‘I do not know; either one could win’. This indicates that the winner of that match is not identified or known to either the speaker or the hearer. However, uniqueness criterion seems to work better in explaining the use of *al*-*N* in this example. Both interlocutors share common-sense knowledge about semi-final soccer matches. In such matches, one of the two competing teams has to win the match; there is no room for such a match to end in a tie. Based on this shared knowledge, reference is made to the unique, but not identified in the real sense, team which will win the match. It refers to the most salient unique entity in the coming soccer match, which satisfies the descriptive content of the noun *faaʔiz* ‘winner’. Other examples of the use of *al*-*N* that resist familiarity or identifiability accounts, and instead entertain uniqueness account are given in (28):
(28) a. [A family looking for a park to have a picnic in found a good one, but discovered that it was closed. One of them says]

\[
\text{al-hadiiqat-}u \ \text{munlaqa.} \ \text{?aduunn-}u \ \text{?anna} \ l-\text{haaris-}a \ \text{yadara mubakkiran}
\]

\[
\text{the-park-Nom closed.} \ 1.\text{sg.think-Nom that the-guard-Acc left early}
\]

(The park is closed. I think the guard left early)

b. [After moving his son to a new school, and never saw the principal before, a parent having an argument with his son’s teacher over some grade issues says:]

\[
\text{?uriid-}u \ \text{?an ?araa} \ l-\text{mudiir-}a \ \text{radgaa?an}
\]

\[
1.\text{sg.want-Nom to see the-principal-Acc please}
\]

(I want to see the principal, please)

c. [An immigration officer talking to an international visitor]

\[
\text{?asi?in-}i \ l-\text{dgawaaz-}a \ \text{min fad?lik}
\]

\[
give\text{-me the-passport-Acc in care.your}
\]

(Give me the passport, please)

In addition to denoting a unique individual salient in a certain context, \textit{al-}\textit{N} can be used to denote a unique kind, species, or, as traditional Arab grammarians like to put it, genus. Here, reference is not made to any member of a kind or species, but rather to the whole kind or species as a unique entity. Consider the examples in (29):

(29) a. \textit{al-kalb-}u \ \textit{hayawaan-u-n} \ ?aliif

\[
\text{the-dog-Nom animal-Nom-N domestic}
\]

(The dog is a domestic animal)

b. \textit{asˤ}sqˤ\textit{aqr-u} \ tˤaʔir-\textit{u-n} \ dgaarih

\[
\text{the-hawk-Nom bird-Nom-N of prey}
\]

(The hawk is a bird of prey)

c. \textit{af-fayˤa}n-u \ \textit{saduww-u} \ l-\textit{ʔinsaan}

\[
\text{the-devil-Nom enemy-Nom the-Man}
\]

(The devil (Satan and his demons) is the enemy of Man)

d. \textit{al-qird-}u \ \textit{mina 0-ʔadiyyaat}

\[
\text{the-monkey-Nom from the-mammals}
\]

(The monkey is a mammal)

The definites in (29) denote kinds or species. They do not refer to any member of that kind or species, but rather to the kind or species as a whole. These kinds are unique, in the sense that each kind is the only salient entity in the context domain, which might be the whole world. In (a), for example, reference is made to the unique individual species of dogs, not to any dog or a group of
dogs. Likewise, in (b) the definite asˤ-sˤaqr ‘the hawk’ is intended to denote the whole unique species of hawks, not to any member of that species. The same interpretation holds for the rest of the definites in (29).

Therefore, the use of al as a pure definite marker in MSA entails that the referent of al-N is familiar, identifiable, or unique. Noun phrases prefixed with al whose referents fail to meet at least one of the definiteness criteria are rendered semantically not definite; hence, al will prove not to be an exclusive definite marker. In this case, it is plausible to claim that morphology per se is not reliable enough to determine in/definiteness in MSA. A semantic analysis based on the categories on which al definite marker is established has to be incorporated to probe the synchronic function of al in MSA, and whether it is still a faithful carrier of definiteness in the language, or has developed other grammatical uses which are not much relevant to its original function. This will be examined in the following section.

5.2 Other Uses of al

The definite article al does not function as a pure definite marker in MSA. In fact, it has developed other uses not necessarily related to definiteness. The presence of al-Ns is much more frequent than the presence of bare Ns. This heterogeneous frequency of al-N and bare N indicates that the presence of al serves other grammatical functions, and that it no longer marks definiteness exclusively. Sub-section (5.2.1) investigates the definite article al introducing nouns whose referents are not semantically definite according to the definiteness criteria. These al-Ns can be replaced with their bare N counterparts with slight change in meaning. Sub-section (5.2.2) discusses cases where al-Ns are used in indefinite environments, and still these al-Ns are not interchangeable with their bare N counterparts. This indicates that the presence of al here serves other grammatical functions not necessarily related to in/definiteness. Sub-section (5.2.3) probes cases where al attaches to nouns, adjectives, and adverbs which are semantically neutral, if not irrelevant, to semantic definiteness. The presence of al in these cases explicitly serves other grammatical functions like phrase-clause distinction.
5.2.1 The Use of al in Indefinite Environments

Although the definite article *al* in MSA is used with nouns whose referents are familiar, identifiable, or unique, this use of *al* is not exclusive. In fact, it is used with nouns whose referents are completely the reversed, i.e. semantically indefinite nouns; they are not familiar, identifiable, or unique. The nouns denote random individuals of certain classes upon which the descriptive content of the nouns used apply. In a situation where a senior high school student is talking with his uncle about his plans after graduating from high school, the uncle asks his nephew:

(30) Uncle: *maādaa tuxat’itˤ ʔan tafʕal baʕda ʔan tunhii l-marhala ʔ-ʔaanawiyya?* what 2.plan to do after that 2.finish the-stage the-secondary?
Student: *kul ihtimaam-ʔii munsʕabbun ʕala l-duxxuli l-ʤaamiʕa* all attention-my focused on admission the-university
Uncle: *hal tufadʕil ʤaamiʕa muʕayyana?* do 2.prefer a university particular?
Student: *laa, ʔay ʤaamiʕa ʤayyeda* no, any university good
Uncle: *bi-t-tawfiq yaa bunay* with-the-luck O sonny
(Uncle: What do you plan to do after finishing the secondary stage? Student: All my attention is focused on being admitted to a university. Uncle: Do you prefer a particular university? Student: No, just any good university. Uncle: Good luck sonny)

According to the traditional view of definiteness, the use of *al* in *l-ʤaamiʕa* ‘the-university’ is sufficient to mark the noun prefixed with it *ʤaamiʕa* ‘a university’ semantically definite. However, it is clear from the scenario that *l-ʤaamiʕa* ‘the-university’ is not semantically definite according to the definiteness criteria. In other words, the referent of *l-ʤaamiʕa* ‘the-university’ is not familiar, identifiable, or unique to both interlocutors. This is explicated by the uncle’s question of whether the student has in mind a particular, definite university he refers to, and the student’s answer expressing his interest to be admitted to any good university, not to a unique one he has in mind. Therefore, no reference is expressed in this grammatically definite instance of *l-ʤaamiʕa* ‘the-university’, yet the definite article *al* is attached to this noun phrase. The exact denotation of *l-ʤaamiʕa* ‘the-university’ is that it denotes any random individual *ʤaamiʕa* ‘university’ in the class of entities to which the

---

19 In all examples of this section, only *al-Ns* in indefinite environments, appearing in bold, will be discussed; NPs with semantically definite referents are irrelevant to this section.
The noun phrase *l-matˤaṣim* ‘the-restaurants’ is definite in form; it is introduced with *al*, the definite article. However, is *l-matˤaṣim* ‘the-restaurants’ semantically definite? Put differently, does this noun phrase denote familiar, identifiable, or unique referents? A straightforward answer is no. This answer is justified by the facts introduced in the context; the speaker has never been or heard of the
country Luxembourg. This indicates that both the speaker and the hearer are not familiar with its restaurants; neither the speaker nor the hearer can identify them. They are thousands of miles far, and they are not unique because both interlocutors cannot even tell how many restaurants there are in Luxembourg. It is clear that \textit{l-mat'aaṣim} ‘the-restaurants’ fails to satisfy any of the definiteness criteria clusters, yet the definite article \textit{al} is felicitously used with \textit{mat'aaṣim} ‘restaurants’. A more plausible interpretation of \textit{l-mat'aaṣim} ‘the-restaurants’ is that it denotes any random members of the class of \textit{mat'aaṣim} ‘restaurants’. More precisely, it denotes semantically indefinite entities, unfamiliar, unidentified, and not unique. That said, \textit{al} here is no longer considered a grammatical carrier of semantic definiteness, hence its presence does not guarantee that the noun it is attached to is semantically definite.

However, one may argue that \textit{l-mat'aaṣim} ‘the-restaurants’ denotes definite referents based on cross-reference familiarity. The ‘restaurants’ has not been mentioned before, but Luxembourg and eating have been mentioned. By commonsense knowledge, any country has restaurants, and the speaker by using the word eating conjures up for the hearer all places relevant to eating like cafés, restaurants, etc. Therefore, both interlocutors refer to the restaurants there, not to any other restaurants. It is worth mentioning that English admits the use of ‘the restaurants’ in the same context, even though the speaker has no particular restaurants in mind. This analysis, however, is untenable. In MSA, the same sentence is felicitous even if the singular \textit{l-mat'aim} ‘the-restaurant’ is used. The use of a singular grammatically definite restaurant cancels cross-reference familiarity; no country in the world is by default has one restaurant which can be invoked as part of the commonsense shared knowledge. It is worth noticing that English does not allow a singular restaurant in the same context.

In another situation where two friends are chatting about a movie currently playing in some movie theatres, one of the two interlocutors asks his friend whether he watched that movie or not. Their conversation is in (32):

(32) Friend1: \textit{hal faaḥdta} \textit{film t-taaytanik?} \\
q \textit{watched.2.sg} \textit{film the-titanic?} \\
Friend 2: \textit{laa, lam ṣufaaḥid-hu bašd, laakin sa-ʔaḍhab li-s-siūnamaa}
no, NEG 1.watched-it yet, but Fut-1.go to-the-cinema li-ʔufaaahoma-hu to-1.watch-it

Friend 1: hal tfakkir fii ʔan taðhab li-siinamaa muʕayyana?
q 2.think in to 2.go to-a cinema particular?
Friend 2: laa, sa-ʔaðhab li-ʔay siinamaa taʕridˤu-hu
No. Fut.1.go to-any cinema play-it

(Friend 1: Did you watch the movie Titanic? Friend 2: No, I have not watched it yet, but I will go to a movie theatre to watch it. Friend 1: Are you thinking of going to a particular movie theatre? Friend 2: No, I will go to any movie theatre (currently) playing it.)

In this scenario, the movie Titanic is familiar to both interlocutors; however, it is clear that the movie theatre(s) currently playing it is not, yet s-siinamaa ‘the-movie theatre’ occurs with the definite article al. This lack of familiarity of the referent of s-siinamaa ‘the-movie theatre’ on the part of both the speaker and the hearer is explicitly stated by the utterances of both interlocutors, which followed the occurrence of the grammatically definite s-siinamaa ‘the-movie theatre’. The hearer ‘Friend 1’ expresses his unfamiliarity of the movie theatre the speaker refers to by asking the speaker if he intends to go to a particular movie theatre. If the referent of the movie theatre were known to the hearer in a way or another, he would not ask this question; notice that this question is felicitous even if the speaker used an indefinite noun phrase siinama ‘a movie theatre, instead. The speaker’s response to the hearer’s question explicitly states that he is not referring to a particular individual movie theatre, but rather to any individual movie theatre out of the class of ‘movie theatre’. This response, also, clearly illustrates that the speaker has no particular movie theatre he wants the hearer to identify. As for uniqueness, the word ʔay ‘any’ which preceded siinama ‘a movie theatre’ in the speaker’s response to the hearer’s question of whether he is thinking of going to a particular movie theatre, invalidates uniqueness account. ʔay ‘any’ co-occurs with indefinite nouns in MSA; notice that it is followed by a bare noun siinama ‘a movie theatre’, and indefinite nouns do not denote unique referents by default since any random individual satisfying the semantic content of the noun used can be picked out.

Another example taken from Quran (18:109) indicates that even when Classical Arabic was used, there were instances of al-N in indefinite environments. Consider (33):
In this verse two instances of *l-bahru* ‘the-sea’ occurred. Both of them are grammatically definite; they are introduced with *al*. The second instance represents anaphoric familiarity since in the linguistic domain, another instance of the same noun phrase is previously mentioned, and both phrases are coreferential. Therefore, the hearer is familiar with the referent of the second instance of *l-bahru* as he/she assumes that it refers to the same referent of the noun previously mentioned. That said, the second instance needs no further investigation. As for the first instance of *l-bahru* ‘the-sea’, there is nothing in the verse, neither in the preceding or following verses, which guarantees that the addressee, Prophet Mohammad, is familiar with, or can identify the particular sea Allah is referring to. For traditional Arab grammarians, the first instance of *l-bahru* ‘the-sea’ is semantically definite because the noun here does not refer to any member of the class ‘sea’, but rather to the ‘genus’ sea. Therefore, reference is successful here on uniqueness account; instead of referring to a unique individual sea, Allah refers to the unique genus sea. However, closer investigation of the phrase “walaw dżi’naa bimiθlihi madadaa” (even if We brought the like of it as a supplement) invalidates this interpretation. Allah wants to emphasize that his words are infinite by saying that even if We brought a similar sea to the one mentioned first, and used that other sea(s) as ink to write the words of Allah, any sea would not be sufficient for writing Allah’s words. This shows that the first instance of *l-bahr* ‘the-sea’ does not refer to the genus sea, but rather to any random member in the genus, for which a similar random individual sea out of the class of ‘seas’ can be also brought. Notice the use of *bimiθlihi madadaa* ‘the like of it as a supplement’, which indicates that a similar individual, not a genus, can be brought. This is interesting because a genus by default is unique; there is no other genus with the same characteristics. However, any genus has members which share the same distinctive attributes which qualify them to be members of that genus. Therefore, when Allah mentioned ‘the like
of it as a supplement’. He intends another individual sea out of the class ‘seas’ which shares with the first random individual sea the same characteristic attributes.

Interestingly, the examples in (30-32) can be restated with indefinite nouns replacing the ones in bold with almost no change in meaning with respect to definiteness. This gives further support for treating these NPs as semantically indefinite, and that the occurrence of the definite article al does not contribute definiteness to these nouns. This also has bearing on treating al as a pure definite marker, in the sense that it carries semantic definiteness, and that though it functions as a definiteness carrier in some cases, it does not do so exclusively. This use of al as not a definite marker might have developed some time after its first stage as a pure definite marker, which indicates that taking its presence/absence as an absolute ground to determine in/definites in MSA is untenable. Example (30)\(^{20}\) is repeated in (34) with a change of the definite noun l-ʤaamiša ‘the-university’ into its indefinite counterpart ʤaamiša ‘a university’.

(34) Uncle: maaða tuxat’il ʔan taʃal baʃda ʔan tunhii l-marhaʃa ʔa-taṇawiyya? what 2.plan to do after that 2.finish the-stage the-secondary?

Student: kul ihtimaamii muns’abbun ʕalaʃ duxuuli ʤaamiʃa all attention-my focused on admission a university

Uncle: hal tufad’il ʤaamiʃa muʃayana? q 2.prefer a university particular?

Student: laa, ʔay ʤaamiʃa ʤayyida no, any university good

Uncle: bi-t-tawfiq ʔaa bunay with-the-luck O sonny

(Uncle: What do you plan to do after finishing the secondary stage? Student: All my attention is focused on being admitted to a university. Uncle: Do you prefer a particular university? Student: No, just any good university. Uncle: Good luck sonny)

Replacing l-ʤaamiša ‘the-university’ with ʤaamiša ‘a university’ does not affect our interpretation of the sentence; notice that the translation given for both noun phrases is similar ‘a university’. Both l-ʤaamiša and ʤaamiša in (30&34) denote a referent which is unfamiliar, unidentifiable, and not unique. The slight difference that I can notice, as a native speaker, is that the use of the definite sounds more natural than the indefinite. This has nothing to do with definiteness, but rather with a

\(^{20}\) The definite nouns in the other two examples (29&30) can be replaced with their indefinite counterparts, with no change on their interpretation in terms of definiteness.
development in the use of *al* which tends to co-occur with most nouns except for those which are intended to be existential like there is/are Ns. This is reflected in the frequent uses of the grammatically definite nouns in MSA, as opposed to a receding use of indefinites. It sounds to me that a tendency of using only *al-Ns in MSA is being developed\(^{21}\).

Other examples of *al-N* in indefinite environments are given in (35):

(35) a. [A student studying in a college far from her home city is telling her friend how she keeps in touch with her family]

\[
\text{nataabaadal-u } \text{l-ʔaxbaa-ʔa } \text{ʔanaa wa-ʔaaʔilat-ii } \text{fii nihaayat-i l-ʔasbuus}-
\]

3.Pl.exchange-Nom the-news-Acc I and-family-my in end-Gen the-week

(My family and I exchange news in the weekend)

b. [A women expressing her fascination with trees]

\[
\text{ʔaʃaq-u } \text{l-dguluus-a tahta } \text{f-ʃaʃara}
\]

1.love-Nom the-sitting under the-tree

(I love sitting under a tree)

c. [A mother complaining about her son not focusing in school]

\[
\text{yumdˤii } \text{muʃʔam-a waqti-hi fii mufaahadat-i l-ʔaflaam}
\]

3.spend most-Acc time-his in watching-Gen the-films

(He spends most of his time watching movies)

d. [A father asks his younger son where his older sister could be; the son answers: she could not find something to drink, so]

\[
\text{ðahabat } \text{li-taʃtarii } \text{l-ʃalib}
\]

went.3.f to-3.f.buy the-milk

(She went out to buy milk)

e. [A college student asks her friend why she did not apply for a college after graduating from high school; the friend answers]

\[
\text{læa } \text{ʔuhibb-u } \text{ʔan } \text{yumdˤii saaʃaat t'awiilah wa-ʔanaa ʔustamiʕi l-ʃmuadˤaraat}
\]

NEG 1.like-Nom to 1.spend hours long and-I 1.listen to the-lectures

(I do not like to spend long hours listening to lectures)

\(^{21}\) Investigating this grammatical tendency falls out of the scope of this dissertation. In addition, no attempt to explain the function of *al* when attached to nouns with semantically indefinite referents will be made. The main concern of this chapter is to prove that relying on the presence/absence of *al* as an exclusive marker of definiteness is problematic, and that a semantic analysis based on the definiteness criteria is required to determine definites and indefinites in MSA.
It is worth mentioning that in all the sentences in (35) the definite nouns in bold can be replaced with indefinite counterparts with no change in meaning in terms of definiteness, though as mentioned above the ones with grammatically definite nouns sound more natural than those with bare nouns.

However, in certain cases *al-N* is used in indefinite environments where it is ungrammatical to replace the *al-N* with its bare *N* counterpart. This ungrammaticality has nothing to do with definiteness because the *al-N* in this case does not denote a semantically definite referent, and if it is an issue related to definiteness, using bare *N* should be non-problematic. These uses are discussed below.

### 5.2.2 Obligatory *al* in Indefinite Environments

Support for claiming that *al* no longer functions as a pure grammatical encoder of semantic definiteness, and that it has developed other uses not essentially related to this semantic ingredient comes from facts that *al* is used with nouns entertaining an indefinite interpretation only; nonetheless, these NPs are not mutually interchangeable with their grammatically indefinite counterparts. This indicates that the use of *al*, in these cases, does not relate to definiteness category, and that its use is enforced by other grammatical requirements entertained by the grammar of MSA. In a situation where a mother is calling her daughter, complaining that her daughter, who is working in a different country, is not sending her enough money; the mother argues that her daughter does not spend her money appropriately. The daughter argues back asking her mother to give her examples of her misuse of her money. The mother utters the sentence in (36):

(36) ḥadī unn-u ḥanna-ki tunfiqiina muḥḍam-a daxli-ki ẓalaa
1.think-Nom that-you 2.f.spend most.Acc income-your on
*1-malaabis/*malaabis
the-clothes/*clothes
(I think you spend most of your income on [buying] clothes)

In this example, for the sentence to be grammatically acceptable the noun *malaabis* ‘clothes’ must be introduced with *al*, otherwise it will be grammatically unacceptable. Interestingly, *l-malaabis* ‘the-clothes’ and *malaabis* ‘clothes’ both denote semantically indefinite referents in this example.
Thus, if ungrammaticality is due to in/definiteness mismatch, there must be no problem in using the bare *N malaabis* ‘clothes’ as it also denotes a semantically indefinite referent. However, we need to verify that *l-malaabis* ‘the-clothes’ is semantically not definite, in the sense that it does not refer to a familiar, identifiable, or unique object. The speaker, in using *l-malaabis* ‘the-clothes’, does not refer to a particular collection of clothing objects she is familiar with; as stated above her daughter lives in a different country. Notice also the use of *ʔadֳ‘unn-u* ‘I think’, which indicates that the speaker is not familiar with which clothes she is talking about, or can identify them. As for uniqueness account, there is nothing in the sentence, or in the situation provided, that indicates that the speaker is referring to a unique collection of clothes on which her daughter spends her money. Notice the use of the verb *tunfiqiin* ‘2.f.spends’ which indicates habituality of the event of spending; one cannot habitually spend money on buying the same unique collection of clothes. A more plausible interpretation is that when the speaker used *l-malaabis* ‘the-clothes’, she meant any random object to which the descriptive content of the noun applies. In addition, this analysis of the denotation of *l-malaabis* ‘the-clothes’ as denoting a semantically indefinite collection of clothes complies with the habitual reading of the sentence introduced by the present-tense verb *tunfiqiin* ‘2.f.spends’. According to this reading, the mother says that generally, whenever her daughter spends money, she spends it on buying different pieces of clothes. This example is quite interesting because it clearly illustrates the interaction between genericity and definiteness, and that accurately distinguishing grammatical definiteness and semantic definiteness is crucial to investigating this kind of interaction.

In another situation, a host in a talk show asks her guest what he usually does whenever he feels upset or depressed. The guest utters (37):

(37) *ʔaxruʤ-u xaaridʤ-a l-madiinat-i fi l-laylʤ-%layl-i*

1.get out-Nom outside-Acc the-city-Gen at the-night-Gen/%night

*wa-ʔamfiit bayna l-ʔasdaarʤ-%ʔasdaar-i*

and-1.walk between the-trees-Gen/%trees

(I go outside the city at night and walk between trees)

This example is quite interesting because it has two instances of *al-N* that cannot be replaced with their bare *N* counterparts, though neither of these *al-Ns* is semantically definite. In using *al-layl* ‘the-
night’, the speaker does not intend to trigger to the hearer that he refers to a particular night she is familiar with in a way or another, or can identify using extralinguistic machinery; notice that the speaker refers to *al-layl* ‘the-night’ as a time correlated with his going outside the city when he feels upset. Since he does this only when he is upset, and there is no specific feeling upset states he is referring to, this night is never the same because he might feel upset every two or three weeks, for example, and no single night lasts that long. This invalidates uniqueness account also. Since feeling-upset states are not known when to happen, the time they happen in is naturally not known or indefinite. Therefore, the speaker, though using the grammatically definite *al-layl* ‘the-night’, he, as well as the hearer, understands that it denotes any random night that coincides with his accidental feeling upset state. The second instance of *al-N*, *l-ʔafār* ‘the-trees’, is clearly semantically indefinite. The speaker does not refer to definite trees he walks between whenever he feels upset and goes outside the city. In fact, being outside the city is itself indefinite; there are many places which are technically considered outside a certain city. The speaker when using *l-ʔafār* ‘the-trees’ does not intend to refer to semantically definite trees in the real sense. He rather refers to any trees which happen to exist whenever he feels upset and decides to go outside the city for a walk between trees. In fact, if someone takes a photo of a certain pair of trees the speaker walked between once and shows it to the speaker, there is nothing in the sentence that guarantees that he would be able to recognize them, or remember that he walked between them once. In addition, if the host in the show asks the speaker about the number of trees he walked between on the last time he did this, or to describe to her every single pair of trees he walked between, those questions would be taken as ridiculous by the speaker and the audience, and probably will cause a burst of laughter. A more plausible interpretation is that this grammatically definite noun phrase is semantically indefinite based on the definiteness criteria; the referents of these trees are not familiar, identifiable, or unique to both interlocutors. This example, the previous one, and the other examples in (38) below show that the use of *al* cannot be taken as a necessary and sufficient condition for distinguishing definites from indefinites in MSA, though *al-N* in certain cases is not in complementary distribution with *N*. 
(38) a. [A judge asks a woman about the grounds for her divorce; the woman replies]

\[
\text{huwa laa yahtamm-u bii wa-la\text{-}aa bi-\text{awlaad-ih, fa-huwa yukarris-u}}
\]

He NEG 3.take care-Nom of me and-NEG of-children-his, as-he 3.dedicate-Nom all time-his to-following-Gen the-news/*news from-a channel to another (He does not take care of his children or me as he spends his time watching newscasts from one channel to another)

b. [A sign on a beach says]

\[
mamnuu s-sibaaha/*sibaaha
\]

forbidden the-swimming/*swimming (Swimming is forbidden)

c. [A host asks his guest to describe things she enjoys doing with her leisure time; the guest answers]

\[
\text{ʔi\text{-}daqad l-qahwa/*qahwa wa-l-masfi/*masfi \text{ʕalaa f-\text{jaat-i'i?}}}
\]

Making the-coffee/*coffee and-the-walking/*walking at the-beach (Making coffee and walking at the beach)

d. [Two men met in a hotel’s restaurant. One of the two told the other that for twenty years now he has been on many vacations all around the world. The other man asks about the times his hotel mate had his vacations; the man replied]

\[
\text{ʔana ʔusaafir-u \text{ʔalaa\text{-}a marrat-i-n fii s-sana/*sana. saafar-tu \text{ʕafraat-a}}}
\text{I 1.travel-Nom three-Acc times-Gen-N in the-year/*year. travelled-1 ten-Acc times-Gen-N in the-year/*year. travelled-1 ten-Acc times-Gen-N in the-year/*year.}
\]

(I travel three times a year. I travelled ten times in spring, and fifty times in winter)

e. [A young woman took her mother who has lost her hearing in both ears to a doctor’s appointment. The doctor noticed that the old woman is too upset, and asked her daughter about this. The daughter said that her mother is a very pious Muslim and she was upset because since she became deaf a year ago she could not hear any calling to prayer, which is heard five times a day]

\[
mun\text{ðu sana, lam tastat\text{-}i\text{i\text{-}a l-\text{ʔadaan}/*\text{ʔadaan}}}
\]

since a year, NEG 3.f.could to 3.f.hear the-calling to prayer/*calling to prayer (Since a year, she could not have heard a calling to prayer)

In addition to the use of \text{al} with nouns in an indefinite environment, where \text{al-}N is sometimes interchangeable with \text{N} and sometimes not, \text{al} has developed other uses where it obligatorily attaches to nouns for grammatical reasons irrelevant to in/definiteness phenomenon. These uses are discussed in (4.2.3) below.
5.2.3 Other Obligatory Uses of *al*

Not all uses of the definite article *al* can be accounted for based on definiteness or indefiniteness accounts. In other words, *al* obligatorily attaches itself to nouns, adjectives, and adverbs which are semantically neutral, if not irrelevant, with regard to both definiteness and indefiniteness. These uses of *al* are triggered by some grammatical rules which, for reasons not related to definiteness phenomenon, entertain *al*-N phrases only. Therefore, in these cases the presence of *al* cannot be taken as an indicator of definiteness. The attributive adjective presents a clear case of these uses of *al*. In MSA, the attributive adjective agrees with the noun it modifies in number, gender, and grammatical definiteness. Consider the examples in (39).

(39) a. haaðihi  l-bint  l-dgamiil-a
    this.f  the-girl  the-beautiful-f
    (This beautiful girl)

b. al-kursiy  l-maksuur
    the-chair  the-broken
    (The broken chair)

c. al-muʃallim-aat  l-mafyuul-aat
    the-teacher-Pl.f  the-busy-Pl.f
    (The busy [female] teachers)

d. atˤ-ʔullaab  l-mudʔahid-uun
    the-students  the-diligent-Pl.m
    (The diligent students)

The examples in (39) are grammatically definite noun phrases in which the adjectives modifying them are obligatorily definite. Grammatical definiteness in adjectives is required for two reasons not related to definiteness. The first reason has to do with syntactic agreement between head noun and complement adjective in MSA. Put differently, the adjective, the complement in this case, is rendered grammatically definite to fulfill agreement feature, which requires that an adjective modifying a head noun have to agree with that noun in number, gender, and definiteness. The adjective itself is not semantically definite; simple semantic definiteness in MSA is restricted to NPs which denote familiar, identifiable, or unique referents. Therefore, the presence of *al* in these adjectives is syntactically motivated. The other more interesting grammatical reason has to do with
distinguishing a subject-predicate sentence from a noun phrase. Changing the grammatically definite adjectives in all the sentences in (39) into grammatically indefinite ones yields in subject-predicate sentences, rather than noun phrases. Consider (40) below:

(40) a. haadīhi l-b'int  ēdamiil-a
    this.f  the-girl  beautiful-f
    (This girl is beautiful)

b. al-kursiyy  maksuur
    the-chair  broken
    (The chair is broken)

c. al-muṣallim-aat  mafyuuul-aat
    the-teacher-Pl.f  busy-Pl.f
    (The [female] teachers are busy)

d. atˤ-tˤullaab  mudqahid-uun
    the-students  diligent-pl.m
    (The students are diligent)

Notice that the word order in the sentences (39) is similar to that in the phrases (40). Therefore, the difference between the two expression types is brought upon by the presence/absence of al on the adjective without any change on the definiteness status of the NPs in both the phrases and the sentences. It is clear that this use of al is purely grammatical, and al here cannot be taken as a definite marker.

A similar case is found in demonstrative-noun expressions. In a demonstrative-noun expression, the expression is rendered a phrase if the noun following the demonstrative is al-N, and a sentence if the noun following the demonstrative is bare N. Consider the examples in (41).

(41) a. haadāa  l-walad  ‘this the-boy’
    (This boy)

    a’. haadāa  walad  ‘this boy’
    (This is a boy)

    b. haadīhi  sˤ-sˤuura  ‘this.f the-picture’
    (This picture)

    b’. haadīhi  sˤuura  ‘this.f picture’
    (This is a picture)

    c. haadʔulaaʔ  l-qiʔatˤ  ‘these the-cats’
    (These cats)

    c’. haadʔulaaʔ  qiʔatˤ  ‘these cats’
    (These are cats)

    d. ḥulaaʔik  l-fannaanaat  ‘those the-artists.f’
    (Those [female] artists)

    d’. ḥulaaʔik  fannaanaat  ‘those artists.f’
    (Those are [female] artists)

It is worth mentioning that the presence/absence of al in these examples is not related to definiteness since definiteness is encoded by the demonstratives, rather than the al article here. According to
Lyons (1999), nouns introduced with demonstratives belong to what he dubbed complex definites, in which definiteness is encoded indirectly, not by a definite article. Therefore, since definiteness in these expressions is carried by the demonstratives, the presence/absence of *al* functions not as a definiteness carrier, but as phrase vs. sentence marker, respectively. As with the noun adjective expressions, the definiteness status of the noun introduced by a demonstrative is not affected by the presence/absence of *al*.

Another environment where a noun has to be grammatically definite is quantificational sentences. In a quantificational sentence an *al-N*, rather than a bare *N*, must occur after all quantifiers except *kul* ‘every’\(^\text{22}\). This grammatical requirement is enforced regardless of the semantic definiteness status of the noun. Consider the example in (42).

(42) [A teacher discussing feminism and freedom issues in the Arab world with her students in Gender and Feminism class]

\[
\text{fii } \text{*kul* 1-bilaad} \quad 1-%arabiyya/*bilaad}^\text{23} 1-%arabiyya, 1-%u\text{d}^\text{m-}\text{u} \quad 1-\text{banaat}/*\text{banaat} \\
\text{in all the-countries the-Arabic/*countries Arabic, most-Nom} \quad \text{the-girls/*girls} \\
\text{Sindahunna huriyya ʔaqaq min 1-%u\text{d}^\text{m-i} 1-%awlaad}/*\text{awlaad} \\
\text{have.Pl.f freedom less than most-Gen the-boys/*boys} \\
\text{(In all Arab countries, most girls have less freedom than most boys)}
\]

This example has three instances of *al-N* following different quantifiers; two are semantically indefinite and one is semantically definite. The semantically definite is *1-bilaad 1-%arabiyya* ‘Arab countries’ following quantifier *kul* ‘all’. It is definite on both uniqueness and familiarity accounts. On uniqueness account, it refers to a unique set of individual countries which are Arabic in the world situation. Whereas, on familiarity account, the noun phrase refers to a familiar set of objects known to interlocutors either because it was previously mentioned in class as an example of countries where there is some gender discrimination, or from assumed shared knowledge. However, the other two

\[^{22}\text{In MSA, the quantifier *kul* is homonymic; it stands for either ‘every’ or ‘all’. The one meaning ‘every’ allows only a grammatically indefinite singular noun after it, and the one meaning ‘all’ allows only a grammatically definite plural or a mass noun after it.}\]

\[^{23}\text{Using the indefinite *bilaad* is accepted on the other *kul* ‘every’ reading; in this case *bilaad* means ‘a country’, rather than ‘countries’.}\]
nouns *l-banaat* ‘the-girls’ and *l-ʔawlaad* ‘the-boys’ following the same quantifier *muʕð'am* ‘most’ are not definite in the real sense. Neither the speaker, nor the hearers can identify, or are familiar with the referents of these two noun phrases; there might be at least a hundred million boy and girl referents. No one is familiar with, or can identify this huge number of individuals. It is clear also that the referents of these two noun phrases are not unique. Therefore, these noun phrases introduced with *al* are semantically indefinite. This asymmetry of quantifiers allowing only *al*-Ns after them regardless of the semantic definiteness of the NPs indicates that the obligatory presence of *al* in noun phrases following quantifiers is enforced by a grammatical requirement entertained by the grammar of MSA, which does not observe definiteness. Other examples of semantically indefinite *al*-Ns and semantically definite *al*-Ns following the same quantifiers are given in (43&44), respectively.

(43) a. *kull-u l-ʔurduyyiina/*ʔurduyyiina yuḥibb-uuna balada-hum
    all-Nom the-Jordanians/Jordanians 3.love-pl country-their
    (All Jordanians love their country)

b. *baʕd-i-u s-sʔiiniyyiina/*sʔiiniyyiina yatakallam-uuna 1-ʔinkliiztīyya
   some-Nom the-Chinese/Chinese 3.speak-pl the-English
   (Some Chinese speak English)

c. *muʕð'am-u n-nabaataat/*nabaataat taḥtaad-u 1-ʔuxidʒīn
   most-Nom the-plants/plants need-Nom the-oxygen
   (Most plants need oxygen)

d. *fii ʔafriiqyaa, kaθiir min l-ʔatʕfaal/*ʔatʕfaal yuḥramuuna min t-tašliim
   in Africa, many from the-children/children are deprived of the-education
   (Many children in Africa are deprived of education)

(44) a. *kull-u rʕ-tʕullab/*tʕullab fii haadaa sʕ-sʕaff idʕtaaz-u
    all-Nom the-students/students in this the-class passed-Pl
    l-imtihaan n-nihaaʔi the-exam the-final
    (All students in this class passed the final exam)

b. *baʕd-i-u s-sayaarat/*sayaarat allatii iftaraytu-haa kaanat raxiisʔa
   some-Nom the-cars/cars which l.bought-them were cheap
   (Some of the cars I bought were cheap)

c. *muʕð'am-u d-ʔukuur/*ʔukuur fii ʕaʔilat-iwi wuliduu qабla 1-ʔinaat0
   most-Nom the-males/males in family-my were born before the-females
   (Most male [children] in my family were born before female [children])

d. *kaθiir min l-mudʕawhararat/*mudʕawharat fii haadaa sʕ-sʕunduwq
many of the-jewels/*jewels in this the-chest
ʔahdaa-haa lii ʔabii
gave-them to me father-ny
(Many jewels in this chest are gifts from my father)

Additional support against treating *al* as a pure semantic marker by which definites and indefinites are determined in MSA comes from introductory and concluding phrases. The distribution of *al-N* and *N* in these phrases is slightly random and does not relate to definiteness in the real sense.

Some examples are given (45).

(45) a. Obligatory *al-N*:

\[
\begin{align*}
\text{fiī l-waaqiī} & /\text{waaqiī} & \text{‘in the-reality; in reality’} \\
\text{fiī l-haqīqa} & /\text{haqīqa} & \text{‘in the-fact; in fact’} \\
\text{fiī l-ʔasīl} & /\text{ʔasīl} & \text{‘in the-origin; in origin/originally’} \\
\text{fiī l-yaalib} & /\text{yaalib} & \text{‘in the-most; mostly’} \\
\text{fiī l-mudgumal} & /\text{mudgumal} & \text{‘in the-total; in total’} \\
\text{bi-t-tawfiiq} & /\text{t-tawfiiq} & \text{‘with-the-luck; good luck’} \\
\text{maʕ faaʔiq l-ihtiraam} & /\text{ihtiraam} & \text{‘with most the-respect; most respectfully’}
\end{align*}
\]

b. Obligatory *N*:

\[
\begin{align*}
\text{bi-s'araaaha} & /\text{bi-s-s'araaha} & \text{‘with-frankness; frankly’} \\
\text{bi-ʔamaana} & /\text{bi-l-ʔamaana} & \text{‘with-honesty; honestly’} \\
\text{bi-fakil ʕaam} & /\text{l-ʕaam} & \text{‘in-form general; in general’} \\
\text{bi-xtis'aar} & /\text{bi-l-xtis'aar} & \text{‘in-short; in short’} \\
\text{taariixiyyan} & /\text{at-taariixiyyan} & \text{‘historically’} \\
\text{baadi? ʔii bid?} & /\text{l-bid?} & \text{‘start of a starter; to begin with/first of all’}
\end{align*}
\]

As can be noticed the distribution of *al-N* and *N* in these introductory and concluding phrases is irrelevant to semantic definiteness, and the presence/absence of *al* is formulaic or idiomatic. This shows that *al* is no longer used as a pure definite marker in the real sense, but rather other uses, mostly irrelevant to its original function, have developed, which place *al* in a different stage on Greenberg’s (1978) life cycle of a definite marker. This stage is probably the second, where a definite marker no longer functions as an exclusive definite marker, but rather found in definite and indefinite environments, with few uses slightly irrelevant to in/definiteness. However, it has not reached stage three yet, in which a definite marker is completely grammaticalized and functions as a pure nominal or gender marker.
6. Conclusion

In this chapter definiteness, a semantic phenomenon that has strong bearing on the analysis of both reference to a kind genericity and characteristic genericity in MSA, has been investigated. It has been shown that the traditional Arab grammarians’ view of determining simple definites is untenable and problematic. According to this view, noun phrases are classified into semantically definite and semantically indefinite based on the presence/absence of the definite article *al*, respectively. It has been shown that this view is not useful to this thesis, and misses an overall generalization of what is definiteness in the real sense.

Arguments against the traditional view of determining definites and indefinites in MSA are based on facts related to other uses of *al* which do not seem to be relevant to definiteness. The definite article *al* has been shown to felicitously introduce nouns which entertain indefinite interpretation only. In such cases, the *al-N* are interchangeable with their bare *N* counterparts with a slight change in meaning that has nothing to do with the definiteness status of the minimally contrasting noun phrases. Further support for rejecting the traditional morpho-syntactically oriented dichotomy of noun phrases into definites and indefinites comes from uses of *al-N* in indefinite environments, in which *al-N* is not interchangeable with its *N* counterpart, though both noun phrase types entertain an indefinite denotation only. The ungrammaticality of replacing *al-N*, entertaining an indefinite denotation only, with its minimally contrasting *N*, entertaining the same indefinite denotation, is problematic to the traditional view of simple definites. However, this asymmetry can be easily accommodated if the presence of *al* in these cases is not taken as a carrier of definiteness, but rather is enforced by other requirements entertained by the grammar of the language, like sentence vs. phrase distinction, and demonstrative-noun vs. sentence distinction. Some grammatical structures allow only *al-N* to surface for reasons not relevant to definiteness. A final argument raised against the traditional perspective of determining simple definites in MSA comes from expressions in which the distribution of *al-N* and *N* is completely random and formulaic.
In order to account for these asymmetrical uses of the definite article \textit{al} in MSA, this chapter adopted Lyons (1999) major distinction between semantic definiteness and grammatical definiteness. This distinction draws a clear line between what is semantically definite, determined on the definiteness criteria which require the referent of a noun phrase to satisfy at least one of its three conceptual components: familiarity, identifiability, or uniqueness to be definite, and what is grammatically definite. The latter draws upon the presence/absence of the grammaticalization of semantic definiteness, which is susceptible to diachronic change and often develops other uses not related to the function it was first established to fulfil. Therefore, as the case with other grammatical or functional categories, grammatical definiteness might start its life as a pure semantic marker, and ends up a pure syntactic marker, as gender or nominal marker. That said, grammatical definiteness per se cannot be taken as a solid ground for determining definites and indefinites in MSA, as well as many other languages (see Lyons 1999).

It is worth noting that precisely identifying semantically definite and indefinite noun phrases, based on the definiteness criteria, has strong bearing on the analysis of generics in MSA. For instance, some habitual sentences require a semantically indefinite referent for the interpretation of the sentence to be felicitous in MSA (see ex. 37 above); however, MSA has some restrictions on the distribution of noun phrases in terms of grammatical definiteness that has nothing to do with the definiteness status of the referent of the noun phrase semantically. Therefore, according to these grammatical restrictions, only a grammatically definite noun phrase is allowed to surface in certain structures. That said, genericity manifestations examined in NPs and verbless sentences in the remainder of this dissertation will be guided by this understanding of definiteness and indefiniteness based on the definiteness criteria.
CHAPTER 4: ON GENERICITY IN NPS IN MSA

1. Introduction

Assuming the analysis of *al* developed in the previous chapter in which *al-N* does not invariably mark definite noun phrases, characterizing sentences in MSA will be shown to share the semantic structure of characterizing sentences in English, Italian, German, and other studied languages. The semantic structure of a characterizing sentence requires a semantically indefinite subject argument that denotes variable individuals for the sentence to generalize over. This is a crucial condition for a characterizing sentence, as opposed to a particular sentence, which is tied to a particular entity or a group of particular entities (Krifka et al. 1995). If we accept the traditional view of definiteness in MSA - according to which every *al-N* is definite in form and sense - we will end up with two unwelcomed results. The first outcome, which is the more adverse and implausible, is that MSA lacks characterizing sentences as a whole. This result contradicts native speakers’ intuitions who use characterizing sentences widely to express generalizations about individuals out in the world. Moreover, this result is incompatible with speakers’ intuitions about sentences like those in (1) pertaining to the generic interpretation which such and similar sentences express.

(1) a. *tu-hibb-u*  
    Pres.f-love-Nom the-girl-Nom mother-Acc-her  
    (A/The girl loves her mother)

b. *ya-htaah-yu*  
    Pres-need-Nom the-young-Nom to mother-Gen-his  
    (A/The young (baby) needs his mother)

**24** Syntacticians debate on whether or not determiners project full phrases, and are thus the heads of DPs, which take NPs as their complements. The analysis that follows in this project is consistent with treating determiners as specifiers of NPs.

**25** The traditional grammarians’ view dictates also that bare NPs are mostly existential, and hence cannot be used as subject arguments in characterizing sentences, see (section 3) below.
The second less awkward, but still implausible outcome, is that characterizing sentences in MSA are idiosyncratic and language peculiar in that the subject argument NP is tied to a particular individual. This chapter will argue against both of these repercussions, showing that the accurate characterization of the semantic definiteness status of *al-Ns* based on the definiteness criteria elaborated in chapter 3 eliminates these two counterintuitive results. According to the analysis proposed in this chapter, characterizing sentences are pervasive in MSA, and their semantic structure maps to that of characterizing sentences in natural languages. The line between a characterizing sentence and its particular counterpart is drawn based on the correct semantic distinction of *al-Ns* in terms of definiteness, where a semantically definite *al-N* subject is compatible with an episodic reading of the sentence, and a semantically indefinite *al-N* is compatible with a characterizing sentence and a generic reading of the sentence. As discussed in chapter 3, the predicate, the context, and other constituents in a sentence are crucial clues to determine whether an *al-N* subject is semantically definite/indefinite, and hence the exact reading of the sentence, generic/non-generic.

In addition to investigating characterizing sentences incorporating *al-N* subjects in MSA, this chapter will examine the variety of noun phrases in MSA with special focus on nominal forms which are allowed/disallowed in argument positions in generic sentences. Bare, mass, definite, and referential NPs will be scrutinized in episodic and generic environments to see whether the language dedicates any of these nominal forms to formally mark generic or episodic readings, or whether a specific nominal expression is marked as a generic NP or a kind-referring expression.

Another important goal of this chapter is to investigate the preliminary claim of Modern Arab scholars (see Fassi Fehri 2012, chapter 7) that unlike English and other languages, MSA does not allow a bare NP in a sentence that does not express an episodic reading, hence the bare NP being an episodicity marker in MSA. This chapter, however, rejects this claim and shows that in certain syntactic and semantic environments, bare NPs can be used in generic sentences. More strikingly -
contra to the commonsense view of bare NPs in MSA - when a bare NP is felicitously used in a characterizing sentence, it can only receive a non-existentential interpretation. A significant distinction will be made between the referential use of bare NPs when unmodified, and the quantificational use when modified. The latter is compatible with characterizing sentences, while the former seems compatible with episodics only.

The construct state (CS) in MSA, a Semitic peculiar nominal construction, will be examined to explicate its compatibility with generic sentences. The focus of this section will be to investigate whether the indefinite in form CS phrases behave differently from bare simple nouns pertaining to their compatibility with characterizing sentences, and whether the grammatically definite CS phrases behave differently from their simple definite counterparts pertaining to both characterizing sentences and reference to kind sentences. It will be shown that unlike unmodified bare NPs, bare CS phrases are compatible with characterizing sentences. However, grammatically definite CS phrases behave in a way similar to that of al-Ns with regard to both characterizing and particular sentences. In reference to kind sentences, however, grammatically definite CS phrases can be used as kind-denoting nominal expressions, but unlike kind-referring al-Ns, kind-referring CS phrases denote sub-kinds or sub-species, rather than whole kinds or species.

2. The Noun Phrase in MSA: A Syntactic Overview

A noun phrase in MSA consists of a noun and its modifiers. Morphologically speaking, nouns are classified into two categories: derivational and non-derivational. Derivational nouns, termed mas'dar ‘source’ in traditional grammar, are derived from verbs, and are semantically close to the root. The patterns of derived verbs are fixed\textsuperscript{26}. Non-derivational nouns are not derived from verbs. Consider the examples in (2& 3).

(2) Derivational Nouns

\begin{align*}
\textit{kitaab(t)} & \quad \text{‘writing’} \quad \text{As in al-\textit{kitaaba muntis'a} \quad \text{‘Writing is fun’} \\
\textit{mu}s\textit{allim} & \quad \text{‘teacher’} \\
\textit{\$ilm} & \quad \text{‘science/knowledge’}
\end{align*}

\textsuperscript{26} Arabic is highly derivational language that adopts root and pattern morphological model.
mustawdaʕ ‘store’

(3) Non-Derivational Nouns

radjul ‘man’
bint ‘girl’
ʔanf ‘nose’
namir ‘tiger’

The nouns in (2) have corresponding verbs from which they are derived. For the noun kitaaba(t) ‘writing’ there exists the verb katab ‘to write’, from which the noun is derived. Likewise, muʕallim ‘teacher’ is derived from the verb ʕallam ‘to teach’. Similarly, ʕilm ‘science/knowledge’ is derived from ʕalim ‘to know’, and mustawdaʕ ‘store’ from ʔawdaʕ ‘to store/deposit’. It is clear that the verb and its derived noun exhibit some semantic similarity obtained from the meaning of the root they both share. The derivational process yields different fixed noun patterns on par with the varying verb patterns.

The nouns in (3), however, have no verb counterparts. The noun radjul ‘man’, for example, has no corresponding verb. Similarly, bint ‘girl’ is not derived from a corresponding verb; simply because there is no verb relevant to this noun. The same rationale holds for the rest of the nouns in (3). Therefore, these nouns could be considered primitive nominal expressions that belong to a semi-closed, less-productive set. The derivational noun set, on the other hand, is highly productive; whenever a new verb is coined, borrowed, or made up, it is expected that a corresponding verbal noun or masˤdar, as well as other nominal forms, will be automatically derived.

Noun phrases in MSA can be classified according to morphosyntactic and semantic properties that include - among other things - definiteness, number, case, and gender. These properties affect the syntactic distribution and the semantic interpretation of noun phrases. In the following paragraphs, I will briefly address these listed properties in order.
The nouns in MSA, count and mass, are either grammatically definite or indefinite\textsuperscript{27}. A definite noun (4) is overtly marked by the article \textit{al} ‘the’ being prefixed to it. The article and the noun are fused together in one word.

(4) \textit{al-kitaab} ‘the book’
\hspace*{1em} \textit{al-kutub} ‘the books’
\hspace*{1em} \textit{al-džundiyy} ‘the soldier’
\hspace*{1em} \textit{al-maʔ} ‘the water’
\hspace*{1em} \textit{al-maal} ‘the money’

Grammatically indefinite nouns (5), on the other hand, lack an overt D\textsuperscript{28}. Therefore, the absence of the definite article indicates indefiniteness of the noun. It is worth mentioning that indefinite nouns are not overtly marked for indefiniteness, though \textit{-n} suffix, which always appears on indefinite nouns, is taken by some linguists as an indefinite marker.

(5) \textit{kitaab-\textit{-u-n}} ‘a book-case-N’
\hspace*{1em} \textit{kutub-\textit{-u-n}} ‘books-case-N’
\hspace*{1em} \textit{džundiyy-\textit{-u-n}} ‘a soldier-case-N’
\hspace*{1em} \textit{maʔ-\textit{-u-n}} ‘water-case-N’
\hspace*{1em} \textit{maal-\textit{-u-n}} ‘money-case-N’

Nouns in MSA have three numbers - singular, dual, and plural. Dual and plural nouns are explicitly marked, while singular nouns lack an explicit number morphology. The lack of overt number morphology marks singularity. Put differently, grammatical singularity in MSA is taken as the default number that needs no explicit marker, whereas other numbers require overt marking with respect to number. See (6) below for number morphological classification and examples.

(6)

<table>
<thead>
<tr>
<th>Singular masculine</th>
<th>Singular feminine</th>
<th>Dual masculine</th>
<th>Dual feminine</th>
<th>Plural masculine</th>
<th>Plural feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>extit{muslim} 'Muslim'</td>
<td>extit{Muslim-\textit{-a(t)}}</td>
<td>extit{muslim-\textit{-a(aan)}}</td>
<td>extit{muslima-\textit{-i-a(an)}}</td>
<td>extit{muslim-\textit{-uun}}</td>
<td>extit{muslim-\textit{-aat}}</td>
</tr>
<tr>
<td>extit{mușallim} 'teacher'</td>
<td>extit{mușallim-\textit{-a(t)}}</td>
<td>extit{mușallim-\textit{-a(aan)}}</td>
<td>extit{mușallim-\textit{-i-a(an)}}</td>
<td>extit{mușallim-\textit{-uun}}</td>
<td>extit{mușallim-\textit{-aat}}</td>
</tr>
</tbody>
</table>

\textsuperscript{27} Definiteness here refers to grammatical definiteness, marked by the presence/absence of \textit{al} ‘the’, not to semantic definiteness established on the definiteness criteria.

\textsuperscript{28} There are a number of arguments against treating \textit{-n} morpheme - usually termed nunation - as indefinite marker (e.g. Fassi Fehri 1993, 2012; Hallman 1999), so this study treats it as declension marker, and capital N is used to stand for nunation, (see section 7).
As is clear from the data in (6), singular nouns in MSA are not overtly marked for number. Dual nouns, however, are marked using the suffix –ään, which is used invariably with masculine and feminine nouns. Plural nouns are more complicated than what is introduced in the data; all the example nouns belong to the traditionally termed *sound plural nouns*. This class employs the suffixes –uun and –aat to mark plural number on masculine and feminine nouns, respectively. It is worth mentioning that the quality of the long vowel in –uun suffix varies from uu, aa, to ii according to the grammatical status of the noun; in particular the case assigned to it, where the three vowels mark nominative, accusative, and genitive case, respectively. Unlike sound plurals, which are marked with suffixal endings, the other class of plurals (7), termed *broken plural* in traditional grammar, is expressed with a range of roughly thirty different forms of varying degrees of productivity and semantic specificity (Hoyt 2008). Mass nouns take two plural forms: the collective noun (8), usually expressed as a broken plural, and the paucal plural (9), which is usually a sound plural, and denotes the plural of a singulative object of relevant mass noun.

(7)  

<table>
<thead>
<tr>
<th>Singular Noun</th>
<th>Plural Noun</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>kitaab</td>
<td>kutub</td>
<td>‘book’</td>
</tr>
<tr>
<td>fayʔ</td>
<td>??fyaaʔ</td>
<td>‘thing’</td>
</tr>
<tr>
<td>sitara(t)</td>
<td>sataaʔr</td>
<td>‘curtain’</td>
</tr>
<tr>
<td>qanniina(t)</td>
<td>qanaani</td>
<td>‘bottle’</td>
</tr>
</tbody>
</table>

(8)  

<table>
<thead>
<tr>
<th>Singular Noun</th>
<th>Plural Noun</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>yanam</td>
<td>yanam-aat</td>
<td>‘sheep-countable’</td>
</tr>
<tr>
<td>samak</td>
<td>samak-aat</td>
<td>‘fish-countable’</td>
</tr>
<tr>
<td>baqar</td>
<td>baqar-aat</td>
<td>‘cows-countable’</td>
</tr>
<tr>
<td>fadjar</td>
<td>fadjar-aat</td>
<td>‘trees-countable’</td>
</tr>
</tbody>
</table>

(9)  

<table>
<thead>
<tr>
<th>Singular Noun</th>
<th>Plural Noun</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>yanama(t)</td>
<td>yanam-aat</td>
<td>‘sheep-countable’</td>
</tr>
<tr>
<td>samaka(t)</td>
<td>samak-aat</td>
<td>‘fish-countable’</td>
</tr>
<tr>
<td>baqara(t)</td>
<td>baqar-aat</td>
<td>‘cows-countable’</td>
</tr>
<tr>
<td>fadżara(t)</td>
<td>fadżar-aat</td>
<td>‘trees-countable’</td>
</tr>
</tbody>
</table>

Among all the descendants of Classical Arabic, MSA is exclusively marked for case. Other Arabic dialects have lost this morphology and employed word order instead for theta-role assignment. In MSA, declinable nouns are assigned nominative, accusative, or genitive case based on their
syntactic position in the sentence. These cases are demonstrated by –u, -a, or -i suffixes\textsuperscript{29}, respectively. Consider the data in (10).

(10) a. ʤaaʔ-a \hspace{1cm} l-walad-u
came-3 \hspace{1cm} the-boy-Nom
(The boy came)

b. raʔay-tu \hspace{1cm} l-walad-a
saw-1 \hspace{1cm} the-boy-Acc
(I saw the boy)

c. marar-naa \hspace{1cm} bi-l-walad-i
passed-1.Pl \hspace{1cm} by-the-boy-Gen
(We passed by the boy)

Gender morphology appears on certain classes of words, but it plays a major role in agreement marking and other concordial relations (Hoyt 2008). Grammatical gender in MSA relates to a variety of semantic categories like biological sex (11.a), individuation of non-human mass nouns (11.b), and inanimate mass plurality (11.c). Feminine gender is realized by the bound –at\textsuperscript{30} suffix, where its absence indicates masculinity. In other words, masculine gender is morphologically unmarked on singular nouns. Other feminine nouns are not morphologically marked for gender, but they syntactically and semantically behave as feminine, either because they denote biologically feminine individuals (11.d), or by convention (11.e). As for plural forms, only sound plural shows gender morphology (11.f). The data below illustrate these syntactic facts.

(11) a. xaadim \hspace{1cm} ‘a male servant’ \hspace{1cm} xaadim-(t) \hspace{1cm} ‘a female servant’
b. samak \hspace{1cm} ‘fish-mass’ \hspace{1cm} samak-(t) \hspace{1cm} ‘an individual fish’
c. hidqaar-(t) \hspace{1cm} ‘stones’ \hspace{1cm} nifaar-(t) \hspace{1cm} ‘mulch/sawdust’
d. bint \hspace{1cm} ‘girl’ \hspace{1cm} ʕaruus \hspace{1cm} ‘bride’
e. ʃams \hspace{1cm} ‘sun’ \hspace{1cm} ʃarb \hspace{1cm} ‘war’
f. mudarris-uun \hspace{1cm} ‘male instructors’ \hspace{1cm} mudarris-(t) \hspace{1cm} ‘female instructors’

In general, all modifiers; simple adjectives, adjective phrases, relative clauses, and prepositional phrases; follow the modified noun in a noun phrases. Consider the examples below.

\textsuperscript{29} These vowels are lengthened when used to mark case on dual and plural nouns.

\textsuperscript{30} The phonological status of \textit{t} in –at suffix depends on whether there is a pause or not following the noun which carries it. More accurately, this consonant is expressed in continuous speech, and is null at pausing, hence written in brackets to mark its phonological optionality.
(12)  
\[af\text{-}faix-u \quad l\text{-}qa\text{-}si\text{ir}\]
the-old man-Nom the-short
(mid-old man)

\[al\text{-}kalb-u \quad lla\text{ðii} \quad y\text{-}anbah\]
the-dog-Nom which 3-bark
(mid-dog which is barking)

\[balad-u\text{-}n \quad fii \quad \dot{\text{f}}\text{riiqya}\]
a country-Nom-N in Africa
(a country in Africa)

However, relative adjectives and ordinal numerals can precede the nominal modifier in the so-called
construct state, (see section 6).

(13)  
\[\ddot{\text{ad}}\text{'xam-u} \quad \text{huut}\]
largest-Nom whale
(mid-largest whale)

\[\ddot{\text{at}}\text{'wal-u} \quad \text{nahr}\]
longest-Nom river
(mid-longest river)

\[\ddot{\text{awwal-u} \quad fahiid}\]
first-Nom martyr
(mid-first martyr)

\[\text{saadis-u} \quad r\text{-}ru\text{'asaa}\]
sixth-Nom the-presidents
(mid-sixth president)

It is worth mentioning that the head noun and the attributive adjective show full agreement in
case, number, gender, and definiteness; whereas a predicative adjective concords with the noun in
gender and number only. This is a syntactic benchmark distinguishing a nominal phrase from a
clause.

(14) a.  
\[at\text{-}\text{t'aalib-}\text{at-u} \quad l\text{-}muhtaar-}\text{at-u}\]
the-student-f-Nom the-confused-f-Nom
(mid-confused female student)

b.  
\[\text{kaan-at} \quad at\text{-}\text{t'aalib-}\text{at-u} \quad \text{muhtaar-}\text{at-a-n}\]
was-f the-student-f-Nom confused-f-Acc-N
(mid-The female student was confused)
While the attributive adjective in (a) agrees with its modified noun in case, gender, number, and definiteness, the predicative adjective in (b) agrees with the nominal subject in gender and number only.

3. **Genericity and the Nominal Phrase Distinction in MSA**

Krifka et al. (1995) emphasize that cross-linguistically, generics can be manifested by a variety of different syntactic structures. Generic sentences in English, for instance, can be expressed using an array of syntactically different NPs: indefinite singular NPs (a), bare plural NPs (b), definite NPs (c), mass noun NPs (d), and proper names (e).

(15) a. A dog barks.
    b. Dogs bark.
    c. The dog barks.
    d. Gold is a precious metal.
    e. Jill drives to school.

Despite their syntactic distinctions with respect to the nominal phrase type used in subject argument, each sentence in (15) is clearly generic; no sentence reports an isolated episode about a particular NP referent. All the sentences can be rephrased using the adverb *in general*, for example, with slight change in meaning in the rephrasing sentence.

MSA shares this crosslinguistic feature in that a variety of nominal expressions can be incorporated in generic sentences. However, the traditional view of definiteness indicates that only grammatically definite nouns or nouns occurring in maximal definite phrases can appear in generics because grammatically indefinite nouns force existential interpretation only (Fassi Fehri 2012). This is illustrated in (16-21):

(16) Definite singular subject

\[
\text{al-sarabiyy}-u \quad \text{kariim}
\]

the-Arab-Nom generous

(The Arab is generous)

(17) Definite plural subject

\[
\text{al-sarab}-u \quad \text{kuramaa}?
\]

31 This claim is untenable as bare NPs can be used in generics; see below section (7.2).
the-Arabs-Nom generous.Pl
(Arabs are generous)

(18) Mass noun subject

<table>
<thead>
<tr>
<th>al-maaʔ-u</th>
<th>ʔsaas-u</th>
<th>l-hayaa</th>
</tr>
</thead>
<tbody>
<tr>
<td>the-water-Nom</td>
<td>basis-Nom</td>
<td>the-life</td>
</tr>
</tbody>
</table>
(Water is the basis of life)

(19) Proper noun subject

<table>
<thead>
<tr>
<th>y-amsii</th>
<th>ᵈaliyy-ʔ-u-n</th>
<th>li-l-madrasa</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-walk Ali-Nom-N</td>
<td>to-the-school</td>
<td></td>
</tr>
</tbody>
</table>
(Ali walks to school)

(20) Bare singular occurring in definite construct state (*idˤafa construction*)

<table>
<thead>
<tr>
<th>y-anbah-u</th>
<th>kalb-u</th>
<th>l-dжаiraan-i</th>
<th>laylan</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-bark-Nom</td>
<td>dog-Nom</td>
<td>the-neighbors-Gen</td>
<td>at night</td>
</tr>
</tbody>
</table>
(The neighbors’ dog barks at night)

(21) Bare plural occurring in definite construct state (*iˤdaafa construction*)

<table>
<thead>
<tr>
<th>y-atahaddaʔ-u</th>
<th>ʔawlaad-u</th>
<th>ʕamm-ii</th>
<th>bi-sˤawt-i-n</th>
<th>xaafit</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-speak-Nom</td>
<td>sons-Nom</td>
<td>uncle-my</td>
<td>in-voice-Gen-N</td>
<td>soft</td>
</tr>
</tbody>
</table>
(My uncle’s sons speak softly)

All the sentences in (16-21) are generic in that they express a generalization of some kind. For example, sentences (16&17) do not talk about particular Arabs, but rather about Arabs in general. Similarly, (18) does not relate a property to a particular amount of water, but it gives a property of water in general, (19) expresses a fact about how Ali goes to school in general, and this kind of interpretation holds for (20&21).

It is worth noticing that the same kind of NP subjects can occur in episodic sentences as well (22); a fact that needs a lot of investigation in order to tease apart generic/no-generic readings of a sentence incorporating a grammatically definite NP. Put differently, what avails/blocks a generic/non-generic reading of a sentence incorporating a definite NP subject? This question is addressed in section 4 below.

(22) a. haaḍa | l-ʕarabiyy-ʔ-u | kariim |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>this</td>
<td>the-Arab-Nom</td>
<td>generous</td>
</tr>
</tbody>
</table>
(This Arab is generous)
However, traditional grammarians and even modern Arab linguists (see, Fassi Fehri 2012, chapter 7) claim that unlike English and other languages, MSA allows both indefinite singular and bare plural subjects in existential sentences only. This is because these NP forms only express existential denotation that semantically contradicts genericity.  

(23) y-alsab-u walad-u-n ?amaama bayt-i-hi  
3-play-Nom boy-Nom-N in front of house-Gen-his  
(A boy is playing on the street)  
(*generic /episodic interpretation)  

(24) y-alsab-u ?awalaad-u-n ?amaama bayt-i-him  
3-play-nom boys-nom-N in front of house-Gen-their  
(Boys are playing on the street)  
(*generic /episodic interpretation)  

The only interpretation that (23&24) receive is episodic. The two sentences report accidental playing episodes true of existentially calculated boy individuals at a certain time and place. It is noteworthy that these two sentences can be rephrased by combining the existential expression θammata ‘there is/are’ with no change in meaning in the resulting sentence. These two sentences are incompatible with time reference adverbs like haaðihi l-?ayyaam ‘these days’, which enforce a non-existential denotation of the bare NP subjects in (23&24), and promote a habitual reading of the sentences, (see chapter 5).  

32 It will be shown that this claim is inaccurate and needs refinement, (see section 6).
Therefore, it seems that MSA, according to the traditional grammarians’ view, allows only definite nouns tied to particular individuals to be used in generics, and the presence of indefinite nouns blocks the generic reading of a sentence and avails a non-generic reading only. This analysis, though looks simple and straightforward, is problematic as it shows that MSA either lacks characterizing sentences - a cross-linguistic semantic phenomenon, or that MSA is peculiar in its representation of characterizing sentences. This analysis sets the characterizing sentence structure in MSA apart from the semantic representation of characterizing sentences in world languages. Contra to genericity phenomenon in world languages that requires the NP in a characterizing sentence about which a generalization is made be semantically indefinite, MSA - according to the traditional view of definiteness which invariably classifies an al-N as definite in form and sense - requires that the NP must be definite to be incorporated in a characterizing sentence. This asymmetry is investigated thoroughly in section (4) below, where arguments are made against both asymmetrical options.

4. Genericity and the Definite NP in MSA

The variety of noun phrases that can be used in generic sentences in MSA, characterizing sentences in particular, is problematic to genericity investigation. According to the syntactic facts in section (3), only definite nouns can be incorporated in sentences expressing a generic reading, and indefinite nouns can only be inserted in sentences entertaining a non-generic reading. In order to see how problematic this asymmetry is to the semantic representation of characterizing sentences, I will brief the traditional view of the interaction between NPs and characterizing sentences in MSA in (4.1), and then discuss the semantic essence of a characterizing sentence in (4.2). In (4.3), I will argue against the traditional view introduced in (4.1), showing that the semantic representation of characterizing sentences in MSA conforms to the semantic representation of characterizing sentences introduced in (4.2).

4.1 Characterizing Sentences in MSA: The Traditional View

Although genericity has received very little attention in modern Arabic linguistics, the commonsense viewpoint is that in general, the different interpretations of NP as generic or existential
depend mainly on the overt presence/absence of the definite article *al* ‘the’. The overt presence of *al* avails generic and existential readings of the NP, in the sense that *al*-*N* is either kind-denoting expression, generic NP, or object denoting expression, existential NP (Fassi Fehri 2012). The lack of *al*, however, forces an existential interpretation only. This view is on par with the traditional view of simple definites in MSA, which was shown problematic in chapter three. According to this view, NPs are dichotomized as definite or indefinite based on the overt presence/absence of *al* determiner, respectively. Fassi Fehri (2012:179) states, “Arabic NP/DP interpretations as generic (Gen) or (only) existential (Ex) depend on whether they express overtly the definite determiner or article D, or whether they lack such an overt expression of D.” Notice the characterization of the denotation of NPs, not sentences, as generic or existential. The contrast is illustrated in (25) and (26), respectively:

(25) a. *y-aʔkul-u*  
3-eat-Nom

*ʔasad-u*  
the-lion-Nom

*l-lahm*  
the-meat

(The lion eats meat)

b. *t-aʔkul-u*  
3.f-eat-Nom

*ʔusuud-u*  
the-lions-Nom

*l-lahm*  
the-meat

(Lions eat meat)

(26) a. *y-aʔkul-u*  
3-eat-Nom

*asad-u*-n  
lion-Nom-N

*l-lahm*  
the-meat

(A lion is eating meat)

b. *t-aʔkul-u*  
3.f-eat-Nom

*ʔusuud-u*-n  
lions-Nom-N

*l-lahm*  
the-meat

(Lions are eating meat)

In (25) both the generic reading, *the kind Lion*; and the existential reading, *some contextually salient lion(s)*, are available. In the generic reading, the property of *eating meat* is predicated to the Lion kind. The existential reading of (25) focuses on uniquely salient lions in the domain of discourse that are involved in an episodic eating event at the time of utterance. In (26), however, only an existential interpretation of the NPs, and hence the sentences as a whole, are available. The only interpretation that can be gleaned is a report of accidental eating event of some lion(s) eating meat at a particular time and location. Using a definite NP as a kind-referring expression is felicitous as a kind is semantically definite.
The same contrast holds with respect to mass nouns. The illustration is in (27) and (28). In the examples, nouns with definite *al* obtain generic- kind-referring-, and existential- object-referring, denotations; whereas, bare nouns which lack overt *al* can only receive existential interpretation. It is noteworthy that number and countability have no bearing on this distinction as is clear from (25&26) and (27&28), respectively.

(27) a. *laa n-astattiiʕ-u l-Say f-a biduuni l-maaʔ*
    NEG 1.Pl-can-Nom the-living-Acc without the-water
    (We cannot live without water)

b. *al-haliib-u yiðaaʔ-u-n sʔiḥhiy*
    the-milk-Nom food-Nom-N healthy
    (Milk is a healthy food)

(28) a. *ʔ-uriid-u maaʔ-a-n radgaaʔ-an*
    1-want-Nom water-Acc-N please
    (I want water, please!)

b. *faribnaa haliib-a-n ?ams*
    drank.3.pl milk-Acc-N yesterday
    (We drank milk yesterday)

Based on the data above, it seems reasonable to generalize, according to the traditional view, that the presence of an indefinite NP blocks the generic reading of a sentence and promotes an episodic reading as the only available interpretation. The presence of a definite NP, on the other hand, gives rise to an ambiguous sentence with two competing interpretations: generic, where the definite NP is kind-denoting, and existential, in the sense that the definite NP refers to one unique entity or entities salient in the domain of discourse. It should be noticed that the traditional way of determining in/definites is purely morphological. However, Fassi Fehri (2012) refines this view about the existential reading of bare nouns claiming that some bare nouns can be used only in habituals and when-clauses, which he calls modalized contexts, as they provide what he calls sentence binding of GEN, provided by the event structure. This is evident from identifying generic definite NPs as denoting kinds only; whereas, a generic object can be expressed by an indefinite (in form) NP in habitual and in when-clauses. In his words, “[s]ummarizing the contextual and interpretational requirements discussed so far, the following descriptive statements appear to apply to Definite and
Indefinite Generics found in Arabic: Kind denoting Gen are expressible only through overtly definite DPs, and are likely to be referential arguments (kind names). Object-denoting Gen are expressible through various sorts of indefinite DPs, bound (unselectively) by Gen operators; they are necessarily quantificational arguments. Gen operators (providing for characterizing environments) are: Habitual Aspect, Q-Adverbs, and appropriate Ps.” (P.200). I will, however, set aside the claim about the indefinite NP to be discussed in section 7 below, and focus on the one relevant to the definite NP for the rest of this section

The traditional view of determining NPs allowed in generic sentences is problematic. On the one hand, it only admits definite NPs in generic sentences, excluding bare NPs from being inserted in generic sentences except in habituals and modalized contexts where generalizing is over events or situations, rather than individuals. On the other hand, due to the received traditional viewpoint that al-N is semantically definite, the only denotation that a definite NP can receive in a generic sentence is kind denotation. This gives the inaccurate impression that either a characterizing sentence is not available in MSA, or that contra to the characterizing sentence semantic representation in natural languages, MSA is peculiar in realizing a characterizing sentence using a semantically definite NP. In the following section, I will briefly discuss the type of generalization a characterizing sentence expresses, and the semantic requirements that need to be satisfied by an NP functioning as its subject to be allowed in a characterizing sentence. This investigation will show clearly that a characterizing sentence is infelicitous with a semantically definite NP; a fact that will be used to argue against MSA’s peculiarity of realizing a characterizing sentence using a definite NP in subject argument in section (4.3). In addition, it will be shown in that section that MSA entertains characteristic genericity phenomenon, and manifests it semantically in accordance to the universal semantic representation of characteristic genericity.

4.2 A Characterizing Sentence: What Kind of Generalization is Expressed?

By definition, a characterizing sentence reports a generalization over patterns of regularities that summarize groups of non-accidental facts, episodes, or state of affairs. Unlike a particular
sentence, which expresses a statement about “particular events, properties of particular objects, and
the like” (Krifka et. al 1995:3), a characterizing sentence expresses a generalization based on
properties not tied to a particular object or event, but rather to any object or event which satisfies the
descriptive content of the NP or VP used. Consider the difference between characterizing sentences
(29) and their particular counterparts (30)

(29) a. A graduate student is diligent.
   b. A dog hates cats.
   c. An Italian loves pizza.

(30) a. The graduate student is diligent.
   b. The dog hates cats.
   c. The Italian loves pizza.

The sentences in (29) are true generics; they each report a generalization of some kind. In (a), being
diligent is reported as a property of graduate students in general, not about a particular, specific
graduate student. In (b), hating cats is reported as a property of dogs in general. Likewise, the
sentence in (c) states something about Italians in general, not a particular Italian. However, the
sentences in (30) are opposite to those in (29) in that they state isolated facts about particular objects.
This distinction is triggered by the different definite/indefinite NPs used in (29) and (30) respectively.
A semantically indefinite NP can trigger a generic reading because indefiniteness assumes a
descriptive characterization true of a variable or place holder that can be fleshed out by any individual
satisfying the descriptive content of the noun. This is a crucial condition for a sentence reporting a
generalization over entities or events. If the referent of the NP is definite in the real sense, a
generalization cannot be made. Krifka et al. (1995: 32) explicitly state this condition as a crucial
benchmark distinguishing a characterizing sentence from its particular counterpart. They argue that
“… a generalization expresses that if an entity satisfies certain conditions (or has certain properties)
A, then it also satisfies certain conditions (or has certain properties) B… [c]haracterizing sentences
must have at least one variable which is not explicitly tied to some particular object. If this were not

---

33 The definite NPs in (29) are non- kind-referring; they refer to definite referents salient in the context.
the case, they would merely state that a particular object (as described by the restrictor) has a certain property (as described by the matrix, and they could no longer express a ‘generic’ fact.”

This condition applies also to habitual sentences where a generalization is made about events or situations. If the denotation of a predicate is tied to a particular event, rather than to indefinite variable events, the sentence can only receive an episodic interpretation. Consider the habituals in (31), and their contrasting episodics in (32).

(31) a. Jill studies hard for finals.
    b. I eat falafel for breakfast.
    c. Bill gets scared when he sees a snake.

(32) a. Jill is studying hard for finals.
    b. I am eating falafel for breakfast.
    c. Bill got scared when he saw a snake yesterday.

It is worth noticing that only the episodic sentences in (32) are compatible with time reference adverbials like now and today which limit the denotation of the predicate to a particular, existential event. Adding such adverbials to the sentences in (31) renders them unacceptable.

It seems that a characterizing sentence requires a semantically indefinite entity or event to quantify and generalize over. A semantically definite entity or event blocks the generic reading of a sentence, and triggers an episodic reading instead. This crucial semantic condition of a characterizing sentence puts the traditional claim that in MSA only definite NPs can be used in generic sentences, and that the use of an indefinite NP is limited to existential sentences on the line. On one hand, it sets characteristic genericity in MSA from its canonical realization in natural languages. On the other hand, it implies that in MSA only a reference to a kind generic can be expressed; whereas, characterizing sentences with object denoting subject arguments are non-existent in MSA. The latter inference is untenable as speakers of MSA use characterizing sentences widely to express generalizations about numerous regularities in the world. Therefore, this outcome is void, and does not need further investigation. The other outcome that MSA is peculiar in its semantic mapping of a characterization sentence is interesting and is investigated below.
4.3 Characteristic Genericity in MSA: Is it Language Peculiar?

In this section, I will argue against the traditional view that only definite NPs with kind denotation are allowed in sentences entertaining a generic reading. My claim is established upon the findings of chapter three, where a major, significant distinction is made between grammatical definiteness and semantic definiteness. Based on this distinction, I argue that the traditional view is confusing these two overlapping phenomena. I agree, to a certain extent, with the traditional view that in general characterizing sentences in MSA are expressed with al-Ns, but disagree with its characterization of definites and indefinites in MSA based on the presence of the definite article al. My claim is that although characterizing sentences in MSA allow al-Ns, these NPs - contras to the traditional view – are semantically indefinite. If the incorporated NP is semantically definite, however, the only reading available is either reference to a kind, or episodic. Therefore, if we put the problematic morphosyntactic way of determining in/definites away, and adopt the semantic criteria of determining definites proposed in chapter 3, this implausible semantic mapping of characterizing sentences in MSA will evaporate, and everything falls in place. This view is on par with the canonical representation of characterizing sentences, which requires that the NP upon which a generalization is made must not be tied to a particular, semantically definite object.

As argued in chapter 3, definiteness, as a semantic phenomenon, is established on at least one of the concepts of the definiteness criteria: familiarity, identifiability, and uniqueness. Although the definite article is taken by many semanticists as the canonical definite determiner which coalesces with a common noun and contributes either uniqueness or familiarity, this grammatical ingredient cannot be taken as a necessary and sufficient condition for determining simple in/definites in MSA (see section 5 of chapter 3).

34 I will set reference to a kind genericity aside, to be discussed in (5) below, and focus on characterizing sentences in this section.
The following section focuses on the canonical definite NP, *al-N* in MSA, and tries to scrutinize its syntactic and semantic properties, and how it fits in and contributes to generic sentences, characterizing sentences in particular.

### 4.3.1 Characterizing Sentences and the Definite NP with *al* Determiner

The basis of my claim that characterizing sentences in MSA conform to the universal semantic structure of characterizing sentences is that correctly identifying *al-N* as semantically definite or indefinite would resolve the asymmetry created by the traditional view’s implausible outcome that characterizing sentences are idiosyncratic, different from the canonical semantic structure of a characterizing sentence. Put differently, we need to investigate the inter-sentence interaction between *al-N* and other constituents in the sentence— the predicate in particular, and the intra-sentence interaction between the sentence and its pragmatic discourse; i.e., the sentence’s semantic and pragmatic sides. By doing so, the goal is to accurately determine the *al-N* used in a sentence as semantically definite, hence identifying the sentence as a particular sentence, or semantically indefinite, hence identifying the sentence as a characterizing sentence. Both the predicate and the discourse have strong bearing on determining the *al-N* as semantically indefinite/definite, and hence interpreting the sentence as a whole as generic/non-generic, respectively.

An interesting question to be addressed here is whether MSA formally encodes the two competing meanings of the grammatically definite NPs, generic\(^{35}\)— quantificational, and existential—referential, and whether the type of predicate incorporated has some bearing on the ultimate interpretation of the sentence, generic or non-generic. Predicates, according to Carlson 1977, are classified into levels based on the type of properties they denote. Carlson distinguishes between three types of predicates, which he dubbed Stage-level predicates, Individual-level predicates, and Kind-level predicates, abbreviated as S-level, I-level, and K-level predicates, respectively. Since Carlson’s

---

\(^{35}\)Generic here means that the noun is semantically indefinite—denoting a variable individual to which the descriptive content of the noun applies— and can be used in a characterizing sentence; it does not indicate that it is a kind-referring NP.
proposal, the syntactic and semantic ramifications of this distinction have been a lively topic among semanticists (e.g. Asher 2010; Chierchia 1995; Diesing 1988, 1992; Kratzer 1995). S-level predicates denote properties that are ephemeral, transient, episodic, and mostly spatio-temporally bounded like sleep, be in the bathroom, scratch, etc. I- level predicates, however, are construed to express properties of individuals that are permanent and stable in a sense. As Carlson (1977) argues, there are three basic constructions of I-level predicates:

(33) a. Stative VPs, like hate, know, love, believe, own, etc.
    b. Predicative NPs, like be a student, be an animal, be a woman, be mammals, etc.
    c. Adjectives like smart, tall, white, transparent, etc. (vs. angry, sad, available, busy, etc.)

Kind-level predicates denote properties that apply to kinds, species, or ‘genus’ only, not any instance or groups of instances of a kind. Examples of K-level predicates are not as many as S-level or I-level predicates, but they can be easily cited; widespread, extinct, die out, common, rare are canonical examples.

The differentiation made in Carlson’s (1977), and maybe before, between S-level and I-level predicates contributes significantly to the explanation of some syntactic asymmetries found in structures incorporating these two predicate types. Many semanticists (e.g. Carlson 1997; Kratzer 1995; Milsark 1974, 1977) have observed that S-level predicates are compatible with some structures, like temporal adverbials, locatives, perception sentences and there-sentences; whereas, I-level predicates are not. In addition, Carlson (1977) argues that bare plurals in English can have two different interpretations, existential and generic. The type of interpretation depends on the associated predicate; I-level predicates, he asserts, give rise to a generic interpretation, but S-level predicates give rise to an existential interpretation. We will investigate the effect of the predicate’s level on the ultimate interpretation of sentences combining grammatically definite NPs functioning as arguments in MSA, and whether either S-level or I-level predicates trigger, block, or give rise to generic/non-generic interpretation of the sentence. These points are furnished by the following sets of data in (34&35).
The sentences in (34) can only receive existential or episodic interpretations, in which the NPs used as subject arguments denote unique individuals salient in the domain of context. Scrutiny of the predicates used in the sentences reveals that all of these predicates are S-level predicates which predicate properties true of spatiotemporal slices of individuals, as Carlson (1977) meta-cognitively defines stages of individuals. Notice that the predicates are compatible with temporal adverbials; *munḍu saaṭat-i-n ‘since an hour’, *haaḍḍaa sˤ-sˤabaaḥ ‘this morning’, *ʔaaan ‘now’, *qabla saaḍa ‘an hour ago’; locatives, *fii l-matˤbax ‘in the kitchen’, *fii sˤ-sˤaalat-i ‘in the living room’; and perception verbs like *ʔunðˤr ‘look’. These are tests employed in the literature to identify predicates denoting stage-level properties. As well established in genericity literature, generic sentences do not express spatiotemporally bounded propositions; the propositions expressed, however, are taken to be true with no restrictions on time, place, or number of individuals. Thus, it seems that the presence of S-level predicates blocks genericity, and gives rise to an existential interpretation of the definite NP, and hence the whole sentence. This generalization conforms to the generalization made by Carlson (1977) regarding the interpretation of bare plurals, bare mass, and the definite singular in English when are associated with S-level predicates.
I turn now to consider another set of data, and see the effect of incorporating I-level predicates on the interpretation of the grammatically definite NP and the sentence as a whole.

Consider the data in (35).

(35) a. ar-radžul-u ʃ-farqiy-yu  yayuuru
    the-man-Nom the-eastern-Nom jealous
    (The/A Middle Eastern man is jealous)

b. y-aʃrif-u l-muʃallim-u masʔušiiɣgwat-i-zi
    3-know-Nom the-teacher-Nom responsibilities-Acc-his
    (The/A teacher knows his responsibilities)

c. al-kaʃl-bu ʤakiiyy
    the-dog-Nom intelligent
    (The/A dog is intelligent)

d. t-ʔiḥab-u l-bint-u ʔumma-haa
    3.f-love-Nom the-girl-Nom mother-Acc-her
    (The/A girl loves her mother)

e. al-malika-t-zi mayruu-ra
    the-queen-f-Nom arrogant
    (The/A queen is arrogant)

f. t-aʃiif-u l-bint-u maʃa ʃaaʔi-ta-zi-haa fi l-ʔurduun
    3.f-live-Nom the-girl-Nom with family-f-Gen-her in the-Jordan
    (The/A (unmarried) girl lives with her family in Jordan)

The sentences in (35) are ambiguous between generic/nongeneric readings, though the generic reading is naturally more salient. A quick look at the sentences shows that the predicates used are all I-level predicates; note that inserting a spatiotemporal expression or a perception verb in any of the sentences renders it unacceptable, see (36) for the effect of such insertion on some of the sentences in (35).

(36) a*ar-radžul-u ʃ-farqiy-yu  yayuuru-n fi l-maʃbat/l-ʔawm
    the-man-Nom the-eastern-Nom jealous-Nom-N in the-kitchen/today
    (The/A Middle Eastern man is jealous in the kitchen/today)

b.*al-kaʃl-bu ʤakiiyyu-n fi l-ʔadiwa/ʔa-n
    the-dog-Nom intelligent-Nom-N in the-park/now
    (The dog is intelligent in the park/now)

c.*ʔuŋi-ʔa ʒaa buna, t-ʔiḥab-u l-bint-u ʔumma-haa
    look O sonny, 3.f-love-Nom the-girl-Nom mother-Acc-her
    (Look O sonny, the/a girl loves her mother)
This indicates that, in sentences uttered out of the blue, associating I-level predicates with *al-*N*$_s$ triggers a generic reading of sentences, without blocking the possibility that such sentences can be interpreted existentially depending on the context. The sentence in (35.a), for instance, entertains a generic interpretation where the Middle Eastern men, in general, has the property of being jealous, this reading can be true when uttered out of the blue with little or no contextual support, in the sense that this sentence takes the whole world as its situation. Another episodic interpretation is available; according to which a unique, context-dependent Middle Eastern man is reported to hold the property of being jealous. The same ambiguous reading holds for the other sentences in (35). Therefore, it seems that in pure semantic analysis, such construction promotes the generic reading, and whatever weakens this reading comes from the pragmatic side of the sentence, the context in particular, which helps determining the semantic in/definiteness status of the *al-*N*$_s$ used, and hence the interpretation of the sentence as a whole.

A crucial question about the source of ambiguity in sentences associating definite NPs and I-level predicates poses itself here. More explicitly, what is the source of this ambiguity? A straightforward answer is the denotation of the definite NP; whether it denotes a semantically definite or indefinite referent. The former forces an episodic interpretation, while the latter avails a generic interpretation. Notice that a generic sentence reports a generalization over a group of individuals, events, state of affairs, or facts, not a particular instance of these. Therefore, in order to glean the exact interpretation of a sentence incorporating a definite NP and an I-level predicate, we need to investigate the exact denotation of the definite NP used. In other words, is this definite NP semantically definite, denoting a unique, identifiable, or familiar referent, or the reversed, in this case denoting an indefinite referent. This analysis recalls the crucial distinction made in chapter three between grammatically definite NP and semantically definite NP. This distinction has strong bearing on our interpretation of sentences with definite NPs as generic or non-generic. The context provided plays a major role in the determination of a grammatically definite NP as definite or indefinite in the real sense. To see this, I will sharpen the intended interpretation of the sentences by providing
contexts for the data in (35) that limit the denotation of the NPs to indefinite referents, and see the kind of interpretation these sentences avail. Later, I will provide different contexts to the same data that limit the denotation of the NPs to semantically definite referents, and see which reading, generic/non-generic, the sentences provide. Consider the data in (35) repeated in (37) with indefinite environments.

(37) a. [A Jordanian student at an American school talking to her Jordanian roommate about her future husband, and that he must be from the Middle East. The other friend wants to know why her friend is so strict in this regard, though a woman can find a good husband from outside the Middle East. The student answers:]

\[
\text{ar-radqul-u} \quad f\text{-}f\text{arqiyy-u} \quad y\text{ayuur} \\
\text{the-man-Nom} \quad \text{the-eastern-Nom} \quad \text{jealous}
\]

(A Middle Eastern man is jealous)

b. [In a press conference on the Teacher Day in Jordan, The minister of education expresses his trust in teachers, and that they know their job duties and responsibilities]

\[
y\text{-}a\text{srif-u} \quad l\text{-}\text{mu\text{c}allim-u} \quad mop\?u\text{liyyaat-i-hi} \\
3\text{-}\text{know-Nom} \quad \text{the-teacher-Nom} \quad \text{responsibilities-Acc-his}
\]

(A teacher knows his responsibilities)

c. [A father is asking his little daughter which pet she wants as a birthday present and why. The girl says she wants an intelligent animal, and]

\[
\text{al-kalb-u} \quad \delta\text{akiyy} \\
\text{the-dog-Nom} \quad \text{intelligent}
\]

(A dog is intelligent)

d. [A professor discussing with his students why girls are so close to their mothers]

\[
t\text{-}u\text{hibb-u} \quad l\text{-}\text{bint-u} \quad ?\text{umm-a-haa} \\
3\text{-}\text{love-Nom} \quad \text{the-girl-Nom} \quad \text{mother-Acc-her}
\]

(A girl loves her mother)

e. [In a psychology class in an Egyptian school, students are discussing the effect of power on people’s personalities; they are listing what they think general attributes of a queen. One student says]

\[
al\text{-}\text{malika-t-u} \quad \text{mayruura} \\
\text{the-queen-f-Nom} \quad \text{arrogant-f}
\]

(A queen is arrogant)

f. [In a sociology class, A professor is telling her students that unlike unmarried girls in the West who often leave their parents’ house at 18, in Jordan]

\[
t\text{-}\text{a\text{fi}if-u} \quad l\text{-}\text{bint-u} \quad \text{ma\text{a}a} \quad \text{la\text{-}t\text{-i-haa} fi} \quad l\text{-}?\text{urduunn} \\
3\text{-}\text{live-Nom} \quad \text{the-girl-Nom} \quad \text{with family-f-Gen-her in} \quad \text{the-Jordan}
\]
(A (unmarried) girl lives with her family in Jordan)

The access to the right interpretation of the sentences in (37), generic or non-generic, is the denotation of the *al*-Ns used, whether they are semantically definite or indefinite. According to the definiteness criteria, *al-* is definite if it satisfies at least one of its concepts. More precisely, an NP whose referent is familiar, identifiable, or unique is definite, and indefinite otherwise. None of the *al*-Ns in (37) seems to be definite on the definiteness criteria. In (a), for example, the *al-N* *ar-radqul f-farqiyy* ‘the Middle Eastern Man’ does not denote a familiar, identifiable, or unique referent. Notice that both of them are talking about a possible husband in the future, not about an actual husband in the present, or a man proposing to the speaker that the hearer is familiar with. The right denotation of *ar-radqul f-farqiyy* ‘the Middle Eastern Man’ is any individual out of the class of individuals to which the descriptive content of the NP applies, the Middle Eastern man class. Support for this analysis comes from the context provided, where the hearer wonders why her friend is so strict about the origin of her future husband, though she can find a good husband even from a different origin. If both interlocutors are referring to a familiar, identifiable, or unique referent, this wonder would be out of context and not acceptable, or irrelevant. Similarly, the *al-N* in (b) does not denote a definite referent. *l-muṣallim* ‘the teacher’ the minister mentioning does not refer to a definite referent in the real sense. The minister is delivering a speech commemorating the Teacher Day, so nothing in the context or in the sentence indicates that the minister is referring to a particular definite teacher, but rather to any individual to which the descriptive content of *l-muṣallim* ‘the teacher’ is true. Likewise, the *al-N al-kalb* ‘the dog’ in (c) refers to any dog as being intelligent, not to a particular definite dog known to both interlocutors. In (d), *al-N l-bint* ‘the girl’ is semantically indefinite. Notice that the context provided is a class discussing a social phenomenon of girls being very close to their mothers, not a specific definite case of a girl who happens to love her mother. The referent of *l-bint ‘the girl’* being semantically indefinite is felicitous in this class context. The *al-N al-malika* ‘the queen’ in (e) is also semantically indefinite; nothing in the context indicates that a unique, familiar, or identifiable queen is referred to, but rather the opposite. Notice that the class is discussing the general attributes of a
queen in general, not a particular queen. It is worth mentioning that there is no queen in Egypt, and if reference were made to a particular queen, this queen should be identified whether she is the queen of Jordan, Morocco, Britain, or any other country that has a monarchy. This indefinite denotation holds for the *al-N* in (f), where reference is not made to a particular Jordanian girl, but rather to any Jordanian girl. This analysis is supported by the context provided, where a class is discussing women status in different countries and cultures. The property of living with one’s family is true of any Jordanian girl, not of a particular one only. It seems that all *al-N*s in (37) are indefinite in sense, though they are definite in form.

In sum, the NPs in (37) behave not like definite NPs in English, but rather as indefinite NPs, in the sense that they do not naturally denote unique, identifiable, or familiar groups of individuals, but rather the reversed, (see chapter 3). They denote functions that take situations and return sets of individuals. Informally, *l-mušallim* and *al-kalb* in (37. b&c) denote the set of all individuals which are *mušallim* ‘teacher’, and the set of all individuals which are *kalb* ‘dog’ in any given situation, respectively. In other words, such definite NPs are definite in form but indefinite in sense, in that they denote not familiar, identifiable, or unique entities, but rather sets of individuals to which the content of the nouns used apply. This analysis contradicts the traditional dichotomy of NPs in MSA which states that definites and indefinites are determined based on the presence/absence of a definite determiner like the article *al* ‘the’.

I now return to the interpretation of the sentences in (37). No sentence in (37) reports a single fact, a random event, or an accidental state of affair of a particular individual salient in the context domain. It is clear that the sentences in (37) report generalizations true of any individuals to which the descriptive contents of the *al-N*s apply. The sentence in (a), for example, says that in general, any individual who is Middle Eastern man has the property in the matrix, *be jealous*. This generalization is not true of a particular Middle Eastern man known to both interlocutors, but is true of Middle Eastern men generally. The sentence in (b) says that in general, an individual who is a teacher is an individual who knows his/her duties and responsibilities. In (c), a generalization is made about any
individual dog as being intelligent. The sentence in (d) expresses a generalization about queen individuals being generally arrogant. In (e), a generalization is made about girls stating that in general an individual who is a girl is such that that individual loves her mother. Similarly, (f) expresses a generalization about Jordanian girls. It says that in general a Jordanian girl individual is such that she lives with her parents before marriage. These sentences express true generalizations about non-particular members of classes to which the descriptive content of the NPs apply. What forces a generic interpretation of these sentences as generic only is the I-level predicates and the denotation of the al-Ns incorporated as not referring to familiar, identifiable, or unique individuals, but rather the reversed. This indefinite denotation avails the generic reading of the sentences since generics report regularities true of individuals, events, state of affairs, or facts, not instances of these. This shows how crucial identifying true definites in MSA is.

It is worth mentioning that all the sentences in (37) can be rephrased by inserting the adverbial phrase bifakil šaam ‘in general’ with no or slight change in meaning in the resulting sentence. Krifka et al. (1995) suggest using this test as a tool for distinguishing characterizing sentences from particular sentences. According to this test, if a sentence is combined with this adverb or similar adverbs like typically or usually, with no or slight change in meaning in the rephrasing sentence, the original sentence is characterizing.

However, sentences combining definite NPs in form and sense and I-level predicates, or S-level predicates, avail a non-generic reading only. Put differently, the use of semantically definite al-N forces an episodic reading as the only acceptable interpretation of the sentence. This is the kind of reading often gleaned from sentences reporting isolated facts, events, or states of affair about particular, context salient individuals. To see this, the sentences in (35) are repeated in (38) with contexts limiting the denotation of the NPs to definite referents.

(38) a. [On their way home, and after having dinner in their new Middle Eastern classmate’s house, one of the students made an observation about the Middle Eastern man’s wife not shaking hands with them. One of the students, who has known the Middle Eastern man for a year said]
b. [A school principle trying to calm down a parent who is complaining about his son’s math teacher for not grading homework]

\[ y-a\ṣrīf-u \quad l-mu\ṣallīm-u \quad mαš-ulīyya\ṣ-t-hī \]
3-know-Nom the-teacher-Nom responsibilities-Acc-his
(The teacher knows his responsibilities)

c. [A father asking his daughter about what she likes most about the new dog he brought her as a birthday present]

\[ al-kalb-u \quad ʤakīyy \]
the-dog-Nom intelligent
(The dog is intelligent)

d. [A father explaining to his friend why he decided to give full custody of his only daughter to her mother]

\[ t-uḥibb-u \quad l-bīnt-u \quad ʔumm-a-hāa \]
3.f-love-Nom the-girl-Nom mother-Acc-her
(The girl loves her mother)

e. [A woman whispering to her friend while attending an opening ceremony of the 32nd International Arab Youth Conference in the rose-red city of Petra under the patronage of queen Noor]

\[ al-malika-t-u \quad mayru\ṣura \]
the-queen-f-Nom arrogant-f
(The queen is arrogant)

f. [A Jordanian girl, who was working in a bank in Saudi Arabia, has resigned and returned to Jordan. Her supervisor, who was on a 2-year leave without pay, asked another Jordanian employee if she knows anything about that Jordanian girl. The Jordanian employee says]

\[ t-ʔa\ṣīf-u \quad l-bīt-u \quad ma\ṣa ʔaaʔila-t-ha\ṣa \quad fi \quad l-ʔurdu\ṣu \]
3.f-live-Nom the-girl-Nom with family-f-Gen-her in the-Jordan
(The (unmarried) girl lives with her family in Jordan)

The al-Ns in (38) are definite in form and sense, in that they refer to definite referents in the real sense. In (a), for example, ar-raḍgal-u f-ʃarqiyy-u ‘the Eastern man’ denotes a semantically definite referent as it satisfies the familiarity condition of the definiteness criteria. It refers to a particular individual interlocutors are familiar with from assumed shared knowledge of the referent of the NP. The speaker by using al-N here intends to signal to the hearer(s) that he refers to that particular Eastern man they are familiar with. Likewise, the al-N in (b) refers to a semantically
definite muʕallim ‘teacher’ referent. Reference is successful here because the NP l-muʕallim ‘the-teacher’ has been previously mentioned in the discourse. Therefore, the speaker by using l-muʕallim ‘the-teacher’ wants to indicate to the hearer that this teacher is co-referential with the teacher mentioned before, and now is familiar to both interlocutors through anaphoric familiarity. Similarly, in (c) successful reference of the use of al-N al-kalb ‘the dog’ is due to anaphoric familiarity. The NP al-kalb ‘the dog’ refers to a particular dog known to the speaker and the hearer. The NP al-bint ‘the girl’ in (d) is semantically definite as it denotes a familiar individual bint ‘girl’ to both interlocutors. The referent of al-bint ‘the girl’ is known to the speaker and the hearer based on shared knowledge, or uniqueness account because the hearer knows that the speaker has one daughter only, so she is unique context-dependently. In (e), the al-N al-malika ‘the queen’ denotes a definite referent that is familiar to interlocutors from a physical situation. The queen is present in the visible situation in which the interlocutors are. Lastly, the NP al-bint ‘the girl’ in (f) denotes a semantically definite referent known to the speaker and the hearer from both shared knowledge and anaphoric familiarity.

As expected, the sentences in (38) are non-generic. All sentences report isolated or accidental facts and state of affairs true of particular individuals. No regularity is reported in any of the sentences. To see this, notice that unlike the sentences in (37), which avail generic interpretations only, the sentences in (38) accept inserting personal demonstratives like haadaa ‘this’, haadhi ‘this.f’ with very slight change in meaning. The presence of such demonstratives enforces an episodic reading of the sentence. Consider some of the sentences in (38) repeated in (39) with demonstratives incorporated.

(39) a. [A school principle trying to calm down a parent, who is complaining about his son’s math teacher for not grading homework]

\[
\text{y-aʕrif-u} \quad \text{ʔaʔalika} \quad \text{l-muʕallim-u} \quad \text{masʔuuliyyaat-i-hi}
\]

3-know-Nom that the-teacher-Nom responsibilities-Acc-his

(That teacher knows his responsibilities)

b. [A father asking his daughter about what she likes most about the new dog he brought her as birthday present]

\[
\text{haaʔdaa} \quad \text{al-kalb-u} \quad \text{ʔakiyy}
\]
this **the-dog** Nom intelligent
(This dog is intelligent)

c. [A father explaining to his friend why he decided to give full custody of his only daughter to her mother]

\[
t-uḥibb-u \quad hādāhi \quad l-bint-u \quad ?umm-a-haa
\]
3.f-love-Nom this.f **the-girl** Nom mother-Acc-her

(This girl loves her mother)

d. [A woman whispering to her friend while attending an opening ceremony of the 32nd International Arab Youth Conference in the rose-red city of Petra under the patronage of queen Noor]

\[
hādāhi \quad al-malika-t-u \quad mayruur-a
\]
this.f **the-queen**-f Nom arrogant-f

(This queen is arrogant)

However, inserting demonstratives in the sentences in (37) with the same contexts provided would not only change the meaning radically, but also render these sentences infelicitous. Consider (40).

(40) a. [In a press conference on the Teacher Day in Jordan, The minister of education expresses his trust in teachers, and that they know their job duties and responsibilities]

\[
?y-aṣrif-u \quad hāda/ ḍaalika \quad l-muṣallim-u \quad mas?uḍiyya-t-i-hi
\]
3.know-Nom this/that **the-teacher**-Nom responsibilities-Acc-his

(This/that teacher knows his responsibilities)

b. [A father is asking his little daughter about which pet she wants as a birthday present and why. The girl says she wants an intelligent animal, and]

\[
ʔhaaḍa/aa \quad al-kalb-u \quad ḍakiyy
\]
this **the-dog** Nom intelligent

(This dog is intelligent)

c. [A professor discussing why girls are so close to their mothers]

\[
ʔt-uḥibb-u \quad hādāhi/tilka \quad l-bint-u \quad ?umm-a-haa
\]
3.f-love-Nom this.f/that.f **the-girl** Nom mother-Acc-her

(This/that girl loves her mother)

d. [In a psychology class in an Egyptian school, students are discussing the effect of power on people’s personalities; they are listing what they think general attributes of a queen. One student says]

\[
ʔhādāhi/tilka \quad al-malika-t-u \quad mayruur-a
\]
this.f/that.f **the-queen**-f Nom arrogant-f

(This/that queen is arrogant)
e. [In a sociology class, a professor is telling her students that unlike unmarried girls in the west who often leave their parents’ house at 18, in Jordan]

\( ?t-a\text{ā}i\text{i}f-u \quad haa\text{ā}dhi \quad l-bint-u \quad m\text{ā}\text{s}a \quad ?aa\text{ā}t-i-haa \quad fi \quad l-\text{ʔ}urdunn \)

3.f-live-Nom this the-girl-Nom with family-f-Gen-her in the-Jordan

(This (unmarried) girl lives with her family in Jordan)

In chapter 3, a clear distinction between the syntactic status of definite NPs and their semantic denotation was made. Syntactically speaking, the definite article \( al \) does not mark definiteness exclusively, but serves, among other functions, as a syntactic marker to license NPs to occur in argument position, a characteristic MSA shares with many other languages which prohibit bare NPs from being arguments\(^{36}\), like Spanish (Laca 1990). The semantic denotation of the definite NP in MSA, however, is ambiguous between a definite reading, object/kind-referring NP, and an definite reading, quantificational NP. More to the point, the definite article in MSA may be used with semantically indefinite NPs. This explains the English translations given in (35). Disambiguating these three readings depends mainly on the context and other constituents of the sentence such as the predicate.

In sum, it seems that associating an S-level predicate with a definite NP avails an existential reading of the NP, and hence an episodic interpretation of the sentence as the only acceptable interpretation. Incorporating an I-level predicate in the same syntactic environment, however, yields in an ambiguous interpretation between generic and nongeneric readings. This ambiguity can be explicated if the right denotation of the grammatically definite NP used is obtained. More precisely, if the definite NP is semantically indefinite, the only reading that the sentence avails is generic, in which a generalization is reported about an individual-variable supplied by the semantically indefinite NP. If the grammatically definite NP, however, is semantically definite, the only reading the sentence avails is existential, in which an isolated property is true of a particular, contextually salient individual.

\(^{36}\) This generalization needs some refinement since in MSA bare NPs can occur as arguments in a number of syntactic structures, (see section 6).
This analysis clearly shows that the semantic representation of characterizing sentences in MSA is not language peculiar, but rather complies with the semantic architecture of a characterizing sentence in natural languages. According to this architecture, a characterizing sentence reports regularity over individuals through an individual variable provided by a semantically indefinite NP (see e.g. Diesing 1992, Heim 1982, Krifka et al.1995, among others), regardless of the grammatical definiteness status of that NP which can be realized variably in different languages (Lyons 1999).

5 Reference to a Kind and the Definite NP in MSA

Prasada (2012) argues that a remarkable fact about human cognition is that though we see a limited number of particular individuals, we are able to generalize, characterize, and speak about kinds or species based on these individuals. We are able to use nominal phrases to talk systematically about kinds and specimens of kinds. This furnishes important syntactic and semantic questions pertaining to the peculiarities of nominal phrases that allow us to do so. More precisely, what is it that a nominal expression exhibits that gives rise to this double faceted use? What is the role of the other sentence constituents, the predicate in particular, which help both the speaker and the hearer pick out the intended meaning of the nominal expression used? Languages seem to syntactically encode kind-denoting nominal expressions, but they do so variably. A characteristic property of the genericity phenomenon, both characterizing sentences and reference to kind sentences, is that realizing genericity linguistically is often not straightforward. Put differently, natural languages do not seem to employ specific expressions to unequivocally mark generic sentences. In English, for example, the linguistic realization of kind-denoting NPs takes more than one nominal form, like the definite singular count nouns, bare plural count nouns, and bare mass nouns (Krifka et al. 1995). Spanish, unlike English, does not allow bare plural NPs to be generic, and restricts kind-referring expressions to definite NPs, singular and plural. Russian, Japanese, Chinese, and Turkish do not grammaticalize definiteness since they have no article system, and instead employ bare NPs to realize kind reference (Borik and Espinal 2012; Snape et al. 2009). In these languages, the bare NP is ambiguous in four

---

37 See Chierchia (1998) for a detailed study of cross-linguistic variation in realizing reference to kind.
ways. It can be interpreted as definite, indefinite, kind-referring, or object-referring. To determine the exact denotation of a bare NP, speakers resort to pragmatics, or other morphosyntactic tools, like position in the sentence, number, morphemes with article-like functions as demonstratives, and quantifiers.

Though it is well-known that MSA does not allow bare NPs to be generic since it restricts their distribution in argument position, MSA like many other languages, does not seem to dedicate a specific NP form to formally encode kind reference. The grammatically definite NP, singular and plural, can be used as kind-denoting nominal expression.

(41) a. ad-daynaasˤuur-u hayawaan-un munqaridˤ
the-dinosaur-Nom animal-Nom extinct
(The dinosaur is extinct)

b. ad-daynaasˤuuraat-u hayawaanaat-u-n munqaridˤa
the-dinosaurs-Nom animals-Nom-N extinct
(Dinosaurs are extinct)

(42) [In a class about the history of elephants in the United States, the professor asks students about the date elephants arrived in USA, one students answered (a), and another (b)]

arrived-3 the-elephant-Nom to America year-Acc 1796
(The elephant arrived in the Unites States in 1796)

b. wasˤal-at l-fiyala-t-u ?ila-ʔ ?amriikaa ?aam-a 1796
arrived-3.f the-elephants-f-Nom to America year-Acc 1796
(Elphants arrived in the Unites States in 1796)

The NP subjects in (41&42) are generic NPs; they do not denote an individual or groups of individuals, but rather a species or genus. This can be explicated if we examine the predicates incorporated in the sentences. In (41) the predicate associated, munqaridˤ ‘extinct’, is unanimously classified as K-level predicate. It is semantically and logically implausible to construe it holding for an individual or groups of individuals; only a species becomes extinct, not a specimen of that species. This type of predicates report properties which are true of a species or a kind only, hence s-selecting kind-denoting NPs. Since many semanticists use these predicates as a diagnostic test for the classification of generics into reference to a kind genericity and characterizing sentences (see Borik
and Espinal 2012, Carlson 1995, Krifka 1987, Krifka et al. 1995, among others), it is plausible to interpret the NP subjects in (41) as kind-referring NPs. The predicate in (42), wasˤal ‘arrived’, however, is not a K-level predicate. In fact, it is an S-level predicate. Therefore, what makes speakers of the language under examination interpret the subject definite NPs in these sentences as kind-denoting. An answer might be related to the definite NP itself, being idiosyncratic in the language; i.e., MSA invariably uses this NP form as a generic NP. This claim, however, can be easily countered by noticing that the definite NP, singular and plural, can be used in sentences that unambiguously express episodic propositions (43).

(43) a. haaaða l-fiil-u y-abduu mutʕab
   this the-elephant-Nom 3-seem tired
   (This elephant seems tired)

   b. tilka l-fiyalat-u t-aqtarib-u min-naa
   those the-elephants-Nom 3.f-approach-Nom from-us
   (Those elephants are approaching us)

   A more plausible answer can be constructed if we look at the tense used in the sentences in (42). The tense of these sentences is past. This becomes highly significant if we observe that in MSA all characterizing sentences, with object-referring subjects, are stated in the present tense, and using the past tense invalidates the generic interpretation and forces an episodic interpretation instead (44)38. In reference to kind generics and characterizing sentences incorporating kind-referring NP subjects, however, both present and past tenses can be used with no effect on the genericity expressed. It is worth noticing that both sentences in (42) can receive an episodic reading, where a particular, contextually salient elephant in (a), and elephants in (b), arrived in the United States in 1796. However, the scenario provided excludes the object denotation of the NP subjects, and entertains kind denotation as the only acceptable interpretation.

(44) a. an-numuur-u t-aʔkul-u/*ʔakal-at l-lahm
   the-tigers-Nom 3.f-eat-Nom/ate-3.f the-meat
   (Tigers eat/ate meat)

38 The asterisk in the given examples indicates that the sentence no longer expresses a generalization, and can only receive an episodic interpretation.
b. tamfii/*maʃ-at manaal-u ?ilaal l-madrasat-ʃ fi sˤ- sˤabaah
3.f-walk/walked-3.f Manal-Nom to the-school-Gen in the-morning
(Manal walks/walked to school in the morning)

In addition to past tense being compatible with reference to a kind genericity in MSA, reference to kind generics and habituals seem compatible with S-level predicates. Characterizing sentences, however, do not allow S-level predicates. This can be explained if we observe that reference to kind sentences are intended to establish the NP as denoting a kind or species, and thus such sentences can tolerate S-level predicates as long as the NP used is generic. In other words, the role played by the predicate is peripheral compared to the NP. Habitual sentences, however, generalize over sums of events that can be extracted from S-level predicates that have been changed into stative predicates through the regularity expressed, and hence such predicates function in a way similar to I-level predicates. Therefore, it seems that S-level predicates block genericity in characterizing sentences, but allow it in both reference to kind sentences and habituals. These asymmetries confirm a claim made in Krifka et al. (1995: 63) that unlike characteristic genericity, which is gleaned from the collaboration of all constituents in a sentence, reference to a kind genericity is a phenomenon “tied to the NPs in question and not to the sentences as a whole.”

Although MSA refrains from using bare NPs as generic NPs, and employs definite in form NPs as a medium to express reference to a kind genericity, not any definite NP is qualified to do so. The syntactic status must be coupled with a special semantic privilege for a definite NP to be kind-referring expression. Semantically speaking, natural languages appear to establish kind reading of a nominal expression only if there is an inherent characteristic property obtained by the referent denoted by the nominal expression; “the noun or complex nominal constituent must be semantically connected with a “well-established kind” to which the noun phrase then can refer.” (Krifka et al. 1995: 11)

This crucial criterion is based on a connection between a kind and a characteristic property. In the literature, two connection types have been observed: non-accidental statistical connection, and non-accidental non-statistical connection (Prasada 2012). While the former is established statistically
by observing a number of specimens, and based on this observation a generalizing property is inductively attributed, the latter seems more essential, inherent, and definitional (Burton-Roberts, 1977; Carlson, 1977, 1995; Krifka et al., 1995; Greenberg, 2012, Pelletier, 2010b). This semantic distinction explains speakers’ intuitions pertaining to the semantic status of the definite NP subject arguments in (45) and (46), where only the latter is construed to denote a kind.

(45) *al-wuruud-u al-muhdaat-u fii ʕiid-i l-ʔubb-i ʔamraa?
    the-roses-Nom the-gift-Nom in Eid-Gen the-love-Gen red
    (Valentine-gift roses are red)

(46) *an-namir-u muxatˤatˤ
    the-tiger-Nom striped
    (The tiger is striped)

The connection between the NP and the predicate determines whether this NP denotes a kind or not. In (45), the connection between ‘valentine-gift roses’ and the property attributed ‘be red’ is not accidental since almost all roses presented on that day are red. However, this connection is statistical in nature; in the sense that it has been established inductively by observing a number of roses, being presented on valentine day and based on this limited observation a generalization is expressed. Except for this, there is no inherent or essential *redness property of these roses that would establish them as a kind. In (46), however, the connection is different; being striped is an essential property of being a tiger. This analysis is bolstered by the significantly variable truth judgments of these two sentences when initiated by the phrase *haʔa n-nawʕ min ‘this kind of’ which is taken as a canonical expression that forces kind reference reading of the following nominal expression (see Carlson 1977). If someone points to a tiger and a valentine-gift rose and utters the sentences in (47) and (48), the former will be invariably true, but the latter false.

(47) *haʔa n-nawʕ-u mina l-gitˤatˤ-i muxatˤatˤ
    this the-kind-Nom from the-cats-Gen striped
    (This kind of cats is striped)

(48) ?haʔa n-nawʕ-u mina l-wuruud-i ?ʔahmar
    this the-kind-Nom from the-roses-Gen red
    (This kind of roses is red)
The sentence in (48) is judged false because it tries to establish the roses presented on Valentine day as a kind based on their connection with statistically computed redness property. This contradicts our intuition, and the way we construe and encode kinds.

Prasada (2010; 2012) and Prasada & Dillingham (2006) seek to present a logical explanation for this problematic connection issue holding between kinds and properties. They argue that in order to dig deep in these two connection types, reference must be made to the notion of non-accidentalness, rather than to parasitic assimilations to properties designed essentially to serve other semantic structures like essence, definition, and analyticity. They distinguish between two types of nonaccidental connections; namely, special connection that exhibits principled properties which specimens of a kind obtain in virtue of being the kind or kinds of things they are, and statistical connection which involves properties holding for instances of a kind in general. Properties with a principled connection to a kind were dubbed k-properties, while those with statistical connection were dubbed t-properties. K-properties, but not t-properties, are characterized as having explanatory, definitional, and analytical dimensions. The distinction can be explicated in that only sentences with t-properties can be rephrased with a sentence combining the adverbial phrase bifakl-i-n ʕaam ‘in general’, with almost no change in meaning:

(49) a. al-ʔusuud-u  mina 0-ʔadiyyaat
     the-lions-Nom  from  the-mammals
     (Lions are mammals)

     b.? bifakl-i-n ʕaam  al-ʔusuud-u  mina 0-ʔadiyyaat
     in-form-Gen-N  general  the-lions-Nom  from  the-mammals
     (In general, lions are mammals)

(50) a. fustan-u  l-ʕaruus-i ʔabyadʕ
     gown-Nom  the-bride-Gen  white
     (A bridal gown is white)

     b. bifakl-i-n ʕaam  fustan-u  l-ʕaruus-i ʔabyadʕ
     in-form-Gen-N  general  gown-Nom  the-bride-Gen  white
     (In general, a bridal gown is white)

39 In certain cases, Kind-referring expressions associated with k-properties can be modified by the adverb in general, but still other asymmetries can be shown to distinguish k-properties from t-properties and their relation to the two connection types, see Prasada (2012) for convincing arguments in favor of this distinction.
In statistical connection, Prasada (2012) argues, the strength of the association is established based on the degree of prevalence of a property among instances of a kind. In other words, statistical connections are mathematical in nature, and hence promiscuous and susceptible to hold between any two things. However, “principled connections are formal connections that are limited to holding between conceptual representations of kinds and properties” (P.53). In such association, an instance of a given kind is judged to hold a property just in virtue of being the kind of thing it is. This representation explains why principled connections are normative, analytical, and definitional in nature.

In sum, MSA employs both grammatically definite singular and grammatically definite plural nouns as kind-referring expression. However, on par with reference to a kind genericity in many other natural languages (such as English, Indonesian, Hindi, and German), only nominal expressions that denote semantically well-established kinds can function as kind-denoting NPs. Although reference to kind sentences, but not characterizing sentences, can use past tense and S-level predicates, this cannot be taken as a necessary and sufficient condition for encoding reference to a kind genericity because the present tense can be used also, and S-level predicates can be used in episodic sentences. In order to accurately identify reference to a kind genericity in MSA, the semantic status of the nominal subject argument - being connected to a well-established kind, other constituents in the sentence - including tense and predicate level, and the context must be investigated.

6. Genericity in the Construct State

The interaction between genericity phenomenon and nominal expressions in MSA is not exclusive to simple NPs. Genericity manifests itself in another complex nominal expression termed the construct state. However, manifestations of genericity in the construct state are significantly different from its manifestations in simple NPs. In particular, unlike bare nouns which are allowed in episodic sentences only, bare construct state phrases can be used in characterizing sentences, and more interestingly, sentences combining bare construct state phrases and I-level predicates express a generic reading only, a peculiar characteristic which sets them apart from both bare nouns and *al-Ns,
see 6.2. The semantic nature of kinds or species denoted by kind-referring construct state expressions is also different from kinds or species denoted by *al-Ns* when used as kind-referring nominal expressions. While the latter denote whole kinds or species, the former denote sub-kinds or sub-species; see 6.3 below.

Before investigating genericity in construct state phrases, I present a brief syntactic overview of construct state in (6.1) below.

### 6.1 A Syntactic Overview

The construct state (CS) is the peculiar Arabic, and probably Semitic, syntactic construction par excellence (Hoyt 2008). Due to its interesting properties, it has attracted much attention particularly within the modern syntactic theories (e.g. Benmamoun 2000, 2006; Fassi Fehri 1993; Kaplan 1993; LeTourneau 1995; Mohammad 1999; Shlonsky 2004). CS is a type of annexation consisting of at least two members, most often nouns. CS incorporates two nominal expressions and collapses them into one constituent. The first noun is the head of the phrase, and is traditionally called possessee noun; the last noun, the complement, is termed the possessor. In principle, there is no upper limit on the number of the construct state’s members as long as all the pre-final members are grammatically indefinite. In fact, CS is closed and no further embedding is possible once a grammatically definite noun is used (Mohammad 1999). The possessed noun is never overtly marked for in/definiteness. By in/definiteness here I mean grammatical in/definiteness. Semantic in/definiteness is determined based on the definiteness criteria, as in simple NPs. In/definiteness marking is restricted to the last member of the construct state. More precisely, the in/definiteness of CS members is determined solely by the last member; if in/definite, other members will inherit this value, and hence being in/definite accordingly. The first member always carries the main case of the phrase, which varies according to its position and syntactic function in the sentence. However, the last

---

40 The first member in a construct state does not have to be always a noun; adjectives, verbal nominals, quantifiers, and other expressions can occur as the first member of a construct state; see Mohammad (1999), and Benmamoun (2006) for a detailed discussion.
and the other members, if any, are invariably assigned genitive case. The examples in (51) illustrate these properties.

(51) a. maktab-u/a/i  l-ʕamiid-i  
    office-Nom/Acc/Gen  the-dean-Gen
    (the dean’s office)

b. maktab- u/a/i  bn-i  ʕamm-i  t'aailib-i  l-ʕamiid-i
    office- Nom/Acc/Gen  son-Gen  uncle-Gen  student-Gen  the-dean-Gen
    (the office of the son of the paternal uncle of the dean’s student)

In MSA, the construct state expresses a diverse range of possessive and partitive relationships. These include material and inalienable possession (a&b), location (c), part-whole relationships (d), measure or quantity (e), and comparison (f) (Mohammad 1999; Hoyt 2008).

(52) a. sayyaara-t- u/a/i  muʕallim-i-n
    car-f. Nom/Acc/Gen  teacher-Gen-N
    (a teacher’s car)

b. ibn- u/a/i  l-waziir-i
    son- Nom/Acc/Gen  the-minister-Gen
    (the son of the minister)

c. ɣurfa-t- u/a/i  n-nawm-i
    room-f- Nom/Acc/Gen  the-sleep-Gen
    (the bedroom)

d. baab- u/a/i  s-sayyaara-t-i
    door- Nom/Acc/Gen  the-car-f-Gen
    (the car’s door)

e. kaʔs- u/a/i  maaʔ-i-n
    glass- Nom/Acc/Gen  water-Gen-N
    (a glass of water)

f. ʔaqdam- u/a/i  l-mudun-i
    oldest- Nom/Acc/Gen  the-cities-Gen
    (the oldest (city) of cities)

It is worth mentioning that no modification is permitted to intervene between the members of a construct state; all modifiers must follow the construct state phrase. Mohammad (1999: 29) claims that modification being placed outside the CS in observance of the strict adjacency constraint makes it applicable to any member in the CS; that is the CS is ambiguous in this regard. He cites the following example arguing that it has four different readings:
1. The tall son of the maternal uncle of the father of the boy
2. The son of the tall maternal uncle of the father of the boy
3. The son of the maternal uncle of the tall father of the boy
4. The son of the maternal uncle of the father of the tall boy

However, it should be emphasized that the above flexibility is controlled by noun-adjective agreement. If there is concord mismatch between a modifier and any member in the CS, this mismatch forces an interpretation where no modification of that member holds.

After this brief syntactic overview of the construct state in MSA, I turn to discuss manifestations of genericity in sentences incorporating construct state phrases. In (6.2) I investigate characterizing sentences which associate construct state phrases as arguments. Reference to a kind of genericity in sentences incorporating construct state phrases is investigated in (6.3).

### 6.2 Generics Associating Construct State Phrases

As mentioned in the beginning of this chapter, MSA does not put restrictions on nominals used in generic sentences, particularly characterizing sentences, unless a nominal expression is grammatically indefinite. Therefore, one would intuitively expect to find generic sentences which use construct state phrases as arguments, provided that the CS phrases are definite in form. Consider the data in (54)\(^{41}\).

\[(54) \begin{align*}
a. & \text{ʕaamil-u l-matˤam-i y-astayqiðˤ-u mubakkiran} \\
& \text{worker-Nom the-restaurant-Gen 3-wake-Nom early} \\
& \text{(The/A restaurant worker wakes up early)} \\

b. & \text{ʔustaað-u l-dgaamiʃat-i y-aḥḍˤaa bil-ihtiraam} \\
& \text{teacher-Nom the-university-Gen 3-enjoy with-the-respect} \\
& \text{(The/A professor is well respected)} \\

c. & \text{t-uɣlaq-u bawwaabat-u l-madrasat-i fi l-tˤutiːli} \\
& \text{3.f-close.Pass-Nom door-Nom the-school-Gen in the-holidays-Gen} \\
& \text{r-rasmiyy-ʃ the-official-f} \\
& \text{(School doors close on official holidays)}
\end{align*}\]

---

\(^{41}\) The habitual reading of these sentences is irrelevant to this discussion. Both the episodic and lexically characterizing readings are investigated based on the semantic definiteness status of the CS subjects.
The NPs used in (54) as subject arguments are definite in form CS phrases. As the behavior of simple definite NPs used in sentences combining I-level predicates, or in habitual sentences, the sentences here are ambiguous between two interpretations: an episodic reading in which an accidental property or event is reported of a salient individual, and a generic interpretation. The sentence in (a), for instance, can be either interpreted existentially as reporting that there is a familiar or unique, contextually salient restaurant worker who happens to wake up early at a specific time and location\(^{42}\), or it can receive a generic interpretation; in the sense that restaurant workers in general wake up early. The exact semantic status of the \textit{al-N} member, and hence the maximal CS phrase, in terms of definiteness is decisive in determining which reading the sentence avails, generic or episodic. The context and other pragmatic tools play a crucial role in teasing apart the definiteness status of the CS. Similarly, (b) can be interpreted as ‘there exists a contextually prominent professor who is well respected’, or ‘professors in general are well respected’. The sentence in (c) does not use an I-level predicate; \textit{tuylaq} ‘is closed’ is in fact an S-level predicate, but still this sentence can entertain a habitually generic interpretation that can be verbalized as follows: ‘situations of official holidays are such that school doors close in these situations’. It is worth mentioning that (c) can also receive an episodic interpretation with a semantically definite individual school.

Therefore, if we provide contexts that limit the denotation of CS phrases in (54) to semantically definite referents, the only interpretation the sentences express is episodic. However, providing contexts which limit the denotation of CS phrases to variable individuals; i.e., semantically indefinite referents, enforces a generic reading of the sentences in (54) as the only available reading. Put differently, sentences with grammatically definite CS and I-level predicates behave exactly like the sentences with simple grammatically definite nouns and I-level predicates discussed in (4.3). To see this, and to save ink, sentence (54.d) is repeated in (55.a) with a context limiting the denotation of

\(^{42}\) Even in this context, the sentence can still receive a generic interpretation, in particular a habitual one.
CS to a definite referent, and in (55.b) with a context enforcing a semantically indefinite denotation of CS.

(55) a. [A girl is arguing with her mother about buying new clothes. The mother told her not to argue with her, but talk to her father because]

\[
\text{rabb-}u \ 1-\text{?usrat-}i \ mas\?uul-u-n \ ?an \ ihtiyaadgaat-i \ ?abnaa?-i-h} \\
\text{head-Nom \ the-family-Gen \ responsible-Nom-N \ for \ needs-Gen \ children-Gen-his}
\]

(The head of the family is responsible for his children’s needs)

b. [A teacher discussing with his senior students responsibilities of head of the family in the Arab world]

\[
\text{rabb-}u \ 1-\text{?usrat-}i \ mas\?uul-u-n \ ?an \ ihtiaalgaat-i \ ?abnaa?-i-h} \\
\text{head-Nom \ the-family-Gen \ responsible-Nom-N \ for \ needs-Gen \ children-Gen-his}
\]

(Head of the family is responsible for his children’s needs)

As expected, (a) is a particular sentence reporting a random fact about a particular head of the family father known to both interlocutors from shared knowledge familiarity as being responsible for his children’s needs; no generalization is expressed. The sentence in (b), however, where the CS denotes a variable individual head of the family, avails a generic interpretation as the only available reading. It says that in the Arab world situation, head of families in general are responsible for their children’s needs. The latter is a characterizing sentence, but the former is a particular sentence. The crucial difference in both sentences is brought by the different semantic denotations entailed by the same CS in different contexts. The same rational holds for the other sentences in (54) and any similar sentences with the same structure.

However, unlike simple nominal expressions that promote only episodic readings when are grammatically indefinite, grammatically indefinite CS phrases can be used in characterizing sentences, as illustrated in the following sentences.

(56) a. \[
\text{kalb-}u \ hiraasat-}i-n \ \text{xayr-}u-n \ \text{min} \ \text{djihaaz-}i \ \text{?indaar} \\
\text{dog-Nom \ guard-Gen-N \ better-Nom-N \ than} \ \text{system-Gen \ alarm}
\]

(A guard dog is better than a security alarm system)

b. \[
\text{kalb-}u \ s'}ayd-}i-n \ ?fd'al-}u \ \text{min} \ \text{bunduqiyat-i} \ s'}ayd \\
\text{dog-Nom \ hunt-Gen-N \ better-Nom \ than} \ \text{gun-Gen-N \ hunt}
\]

(A hunting dog is better than a hunting gun)

c. \[
\text{fadgarat-}u \ zaytuun-}i-n \ \text{t-}\text{ut'im-}u \ \text{?usrat-a}n \ \text{kaamil} \\
\text{tree-Nom \ olive-Gen-N \ 3.f-feed-Nom \ family-Acc-N \ whole}
\]
(An olive tree is enough to feed a whole family)

d. nuur-u misbaah-i-n y-ubaddid-u 3-dispel-Nom the-darkness
   light-Nom lamp-Gen-N (The light of a lamp dispels the darkness)

e. t-udxil-u kalimat-u haqq-i-n al-danna
   3-f-enter-Nom word-Nom truth-Gen-N the-heaven
   (A word of truth leads to entering Heaven)

f. t-uynii fabkat-u s'ayyd-i-n san qaarib
   3-f-substitute net-Nom fishing-Gen-N for boat
   (A fishing net substitutes for a boat)

g. t-uhyii zaxxat-u matar-i-n l-card-a l-mayta
   3-f-revive shower-Nom rain-Gen-N the-land-Acc the-dead
   (A shower of rain revives a dead land)

These examples are true characterizing sentences, though their subject arguments are grammatically
indefinite CS phrases. Intriguingly, these sentences can only be interpreted as expressing generic
readings, and existential interpretations seem to be completely unavailable. The sentence in (a), for
example, expresses a generalization about guard dogs, stating that in general, a guard dog is better
than a home security alarm system. The sentence does not have another interpretation where an
existential guard dog referent is accidentally reported to be better than an alarm system; the property
of being better than a security alarm is predicated to guard dogs in general. Similarly, (b) expresses a
generalization about hunting dogs in general being better than hunting guns. In (c) a generalization is
reported about olive trees in general, not an existential olive tree. (d) does not report a property of
light of a certain lamp, it is about light of lamps in general. The rest of the sentences are generically
interpreted in the same fashion.

It is worth mentioning that the position of the indefinite CS does not affect the interpretation
of the sentence. Sentences (a-d) are SV(O), whereas sentences (f-g) are VS(O). What is puzzling is
that in these sentences the generic reading is the only reading available; while their minimal
counterparts with definite CS phrases can have both generic and existential readings, as mentioned
above. More precisely, I am comparing the sentences in (56) with their minimally counterparts with
grammatically definite CS uttered out of the blue; i.e., with little explicit contextual support. In this
case, the latter are ambiguous between generic and non-generic readings depending on the semantic definiteness status of the grammatically definite CS subject. Sentences with grammatically indefinite CS, however, are not ambiguous, and only express a generic interpretation. This can be clearly shown by inserting a demonstrative pronoun like *haaða ‘this’ or *haaðihi ‘this.f’, which forces an existential reading of the sentence as the only reading possible. The demonstrative *haaða ‘this’ seems compatible with sentences incorporating grammatically definite CS phrases; in this case the generic reading is excluded, and the episodic reading with a contextually salient referent is promoted instead. However, *haaða is incompatible with any of the sentences in (56). This becomes clear if we notice that these sentences can only be interpreted generically. The pair in (57) illustrates this point.

(57) a. 

\[t-\text{uynii} \quad \text{fabakat-u} \quad s^\text{3}-s^\text{a} \text{yyd-i} \quad \text{haaðihi} \quad \text{san} \quad \text{qaarib}\]

3.f-substitute net-Nom the-fishing-Gen this.f for boat

(This fishing net substitutes for a boat)

b. *

\[t-\text{uynii} \quad \text{fabakat-u} \quad s^\text{a} \text{yyd-i-n} \quad \text{haaðihi} \quad \text{san} \quad \text{qaarib}\]

3.f-substitute net-Nom fishing-Gen-N this.f for boat

(This fishing net substitutes for a boat)

An interesting question relevant to the semantic status of indefinite in form CS compared to both indefinite simple nouns and definite CS is to be addressed here. More precisely, what semantic privilege does bare CS exhibit that renders it behaving differently from both bare simple nouns which relatively cannot be inserted in characterizing sentences, and definite in form CS phrases which render their sentences ambiguous between generic/nongeneric readings? If we examine the internal structure of bare CS, we can get a helpful insight. The essence of a CS is that it contains two nouns in which the last seems to indirectly modify the first head noun. This indirect modification cancels the existential interpretation that bare nouns entertain, and promotes an indefinite denotation of the noun in the real sense. The head noun becomes non-specific in the sense that it no longer denotes individual(s) existentially. Modification of the noun in a grammatically indefinite CS renders it denoting a content-descriptive frame upon which a set of individuals to which this description applies is picked. Therefore, this noun no longer denotes existentially computed individual(s), as simple bare nouns do. In (56.a) kalb-u hiraasat-i-n ‘a guard dog’ is not an existential dog; it is a dog with a
specific property. Similarly, *fadjarat-u zaytuun-i-n* ‘olive tree’ is not an existentially calculated tree; it refers to any individual that satisfies the descriptive content of the CS. Modification in bare CS changes the semantics of the head noun from denoting existentially computed individuals in a specific situation to denoting a set of individuals or objects that satisfy the properties provided by the head noun and the complement in all relevant situations. This explains the difference in denotation between bare CS and simple bare nouns, and their asymmetrical behavior with regard to characterizing sentences. It is noteworthy that unlike sentences combining grammatically simple indefinite subjects, sentences with indefinite in form CS subjects are incompatible with time adverbials which enforce an episodic reading of the sentence like *alʔaan* ‘now’, *al-yawm* ‘today’, *haadaa lʔusbuu* ‘this week’, or *haadiihi lʔayyaam* ‘these days’, which supports the claim that a bare CS does not denote an existential individual or individuals. Consider the examples below.

(58) a. *y-alsab-u walad-u-n fi f-faariς*  
3-play-Nom boy-Nom-N in the-street  
(A boy is playing on the street)

b. *y-alsab-u walad-u-n fi f-faariς alʔaan*  
3-play-Nom boy-Nom-N in the-street the-now  
(A boy playing on the street now)

(59) a. *nuur-u misʔbaah-i-n y-ubaddid-u δʔulma*  
light-Nom lamp-Gen-N 3-dispel-Nom the-darkness  
(The light of a lamp dispels the darkness)

b. *nuur-u misʔbaah-i-n y-ubaddid-u δʔulmat-a alʔaan*  
light-Nom lamp-Gen-N 3-dispel-Nom the-darkness-Acc the-now  
(A light of a lamp is dispelling the darkness now)

Both sentences in (58) are acceptable and interpretable since they express an existential reading, brought by the indefinite NP in (a), and the existential time adverbial *alʔaan* ‘now’ in (b). Both sentences express the same proposition that in a particular situation in the actual world, an existentially computed boy individual is involved in an episodic playing event at a particular place and time. The pair in (59), however, is not congruent. While (a) is a characterizing sentence reporting a generalization about light of lamps in general as holding the property of dispelling darkness, the sentence in (b) is completely ungrammatical. Its ungrammaticality is due to the presence of *alʔaan*
‘now’ which forces an existential reading of the CS subject, and the sentence as a whole. As the indefinite CS nuur-u mis‘baah-i-n ‘a light of a lamp’ does not denote an existentially computed specific object, it is naturally incompatible with existential constructions like al-ʔaan ‘now’, hence the ungrammaticality of (59.b).

As for the asymmetry between definite and indefinite in form CS phrases, the indefinite CS, unlike its definite counterpart, avails only a generic reading of the sentence, contributed, in part, by the variable individual denoted by the indefinite in form and sense CS. The other episodic reading, which a sentence with a definite CS can also express when the grammatically definite CS is semantically definite, is not available in sentences combining indefinite CS because an indefinite in form CS, as well as simple bare nouns, cannot be semantically definite in MSA. Indefinite CS is on par with semantically indefinite al-N, but not with semantically definite al-N. The semantically definite CS can be interpreted as denoting a particular, salient or unique individual. The indefinite CS, however, provides a descriptive frame upon which a set of individuals is picked. Therefore, as expected sentences incorporating grammatically definite subjects CS phrases are compatible with existential time adverbials like al-ʔaan ‘now’, but their minimally contrasting sentences with grammatically indefinite CS phrases are not.

(60) a. nuur-u l-mis‘baah-i y-ubaddid-u ʔolma-
light-Nom the-lamp-Gen 3-dispel-Nom the-darkness
(The light of a lamp dispels the darkness)

a’. nuur-u l-mis‘baah-i y-ubaddid-u ʔolma-
light-Nom the-lamp-Gen 3-dispel-Nom the-darkness-Acc the-now
(The light of the lamp is dispelling the darkness now)

b. nuur-u mis‘baah-i-n y-ubaddid-u ʔolma-
lamp-Gen-N light-Nom 3-dispel-Nom the-darkness
(The light of a lamp dispels the darkness)

b’. nuur-u mis‘baah-i-n y-ubaddid-u ʔolma-
lamp-Gen-N light-Nom 3-dispel-Nom the-darkness-Acc the-now
(A light of a lamp is dispelling the darkness now)

The only difference in meaning between (a&a’) is that (a’) no longer avails a generic reading; it can only be interpreted existentially. This explains why sentences that incorporate the indefinite CS can
only express generic readings, and refrain from expressing existential readings similar to their minimally contrasting counterparts with definite CS phrases.

6.3 Construct State and Reference to a Kind Genericity

Indefinite CS cannot be used as kind-denoting expressions. This becomes intuitively explicable if we remember that the indefinite CS denotes variable individuals to which the descriptive content of the CS applies; it cannot be used as a referential expression referring to a unique entity or kind. More precisely, a kind is definite by default as it is unique, and hence any nominal expression that denotes an indefinite entity is naturally incompatible with reference to a kind genericity. The indefinite CS behaves semantically as a predicate that picks out groups of individuals that satisfy the description content provided by the head and complement nouns in the phrase. Kind-referring NPs, however, behave in a way similar to that of proper names in rigidly referring to a specific definite kind or species (Krifka et al. 1995), and thus the indefinite CS is incompatible with this semantic rigidity.

The semantically definite CS, however, obtains the privilege of being referential, and thus can be used as object or kind-denoting expression. The data in (61) lay this point out.

(61) a. *dub-u l-baandaal saala wafak-i l-inviraadˤ*
bear-Nom the-panda at close-Gen the-extinction
(The panda bear is almost extinct)

b. *fadgarat-u z-zaytuun-i muntafirat-u-n fi f-faq-i l-ulawsatˤ*
tree-Nom the-olive-Gen common-Nom-N in the-east-Gen the-middle
(The olive tree is common in the Middle East)

c. *y-a?kul-u fa?r-u l-haql-i l-hubuub*
3-eat-Nom mouse-Nom the-field-Gen the-grains
(The field mouse eats grains)

d. *y-uwaadjih-u ?atfaal-u l-inanaabiib-i ma?akil-a s?ihiyya*
3-face-Nom babies-Nom the-tube-Gen problems-Acc health
(IVF babies face health problems)

The sentences (a&b) and (c&d) represent reference to a kind genericity and characteristic genericity, respectively. The pair in (a&b) uses K-level predicates, *saala wafak-i l-inviraadˤ ‘almost extinct’ and muntafir ‘common’, which are compatible with kind-referring NPs or CS phrases only.
These sentences express reference to a kind genericity. The statement in (a) talks about the species *dub-u l-baandaq* ‘The Panda’ by predicating a property that can only be true of a kind, rather than an individual; no individual panda can be almost extinct. The sentence in (b) also represents a reference to a kind sentence; in the sense that the property attributed *muntafir* ‘common’ can only be predicated to a kind, but never to an individual or groups of individuals. The genericity of the sentence in (b), for instance, can be informally verbalized as follows: ‘The kind *Olive Tree* exhibits the property of being common’. The sentences in (c&b) do not represent reference to a kind genericity; rather both sentences are characterizing sentences. This becomes clear if we notice that the whole sentence, rather than the generic CS per se, expresses genericity. Nonetheless, the incorporated CS subjects in both sentences are generic nominal expressions. Put differently, both sentences are characterizing sentences that express a generalization true not of an individual or a group of individuals, but rather of the kind itself.

That an indefinite CS cannot be used as a kind-referring expression can be supported by substituting indefinite CS for each minimally contrasting definite CS in each sentence in (61). Changing the definiteness status of CS in (a&b) renders both sentences syntactically unacceptable and semantically odd. An explanation for this is that the associated predicates are K-level predicates; therefore, the incompatibility of an indefinite CS with these predicates proves that they are not kind-referring expressions. It seems that the indefinite CS fails a major diagnostic test for identifying generic NPs, or CS phrases in this case, as set forward in Krifka et al. (1995). The sentences in (c&d) do not use K-level predicates; nevertheless, replacing the definite CS phrases with indefinite ones significantly changes the meaning of both sentences. More precisely, both sentences that express a generic interpretation will entertain an episodic reading as the only reading available. While (c) expresses a generalization about field mice in general, reporting that they eat grains, its indefinite CS counterpart reports an isolated event true of existential field mice being involved in eating grains event. Similarly, (d) expresses a generalization true of IVF babies in general, that they face health
problems, but its indefinite CS counterpart reports a non-generic proposition about existential IVF babies being involved in health problems at a particular time and place.

As mentioned in (section 5), kind-referring NPs exhibit two peculiar characteristics that distinguish them from other nominal expressions in the language under investigation. Kind-referring NPs can associate with S-level predicates, and can occur in sentences in past tense form; nonetheless, they can remain generic. These two points are used as a diagnostic test for identifying kind-denoting expressions in MSA, as well as English and other languages. This indicates that if definite CS phrases can be inserted in past tense sentences with S-level predicates without affecting their genericity, definite CS phrases receive further support to be established as kind-denoting expressions. This is explicated in (62).

(62) a. sˤawwar-naa dubb-a l-baandaa fi l-hadiiqat-i ?ams
    filmed-1.Pl bear-Acc the-panda in the-zoo-Gen yesterday
    (We filmed the Panda at the zoo yesterday)

    b. faahad-tu waθaʔaʔiqiyy-a-n ?an tays-i l-dgabal
    watched-1 documentary-Acc-N on goat-Gen the-mountain
    (I watched a documentary on the ibex)

The predicates in (62) are S-level, and both sentences are in the past tense. However, it is clear that both sentences do not express episodic propositions true of specimens of a kind. In fact, the CS phrases in (62) are generic in that they do not refer to a particular individual panda or goat, but rather to the species *Panda* and *Capra Ibex*, respectively, being realized through indirect kind reference (Pelletier 2010b)\(^{43}\).

I turn to address a question relevant to the type of kind reference that the definite CS expresses. It is known that in a CS phrase, the complement noun modifies the head noun, resulting in the head noun being more specific. This is realized in the type of kinds referred to in the sentences above. In all examples, the kinds referred to denote not whole species or kinds, but rather sub-kinds or sub-species. In (61), CS phrases are used as sub-kind-referring expressions; they denote sub-

\(^{43}\) See (chapter 2, section 2.1), for a discussion of indirect kind reference through representative object interpretation.
classes of kinds. In (a) for instance, reference is not made to the kind Bear, but to a sub-kind Panda. Similarly, in (b) reference is made not to the kind Tree, but to sub-kind Olive Tree. In (c) the kind Mouse is not the intended referent, but rather a sub-kind or class, namely the Field Mouse. The same rationale holds for IVF babies in (61.d), which is a true sub-kind of the naturally born Human Baby kind, and the Ibex in (62.b), which is a sub-kind of the kind Capra. Thus, it is plausible to suggest that definite CS phrases can be used as sub-kind-referring expressions due to the modification relation that holds between the head noun and its modifying complement, which yields in a more specific sub-kind-referring noun.

7. On Genericity in Bare NPs

The traditional Arab grammarians’ view of dividing nouns into two categories based on the presence/absence of the definite article al ‘the’ indicates that the restricted use of bare nouns in subject argument position is a grammatical by-product of the fact that unlike al-Ns, bare Ns in subject argument position are exclusively existential, see (section 3) above. This indicates that bare NPs are only allowed in subject argument position if and only if they denote existential entities. A direct repercussion of this insight is that bare NPs in MSA cannot be incorporated in characterizing sentences. This section argues against this claim, and shows that in certain syntactic and semantic environments bare NPs can be subject arguments in characterizing sentences. In particular, modified bare NPs, whose meaning is changed from existential denotation to quantificational denotation through the semantic interaction between the modifier and the head noun, are allowed in characterizing sentences. Interestingly, in such environment, the generic reading, rather than the existential, episodic reading entertained by unmodified bare NPs, is the more natural and intuitive reading. Another environment in which bare NPs can be used in characterizing sentences is habituals; sentences which generalize over variable situations, not variable individuals. After presenting some relevant syntactic facts about bare NPs in MSA (7.1), this - contra traditional’s - claim is furnished in (7.2).
7.1 The Grammatically Indefinite NP: Syntactic Facts

Traditional Arab grammarians, as well as modern Arab and non-Arab linguists, share a consensus that in general bare NPs in MSA are not permitted in subject position in both equational sentences or subject-predicate sentences, and SV verbal sentences. Bare NPs, however, are allowed in V/P-S order only - where P stands for a predicate of any syntactic form (e.g. Al-haramii 2005, d. 1302; Al-saamirraʔii 2011; Al-wardii 2008 , d.1348; Al-yalaaliimii 1993; Buckley 2004; Haywood and Nahmad 1962; ; Sibawayh 1988, d. 796; Wright 1898) . The examples in (63) lay this out.

(63) *a. raʤul-u-n fii d-daemon
        man-Nom-N  in  the-house
      (A man is in the house)

  a'. fii d-daemon-i raʤul
     in  the-house-Gen man
     (A man is in the house)

*b. raʤul-u-n dgaʔa
        man-Nom-N  came.3
     (A man came)

  b'. dgaʔa raʤul
    came.3 man
    (A man came)

*c. kalb-u-n xayr-u-n min ʔasad
    dog-Nom-N  better  from  lion
    (A dog is better than a lion)

  c'. ʔammata kalb-u-n xayr-u-n min ʔasad
     there  dog-Nom-N  better  from  lion
     (There is a dog (which) is better than a lion)

However, Arab grammarians cited more than forty different environments which permit bare NPs in subject position (Hassan 1980). In the majority of such environments, the bare NP subject does not introduce the sentence in a way or another- the order of the subject and predicate is reversed, or a question particle, preposition, existential there, etc. are inserted in the outset of the sentence.  

---

44 This observation raises an important question about whether Arabic disallows bare NP subject from introducing a sentence, i.e., being the first word in the sentence, or as the traditional view claims, a bare NP subject is not permitted in Arabic regardless of its position in the sentence except for these exceptional
Environments that permit bare NP subjects to introduce a sentence are those in which bare NP subjects are modified either by an adjective, or by a nominal expression in the construct state.

Consider (64) below.

(64) a. zawdżat-u-n t'ayyib-at-u-n tušiin-u šalaa sʰ-sʰišaab
wife-Nom-N good-f-Nom-N help-Nom on the-hardships
(A good wife helps in (alleviating) hardships)

b. dawlat-u salaam-i-n xayr-u-n min dawalat-i ḥarb
state-Nom peace-Gen-N better-Nom-N from state-Gen war
(A state of peace is better than a state of war)

The sentences in (64) incorporate bare NP subjects, which also introduce the sentences, yet the two sentences are grammatical because the two bare NP subjects are modified. This fact has strong bearing on the claim presented in (7.2) that bare NP subjects can be associated in characterizing sentences.

7.1.1 The Grammatically Indefinite NP: True or Pseudo Bare Noun

It has been emphasized above that the grammatical definiteness status of noun phrases in MSA is determined by the presence/absence of the definite article al ‘the’, where al-Ns are grammatically definite, and Ø-Ns are grammatically indefinite\(^{45}\). However, there is a strong debate going on about the exact description of grammatically indefinite noun phrases in MSA. In particular, this debate touches on the issue of the morphological status of indefinite nouns as to whether they are described as true bare nouns, where no indefiniteness marker appears, or pseudo bare nouns, where indefinite nouns are morphologically marked for indefiniteness by a suffix like –n, traditionally termed nunation or mimation (Fassi Fehri 2012). Indefiniteness asymmetry emerges from the fact that

\(^{45}\) Remember that semantic definiteness is determined based on the definiteness criteria though all simple definites are al-Ns, but not the other way around, see (chapter 3, section 4).
when a definite article is absent, a morpheme -\( n \) is suffixed to the stem, and this morpheme appears to be in complementary distribution with the definite article. The data in (65) explains this asymmetry.

(65)
\[
\begin{align*}
\text{al-qalam} & \quad \text{‘the pen’} & \text{qalam-u-}\!\!n & \quad \text{‘a pen-Case-N’} & \text{*al-qalam-u-}\!\!n & \quad \text{‘the pen-Case-N’} \\
\text{al-bint} & \quad \text{‘the girl’} & \text{bint-u-}\!\!n & \quad \text{‘a girl-Case-N’} & \text{*al-bint-u-}\!\!n & \quad \text{‘the girl-Case-N’} \\
\text{al-banaat} & \quad \text{‘the girls’} & \text{banaat-u-}\!\!n & \quad \text{‘girls-Case-N’} & \text{*al-banaat-u-}\!\!n & \quad \text{‘the girls-Case-N’} \\
\text{al-hiḍaara(t)} & \quad \text{‘the stones’} & \text{hiḍaarat-u-}\!\!n & \quad \text{‘stones-Case-N’} & \text{*al- hiḍaarat-u-}\!\!n & \quad \text{‘the stones-Case-N’}
\end{align*}
\]

Fassi Fehri (2012) summarized the three competing views regarding the status of nunation in Arabic. The view of old traditional Arab grammarians suggests that the –\( n \) on nouns is devoid of any syntactic or semantic function regarding definiteness/indefiniteness. Moscati (1964: 100) echoes this viewpoint; in his words, it is “impossible to identify any Proto-Semitic means of expressing definiteness or indefiniteness …there existed a mimation of nouns independent of any semantic function as regards definiteness or indefiniteness.” Brockelmann (1910), on the other hand, argues in favor of an indetermination view. A third view is posited by Kurylowicz (1950), in which he defends an idea that nunation expresses a general form of determination. However, a number of considerations have made some linguists argue against considering the morpheme -\( n \) an indefinite marker (e.g. Fassi Fehri 1993, 2012; Hallman 1999; Hoyt 2008). First, this morpheme appears on proper names that do not take the definite article (66). Second, the definite and indefinite markers are not in complementary distribution on the sound plural and the dual form since they can co-occur on the same noun (67). Third, as observed by Fassi Fehri (1993), the sound plural suffix -\( na \) shares some properties with -\( n \) morpheme in singular indefinite nouns, and singular definite nouns. Particularly, the final -\( n(a) \) and the definite article in all these forms disappear in the construct state (68).

(66) \( dʒaaʔ-u-\!\!n \)
\[
\begin{align*}
\text{came-3} & \quad \text{Zayd-Nom-N} \quad \text{(Zayd came)}
\end{align*}
\]

(67) a. \( \text{al-muslim-u-}\!\!n \)
\[
\begin{align*}
\text{the-Muslim-Nom.Pl-N} & \quad \text{(the Muslims)}
\end{align*}
\]

b. \( \text{as-sabbaah-aa-}\!\!n \)
\[
\begin{align*}
\text{the-swimmer-Nom.D-N} & \quad \text{(the two swimmers)}
\end{align*}
\]
However, there is a limitation on singular common nouns pertaining to the distribution of *al-* and -*n.* Fassi Fehri (2012) claims that such an asymmetry cannot be explicated if definite article *al-* is taken to express uniqueness/familiarity, while indefinite -*n* expresses non-uniqueness/novelty, simply because this characterization will exclude the co-occurrence of both markers on plural nouns, which is grammatical. This supports treating the indefinite marker as not the counterpart of the definite marker. It is a morpheme that represents different syntactic and semantic features, just like *un-* in Spanish, which is compatible with different values of number and gender. Thus, nunation cannot be taken as an indefinite marker, but it might be conditioned by indefiniteness. Therefore, the nunation marker will be treated as a type of declension that marks the majority of declinable nouns lacking *al* ‘the’ in MSA.

### 7.2 Genericity and Bare NPs in MSA Revisited

In MSA the canonical interpretation of bare nouns, both singular and plural, is existential, and thus their presence in a sentence forces an episodic reading. Consider the examples in (69).

(69) a. y- *alhu-tˤifl-u-n* *maʕa* qitˤatˤ- *at-i-hi*  
   3-play.Nom child-Nom-N with cat-f-Gen-his  
   (A child is playing with his cat)

b. t-*ufaahid-u* *niswat-u-n* *t-tilfaaz*  
   3.f-watch-Nom women-Nom-N the-television  
   (Some women are watching TV)

The only interpretation the sentences in (69) avail is episodic. This is enforced by the presence of the bare NPs in subject argument position, *tˤifl* ‘child’ and *niswa* ‘women’. Therefore, it seems that bare

---

46 See Muth (2008) for a discussion of the different functions of nunation in Arabic.

47 See Edzard (2006) for a brief exposition of declension in Arabic and other Semitic languages.
nouns in MSA block any probable generic reading of a sentence, and promote an existential reading as the only available reading. It is worth noting that replacing the bare nouns in (69) with their grammatically definite counterparts would change the semantics of both sentences, rendering them ambiguous between generic, habitual in particular, and nongeneric readings (see section 4.1 above).

Since both bare Ns and al-Ns occurring in the same syntactic and semantic environment bring about a clear distinction in interpretation, it is plausible to assume that bare nouns in MSA do not observe the other constituents of the sentence, particularly the predicate, and give rise to an episodic interpretation to any sentence in which they occur.

The generalization above is inaccurate since it is possible in MSA to use bare NPs subjects in true characterizing sentences. Consider the data in (70).

(70) a. muʔmin-u-n qawiyu-u-n xayr-u-n min muʔmin-i-n d’iʃiif
   believer-Nom-N strong-Nom-N better-Nom-N than believer-Gen-N weak
   (A strong believer is better than a weak believer)

b. kalb-u-n hayy-u-n ?af’dal-u min ?asad-i-n mayyit
   dog-Nom-N living-Nom-N better-Nom-N than lion-Gen-N dead
   (A living dog is better off than a dead lion)

c. kalb-u-n qawiyu-u-n qawiyy-u-n min
   dog-Nom-N strong-Nom-N more-Nom effectiveness-Gen-N than
   dgihaaz-i ?indaar
   system-Gen alarm
   (A strong dog is more effective than a security alarm system)

d. zuhuur-u-n dagmaiilat-u-n tusir-u-naa bi-s-safaada
   roses-Nom-N beautiful-Nom-N feel-Nom-us with-the-happiness
   (Beautiful roses make us feel happy)

e. haakim-u-n ʃ’aalim-u-n ʃaʃaddu ʃalaay f-faʃb-ii min l-ʔaʃdāa?
   ruler-Nom-N unjust-Nom-N tougher on the-people-Gen-N than the-enemies
   (An unjust ruler causes more harm to (his) people than enemies)

f. taʃliim-u-n ʃgayid-u-n yanhadˤ-u bi-l-ʔumam
   education-Nom-N good-Nom-N empower-Nom with-the-nations
   (Quality education empowers nations)

The sentences in (70) can only express generic readings, and the bare NPs used as subject arguments do not entertain existential denotation. (a) reports a generalization about any individual in the set of ‘strong believers’ stating that individual strong believers in general are attributed a
property of being better than individuals in the set of ‘weak believers’. Similarly, (b) expresses a
generalization about living dogs and dead lions in general. In (c), the generalization holds for strong
dogs in general, not any existentially quantified individual dog(s). The other sentences are interpreted
in the same fashion. These bare NPs do not exhibit the canonical existential denotation. This can be
demonstrated if we try to insert a sense verb like ḥaraa ‘I see’ in the sentences in (70), which seem
incompatible with such a verb. However, inserting ḥaraa in the sentences in (69) renders both
sentences felicitous. This is a puzzling insight because it stands in stark contrast to the traditional
conceived interpretation of bare NPs in MSA, set forth at the outset of this section.

However, astute readers must have noticed that the two predicates used in (69) are S-level
predicates, and hence the emergence of the episodic interpretation. The sentences in (70), however,
use I-level predicates. A plausible explanation for the intriguing behavior of bare NPs in (70) might
be attributed to their interaction with the predicates used, which are I-level predicates. More precisely,
the use of I-level predicates, like more effective and better, changes the meaning of bare NPs in MSA
from existential denotation into quantificational denotation. However, this claim seems untenable if
we observe that bare NPs can associate with I-level predicates, and nonetheless maintain their
existential denotation, which blocks the generic reading of the sentence. The sentences in (71) can
only be interpreted existentially, though the predicates incorporated are I-level predicates.

(71) a. y-uḥiḥb-u  faabb-u-n  fataah
   3-love-Nom young man-Nom-N girl
   ((There is) a young man (who) loves a girl)

   b. y-aʔrif-u  saaʔiq-u-n  l-faransiyya
   3-know-Nom driver-Nom-N the-French
   ((There is) a driver (who) knows French)

   c. ya-tˤiʔr-u  tˤaʔiʔr-u-n  fii  s-samaʔ?
   3-fly-Nom bird-Nom-N in the-sky
   (A bird is flying in the sky)

   d. ya-ʔkul-u  nimar-u-n  lahm
   3-eat-Nom tiger-Nom-N meat
   (A tiger is eating meat)
The data in (71)\textsuperscript{48} nullifies the claim that the asymmetrical behavior of bare NPs in (70) is attributed to the associated predicates, where S-level predicates give rise to existential interpretation of bare nouns, and I-level predicates promote a generic denotation. This claim can be further shown to be void when we consider sentences with S-level predicates that nevertheless entertain generic interpretation.

\begin{center}
(72) a. ya-\textgu\textasciitilde u ʔunaas-u-n  min muxtala\textgu\textasciitilde -i  l-ʔa\textgu\textasciitilde maar-i
fi  s\textgu\textasciitilde -s\textgu\textasciitilde uumaal
in  the-Somalia
(People of different ages starve in Somalia)
\end{center}

\begin{center}
b. y-\textswashaab-u ʔafxaas\textgu\textasciitilde -u-n  ʔabriyaa\textgu\textasciitilde -u  fii ħawaadi\textgu\textasciitilde -i
3- injure.Pass-Nom persons-Nom-N innocent.Pl-Nom in accidents-Gen
s-sayyr
the-traffic
(Innocent persons are injured in traffic accidents)
\end{center}

\begin{center}
c. ʔaf\textswashaal-u-n  mina l-djinsayn-i  y-am\textswashal\textasciitilde -uu-na  fii
sinn-i  l-xadaa\textswashal
age-Gen the-prematurity
(Premature babies of both genders get sick)
\end{center}

The sentences in (72) are true characterizing sentences though they use bare NPs as subject arguments and intriguingly S-level predicates - \textgu\textasciitilde starve’, \textswashaab ‘get injured’, \textswashal ‘get sick’. Each sentence obtains a generic reading to the exclusion of an episodic reading. This gives further support to refuting the claim that such sentences entertain non-existential interpretation in virtue of the predicate incorporated per se, particularly I-level predicates. In a nutshell, the predicate’s level per se does not seem to have significant bearing on the kind of interpretation assigned to certain constructions similar to those in (69&71). Thus, it seems that the question about this puzzling behavior of bare NPs in such constructions needs a different kind of explanation.

\textsuperscript{48} It is noteworthy that sentences (c&d) can express a generic reading if the bare NP subjects are changed to \textgu\textasciitilde subjects. This supports the claim proposed here that the existential denotation entertained by the bare NP subjects in such sentences is not a product of their interaction with the predicates associated.
A more plausible explanation can be arrived at if we investigate the internal structure of the bare NPs in (70&72). It is obvious that all the bare NPs are modified in one way or another. In (70.a), for instance, we are not talking about a believer, but rather a special type of believer - one who is physically and spiritually strong or weak. Similarly, (b) does not report about a dog and a lion, but a dog with a specific property, being alive, and a lion being dead. The same holds for the bare subject argument dog in (c) which is modified as being strong. The other NPs in (70) and those in the sentences in (72) are analyzed in the same fashion. Nevertheless, what bearing does modification have on the denotation of bare nouns in MSA? The answer to this question is not different from the answer proposed to a similar one raised above (section 6.2) about the indefinite in form CS. Modification changes the denotation of bare nouns from denoting relevant individuals existentially to denoting intensional sets of individuals to which the descriptive content of the noun and the modifier applies. Therefore, instead of *muhammad* ‘a believer’ being existentially calculated, the NP *muhammad qawiy* ‘a strong believer’ denotes the set of ‘individual strong believers’. Put differently, modified bare NPs are semantically indefinite in the real sense; while unmodified bare NPs are existential. This means that modified bare NPs, unlike their unmodified counterparts, are not tied to a particular individual or a group of individuals, but rather a modified bare NP denotes variable individuals. This is a crucial requirement for any nominal expression to be allowed as a subject argument in a characterizing sentence.

This analysis can be bolstered by the fact that removing modification from the sentences in (70&72) renders them either ungrammatical – in particular those in which bare NPs introduce the sentences, or significantly changes their meanings – particularly sentences in which bare NPs do not introduce the sentences- from expressing a generic reading to expressing an episodic reading. The

---

49 It is worth mentioning that Fassi Fehri (2012) mentioned that bare NPs in MSA can be used in generic sentences when modified. However, he did not provide an explanation for this insightful observation. Providing an explanation for this observation is an important goal of this section.

50 This dissertation will not explore why and how modification changes the denotation of bare NPs. These questions are very interesting and need thorough investigation.
ungrammaticality of such sentences with unmodified bare NP subject arguments can be explicated if we notice that they denote existentially calculated individuals, and are in SV sentences. MSA, as mentioned in (7.1), does not allow bare nouns denoting existential referents in SV order. To allow them in subject argument position, these NPs must fulfil two conditions. First, a syntactic condition enforces VS word order of the sentence for such NPs to be incorporated. Second, a semantic condition restricts their denotation to existentially calculated referents. Therefore, it appears that bare NPs in MSA can only be allowed in characterizing sentences if the head noun is modified. The grammaticality of incorporating modified bare NPs in SV structure, unlike their unmodified bare nouns, shows that they no longer denote existential individuals, hence can be used in characterizing sentences.

However, there are sentences which associate unmodified indefinite NPs, and still entertain generic readings exclusively. Such sentences pose some problems to the analysis proposed above. The sentences in (73) lay this point out.

(73) a. maa marr-a walad-u-n fii haadha l-hayy-i ?illaa
whenever passed-3 boy-Nom-N in this the-neighborhood-Gen but
raʔaa saw-3 ?awlaad-a-n y-alʕab-uu-na fi f-faariʕ
(Whenever a boy passes by this neighborhood, he sees boys playing on the street)

b. sindamaa y-azʔar-u ?asad-u-n t-afirr-u z-zaraafat-u
when 3-rooar-Nom lion-Nom-N 3.f-run-Nom the-giraffes-Nom
haariba escaping
(When a lion roars, giraffes run away)

c. ?aynamaa wallaa naadˤir-u-n wadḥ-a-hu raʔaa ...ʔafdaar
wherever turned-3 beholder-Nom-N face-Acc-his saw-1 trees
(Wherever a beholder turns his face, he sees trees)

The sentences in (73) use unmodified bare NPs in subject arguments and S-level predicates; nevertheless, they can only be interpreted generically. More accurately, the bare NPs in these sentences do not denote existentially calculated individuals, but rather intensional sets of individuals compatible with generic sentences. The sentence in (a) reports a generalization about situations which have a boy and boys in them, stating that if s is a situation of a boy passing by, s is a situation of that
boy seeing boys playing on the street. Similarly, (b) expresses a fact that if s is a situation of a lion roaring, then s is a situation of giraffes running away. As well, (c) generalizes over situations in which a beholder turns his face, saying that these situations are such that the beholder sees trees. It is clear that these sentences are habituals; they generically quantify over situations not entities. They include the so-called temporal when-clause, which quantifies over situation variables that must not be tied to a particular event (Krifka et al. 1995). Since habituals quantify over intensional sets of situations, it is reasonable to deduce that the individuals involved in these situations cannot be existentially calculated, simply because the very situations in habituals are not existentially calculated, but rather intensionally quantified. Situation entities in habituals are picked according to certain properties which decide whether a certain situation satisfies or obtains these properties or not, hence being in or out of the set. It is worth mentioning that habituals change the meaning of S-level predicates into I-level predicates by changing the eventive episodic verbal predicate into a stative predicate. This explains why the habituals in (73) use S-level eventive predicates.

In sum, bare NPs in MSA dominantly receive existential interpretation with both S-level and I-level predicates, and thus cannot be employed in characterizing sentences unless the denotation of these NPs is changed by modification from denoting existentially calculated, particular referents to referring to variable individuals. Another construction that allows bare nouns in generic sentences is habitual sentences, which generalize over situations. In other words, bare nouns in MSA can be only used in generic sentences when are quantificational, rather than referential.

8. Conclusion

In this chapter, the syntactic intricacies of nominal phrases in MSA, a morphologically rich language, were discussed. It has been shown that almost all noun forms can be used in generic sentences. No noun form, however, has been proved to be used as an exclusive generic marker. All

---

51 See Krifka et al. (1995) for a detailed discussion of the semantic interaction between habituals and stative predicates.
the nominal expressions compatible with generic sentences were shown to be compatible with other constructions, like episodic sentences.

The grammatically definite NP in particular appears semantically complicated in MSA. It is ambiguous among three competing denotations: denoting an existentially unique individual, an intensional set of individuals, and a kind or genus. The first two competing denotations were shown to be a reflection of the exact definiteness status of *al*-Ns in MSA. The former is compatible with an *al*-N being definite in the real sense, where an *al*-N refers to a familiar, identifiable, or unique individual. The latter, where an *al*-N denotes a set of individuals to which the descriptive content of the noun applies, is compatible with determining the *al*-N as semantically indefinite based on the definiteness criteria. An interesting consequence of this crucial distinction between definite in form and sense *al*-Ns and indefinite in sense *al*-Ns is refuting the two semantic repercussions of the traditional grammarians’ view of determining definites/indefinites based on the presence/absence of *al* ‘the’. According to the traditional view of definiteness in MSA, every *al*-N phrase is grammatically and semantically definite. Therefore, characterizing sentences, the bulk of genericity phenomenon in natural languages, are either completely absent in MSA, or the semantic architecture of characterizing sentences in MSA is idiosyncratic and language particular. This indicates that, contra to the structure of a characterizing sentence in other natural languages, it does not require a subject argument nominal expression that denotes variable individuals, but rather is tied to a particular individual. These two outcomes are awkward and sound unnatural to native speakers who overwhelmingly use characterizing sentences to express facts and generalizations about the world in a way similar to that used in other natural languages. Discriminating between *al*-Ns in terms of semantic definiteness excludes these two odd consequences. According to the claim elaborated in this chapter, characterizing sentences are not only ubiquitous in the language under investigation, but also conform to the semantic structure of characterizing sentences; in the sense that a generalization over variable individuals provided by an indefinite in sense *al*-N is expressed. The predicate used, along
with other syntactic and contextual clues, helps determine the exact denotation of the definite NP in a sentence, and the reading expressed, generic/non-generic.

It has been shown that both definite singular and definite plural NPs can be associated with kind level predicates, and thus both constructions appear equivalent in their use as kind-referring NPs. However, kind-referring nominal expressions in MSA, on par with those in other languages, must denote well-established kinds to be recognized as kind-referring NPs.

The construct state was shown to be compatible with generic sentences. The grammatically indefinite construct state can be used in characterizing sentences, whereas the grammatically definite construct state can be used in both characterizing sentences and reference to kind sentences. In the latter, the definite construct state denotes not a kind but a sub-kind. This is due to the indirect modification that holds between the head noun and its complement, giving rise to a nominal expression denoting a more specific sub-kind or sub-species.

Contrary to the traditional view, which dominantly identifies bare NPs as denoting individuals existentially, bare NPs have been shown compatible with generic sentences – characterizing sentences in particular - if modified, and with existential sentences if unmodified. It has been argued that modification changes the semantics of bare nouns in MSA from being existential, denoting existentially particular individuals, to be quantificational, denoting intensional sets of individuals. This semantic change qualifies bare nouns to be used in characterizing sentences, but not in reference to kind sentences because they become quantificational rather than referential.
CHAPTER 5: ON GENERICITY IN THE VERBLESS SENTENCE IN MSA

1. Introduction

The present tense verbless sentence in MSA exhibits a syntactic asymmetry in that copula ya-kuun ‘Pres-be’\(^{52}\) is obligatorily null in some sentences (1), optionally expressed in others (2), and obligatorily expressed in still other sentences (3).

(1) a. *ya-kuunu/Ø al-walad-u \(\text{Ø} \) al-walad-u \(\text{Ø} \) al-walad-u ʤaaʔiʕ Pres-be/Ø the-boy-Nom-N hungry (The boy is hungry)

b. *ya-kuunu/Ø al-ʔasad-u hayawaan-u-n muftaris Pres-be/Ø the-lion-Nom animal-Nom-N fierce (The lion is a fierce animal)

c. *ya-kuunu/Ø ʔahmad-u fī l-bayt Pres-be/Ø Ahmad-Nom in the-house (Ahmad is home)

(2) a. ta-kuunu/Ø as-samaaʔ-u sʕaafiyat-a/u-n fi sʕ-sʕayf Pres.f-be/Ø the-sky-Nom clear-Acc/Nom-N in the-summer (The sky is clear in summer)

b. ta-kuunu/Ø al-madaaris-u muylaqat-a/u-n fi sʕ-sʕayf Pres.f-be/Ø the-schools-Nom closed-Acc/Nom-N in the-summer (Schools are closed in summer)

c. ta-kuunu/Ø af-/awaarirš-u muzdahimat-a/u-n fi l-mudun Pres.f-be/Ø the-streets-Nom busy-Acc/Nom-N in the-cities (Streets are busy in cities)

(3) a. ta-kuunu/*Ø l-bint-u fi l-bayt-iʔawqaat-a \(\text{Ø} \) l-ʔahiira Pres.f-be/Ø the-girl-Nom in the-house-Gen times-Acc the-noon

---

\(^{52}\) In MSA, ya-kuun and its past and future forms are not treated as full verbs. They are treated as degenerate (not full verbs), functional verbs. Their presence, for instance, does not change a nominal sentence into a verbal one like other verbs. Subject-predicate sentences with or without them are treated as nominal sentences. This definition of a verbless sentence is shared by many traditional and modern Arab grammarians (e.g. Bakir 1980; Sibawayh 1988, d. 796; Fassi Fehri 1993, Benmamoun 2000). Following this tradition, this project treats any sentence with no full verb as a verbless sentence.
(The girl is home at noon times)

b. ya-kuunu*/Ø muhammad-u-n sahraan-a-n mina l-ʕaafirat-i hattaa l-fadgr
   Pres-be/Ø Muhammad-Nom-N staying up-Acc-N from the-ten-Gen until the-dawn
   (Muhammad stays up late from ten to dawn)

c. ya-kuunu/?Ø l-ʕariis-u muhrad-u-n fii laylat-i z-zafaaf
   Pres-be the-groom-Nom embarrassed-Acc-N in night-Gen the-wedding
   (The groom is embarrassed on the wedding night)

This asymmetry, this chapter emphasizes, reflects a semantic distinction between generic and nongeneric readings of the sentences in question. The claim is that sentences which prohibit an overt ya-kuun (1) vary in the kind of interpretation they express on par with a variation in the incorporated predicate, and the definiteness status of the NP subject. More precisely, sentences with S-level predicates give rise to an episodic interpretation only. Sentences with I-level predicates, however, are ambiguous between generic/non-generic interpretations based on the context which provides the correct denotation of the al-N subject used (see chapter 4, section 4.3.1). Sentences with K-level predicates admit reference to a kind genericity only. In contrast to sentences that prohibit the marker, sentences which optionally allow overt copula ya-kuun (2) obtain an important additional syntactic-semantic ingredient. This ingredient is an adverbial adjunct. These sentences are ambiguous between generic-habitual in particular- and nongeneric readings when copula ya-kuun is null. However, if ya-kuun is overtly realized only a generic reading can be obtained. Put differently, the overt presence of copula ya-kuun forces a generic interpretation as the only acceptable reading of the present tense verbless sentence in MSA. Sentences in which ya-kuun is obligatorily expressed (3) are compatible with a generic interpretation only. Their counterparts with unexpressed copula are either highly odd and unnatural, or require complex pragmatic machinery to validate them. In this case, they can express episodicy only.

In addition to this short introduction, the chapter incorporates another four sections. Section 2 briefly presents the syntactic facts of the verbless sentence in MSA, shedding some light on permissible and impermissible constructions, which in turn allows investigation of the semantic facts associated with these syntactic restrictions on the structure of the verbless sentence in MSA, the
presence/absence of copula \textit{ya-kuun} in particular. The focus of section 3 is to analyze and account for the syntactic facts presented in section 2 in semantic terms. A generic interpretation is attributed to those present tense verbless sentences with an overt copula \textit{ya-kuun}, but not to those without one. This meaning difference might suggest that those sentences that do not show an overt copula do not have one underlyingly. Section 4 investigates the negative verbless sentence. It will be shown that the two negative particles employed to negate the verbless sentences are in complementary distribution with regard to the presence/absence of copula \textit{ya-kuun} in the sentence. This distribution supports the semantic analysis proposed for the asymmetries of the verbless sentence based on generic/non-generic distinction. Finally, a brief conclusion is given in section 5.

2. The Verbless Sentence in MSA: Asymmetrical Behavior of the Copula

The verbless sentence in Arabic syntax has been a topic of discussion and debate since the time of Sibawaih (c.800 AD), the great traditional grammmarian (Al-Horais 2006). It is defined as a sentence that lacks an overt copula ‘\textit{ya-kuun} ’ (Pres-be) in the present tense. It only selects an initial NP, traditionally termed ‘almubtada?’ (topic), and a predicate ‘alxabar’ (comment), which can be NP, PP, AP, or AdP. These two constituents are termed the copula subject and copula complement in the modern literature on copula constructions (Curnow 2000; Dixon 2002; McKay 2010). These points are illustrated in (4) below.

\begin{enumerate}
  \item \textit{al-djaww-u} \textit{dgamiil-u-n} \textit{haaðhi} \textit{l-ʔayyaam}
  \text{the-weather-Nom nice-Nom-N these the-days}
  \text{(The weather is nice these days)}
  \item \textit{al-bint-u} \textit{fi} \textit{l-bayt}
  \text{the-girl-Nom in the-house}
  \text{(The girl is at home)}
  \item \textit{Zayd-u-n} \textit{tˤaalib}
  \text{Zayd-Nom-N student}
  \text{(Zayd is a student)}
  \item \textit{ʔahmad-u} \textit{hunaa}
  \text{Ahmad-Nom here}
  \text{(Ahmad is here)}
\end{enumerate}
The sentences in (4) clearly show the absence of a copula in present tense. (4.a) has an initial NP and an AP that reports something about the initial NP; (4.b-d) are similar to (4.a) in having an initial NP, but differ in the type of predicate each sentence has, PP, NP, and AdP are predicated to the initial NPs in (4b-d), respectively.

Inserting copula *ya-kuun* to any of the sentences in (4) will render it ungrammatical.

(5)  
(a) *ya-kuunu l-ḏjaww-u ḏamiil-a-n haaðihi l-ʔayyaam*
Pres-be the-weather-Nom nice-Acc-N these the-days  
(The weather is nice these days)

(b) *ta-kuunu l-bint-u fi l-bayt*
Pres,f-be the-girl-Nom in the-house  
(The girl is at home)

(c) *ya-kuunu Zayd-u-n t’aaalib*
Pres-be Zayd-Nom-N student  
(Zayd is a student)

(d) *ya-kuunu ?ahmad-u hunaa*
Pres-be Ahmad-Nom here  
(Ahmad is here)

The only difference between the grammatical sentences in (4) and their ungrammatical counterparts in (5) is the presence of copula *ya-kuun*. It is worth mentioning that the present and future forms of copula *ya-kuun*, (*kanna* & *sa-yakuun*), are obligatorily expressed in present and future tense verbless sentences. The data in (6) and (7) give examples of both the past and the future tenses, respectively.

(6)  
(a) *kaana l-ḏjaww-u ḏamiil-a-n l-baariha*
be,past the-weather-Nom nice-Acc-N the-yesterday  
(The weather was nice yesterday)

(b) *kaana-t l-bint-u fi l-bayt-i qablə saaʕa*
be,past-f the-girl-Nom in the-house-Gen before hour  
(The girl was at home an hour ago)

(c) *kaana Zayd-u-n t’aaalib-a-n fi s-sanat-i l-maadʔiya*
be,past Zayd-Nom-N student-Acc-N in the-year-Gen the-last  
(Zayd was a student last year)

(d) *kaana ?ahmad-u hunaa ?ams*
be,past Ahmad-Nom here yesterday  
(Ahmad was here yesterday)

(7)  
(a) *sa-yakuunu l-ḏjaww-u ḏamiil-a-n yadan*
Fut-be the-weather-Nom nice-Acc-N tomorrow
(The weather will be nice tomorrow)

b. sa-takuunu l-bint-u fi l-bayt-i ba'ida saaSa
Fut-f.be the-girl-Nom in the-house-Gen after hour
(The girl will be at home in an hour)

c. sa-yakuunu Zayd-u-n t'aalib-a-n fi s-sanat-i l-qaadima
Fut-be Zayd-Nom-N student-Acc-N in the-year-Gen the-coming
(Zayd will be a student next year)

d. sa-yakuunu ?ahmad-u hunaa yadan
Fut-be Ahmad-Nom here tomorrow
(Ahmad will be here tomorrow)

The data in (6) and (7) make it clear that in the past and future tenses, the copula must be
phonologically overt if such sentences are to be grammatical. In fact, deleting the copula in any
sentence in (6) or (7) automatically renders that sentence unacceptable.

(8) a.*al-djaww-u djamiil-u-n l-baariha
the-weather-Nom nice-Nom-N the-yesterday
(The weather was nice yesterday)

b.*al-bint-u fi l-bayt-i qabla saaSa
the-girl-Nom in the-house-Gen before hour
(The girl was at home an hour ago)

c.* al-djaww-u djamiil-u-n yadan
the-weather-Nom nice-Nom-N tomorrow
(The weather will be nice tomorrow)

d.* ?ahmad-u hunaa yadan
Ahmad-Nom here tomorrow
(Ahmad will be here tomorrow)

As evident from the data presented so far, it can be observed that there is a kind of asymmetry
in the surface structure of the verbless sentence in the present tense on one hand, and its past and
future counterparts on the other. In general, no copula is allowed in the present tense; whereas, the
copula seems obligatory in both the past and future tenses. This asymmetry is not peculiar to the
verbless sentence in MSA. It is a relatively common crosslinguistic phenomenon in languages
entertaining verbless constructions. Dixon (2002:17), in his typological study of over than 250
Australian languages, notes that “a copula is likely to be omitted if reference is to present time, but
included – with the appropriate tense suffix – for past or future reference.” Nordlinger and Sadler (2007) also claim that in a large number of languages, the absence of present tense copula is correlated with the use of overt non-present tense copula forms.

However, investigating more data from MSA reveals that in certain constructions copula ya-kuun presents itself as a phonologically optional constituent in present-tense verbless sentences.

Consider the data in (9) below.

(9) a. al-djaww-u baarid-u-n fi f-fitaaʔ?
   the-weather-Nom cold-Nom-N in the-winter
   (The weather is cold in winter)

   a'. ya-kuunu l-djaww-u baarid-a-n fi f-fitaaʔ?
   Pres-be the-weather-Nom cold-Acc-N in the-winter
   (The weather is cold in winter)

   b. al-muslimuuna kuramaʔ-a-u fii ramadʔaan
   the-muslims.Nom generous.Pl-Nom in Ramadan
   (Muslims are generous in (the month of) Ramadan)

   b'. ya-kuunu l-muslimuuna kuramaʔ-a fii ramadʔaan
   Pres-be the-muslims.Nom generous.Pl-Acc in Ramadan
   (Muslims are generous in (the month of) Ramadan)

   c. at-tanaqul-u sˤaʔb-u-n fi sˤ-sˤabaah
   the-moving-Nom hard-Nom-N in the-morning
   (Commuting is hard in the morning)

   c'. ya-kuunu at-tanaqul-u sˤaʔb-u-n fi sˤ-sˤabaah
   Pres-be the-moving-Nom hard-Nom-N in the-morning
   (Commuting is hard in the morning)

(9.a,b&c) and their minimal pairs in (9.a',b'&c') employ the present tense to express propositions whose truth or falsity is evaluated in the present, past, and future; they express generalizations over a group of situations. In such a construction, the phonological realization of copula ya-kuun is optional. This optionality of the overt presence of copula ya-kuun is interesting, and calls for a thorough investigation to construct a plausible account that would explain the asymmetric behavior of copula ya-kuun in verbless sentences. The optionality raises the question of whether or not there is an underlying, “silent” ya-kuun in sentences where ya-kunn in not expressed.
The asymmetrical behavior of copula *ya-kuun* is further explicated in sentences in which its overt presence is obligatory. This is explained in (10).

(10) a. *ya-kuunu kalb-ii dgaαʔ-iʔ-a-n fī l-masaa?*
Pres-be dog-my hungry-Acc-N in the-evening
(My dog is hungry in the evening)

  a’. ?kalb-ii dgaαʔ-iʔ-u-n fī l-masaa?
dog-my hungry-Nom-N in the-evening
(My dog is hungry in the evening)

b. *ya-kuunu l-μušallim-u sāʔiid-a-n fī l-ʔidʒażaat*
Pres-be the-teacher-Nom happy-Acc-N in the-holidays
(A teacher is delighted in holidays)

b’. *a l-μušallim-u sāʔiid-a-n fī l-ʔidʒażaat*
the-teacher-Nom happy-Acc-N in the-holidays
(A teacher delighted in holidays)

c. *ta-kuunu l-marʔat-u ʕas‘abiyyat-a-n waqt-a d-dawrat-i f-ʃahriyya*
Pres.f-be the-woman-Nom nervous-Acc-N time-Acc
the-cycle-Gen the-monthly
(A woman is nervous during the menstrual cycle)

c’. ?l-marʔat-u ʕas‘abiyyat-u-n waqt-a d-dawrat-i f-ʃahriyya
the-woman-Nom nervous-Nom-N time-Acc the-cycle-Gen the-monthly
(A woman is nervous during the menstrual cycle)

The presence of copula *ya-kuun* in (10.a-c) is the preferred structure as their ‘copula-less’ minimal pairs in (10. a’-c’) are only good under special pragmatic circumstances that well be discussed in section (3) below.

After presenting relevant syntactic facts about verbless sentences in MSA, I turn to discuss these intricacies from a semantic perspective.

3 The Semantics of the Verbless Sentence in MSA

This section restricts itself to discuss the semantics of the present tense verbless sentence in MSA. It proposes a semantic account for the syntactic distribution of copula *ya-kuun* based on a generic/nongeneric distinction. To be more specific, I propose the generalization in (11) to account for the correlation between the overt realization of copula *ya-kuun* and the generic/nongeneric interpretation.
The distribution of copula *ya-kuun* in present tense verbless sentence

a. Copula *yakuun* must be licensed by a spatio/temporal adverbial.
b. Present tense verbless sentences with such an adverbial come in two kinds:

1. Copula *yakuun* is optional, with a generic-nongeneric meaning difference based on the choice of the predicate and the definiteness status of the copula subject.

2. Copula *yakuun* is required based on the choice of the predicate and its interaction with the adverbial.

c. The presence of a spatio/temporal adverbial does not guarantee the presence of *yakuun*.
d. Present tense verbless sentences without adverbial licensors, and therefore without copula *yakuun* can receive varied generic/nongeneric interpretations based on the predicate used as copula complement and the nominal form used as copula subject.

In order to illustrate how this generalization about the syntactic-semantic correlation of copula *ya-kuun* and the generic/nongeneric interpretation works, three sections are presented below. Section (3.1) will be dedicated to investigate the semantic intricacies of verbless sentences with obligatorily absent copula *ya-kuun*. Section (3.2) will investigate the semantics of verbless sentences with optionally expressed/unexpressed copula. Section (3.3) will probe the semantics of verbless sentences with obligatory overt copula.

3.1 Obligatorily-Absent Copula

The canonical structure of the present tense verbless sentence in MSA is that it does not allow copula *ya-kuun* in its Phonetic Form (PF), as opposed to its past and future tense counterparts in which the copula is obligatory. Therefore, the exact semantics of the present tense verbless sentence depends on the nominal form of the copula subject and the copula complement. There is a correlation between the predicate’s level, S-level, I-level, or K-level predicates, and the NP subject, semantically definite/indefinite, object-referring NP, or kind-referring NP, on the one hand; and the overall interpretation of the sentence, generic or episodic, on the other hand. The data in (12) illustrate the correlation between the use of S-level predicates and the presence of an episodic interpretation of the sentence as the only available interpretation. Copula *ya-kuun* is incompatible with these sentences.

(12) a. *laylaa yadbaan-a l-yawm*
    Laila angry-f the-day
(Laila is angry today)

b. \(\text{al-ʕussʕuurr-}\)u \(\text{ʕalaₕ} \) fₕadⱰaₕii bₕayt-\(i\)-naa
   the-sparrow-Nom on the-tree in house-Gen-our
   (The sparrow is in the tree in our house)

c. \(\text{kalb-}u\)-n s'ayyir-\(u\)-n ?amaam-a bₕayt-\(i\)-naa
   dog-Nom-N small-Nom-N front-Acc house-Gen-our
   (A small dog is in front of our house)

d. \(\text{ar-riih-}\)u qawiyy-\(a\)-\(u\)-n hₕaadₜā l-maₕₜa\(ʔ\)
   the-wind-Nom strong-f-Nom-N this the-evening
   (The wind is strong this evening)

The only reading available for the sentences in (12) is episodic. The truth of the propositions depends on reference to some interim situations and contextual clues; thus, these propositions cannot be construed generically. Episodicity is contributed by the use of S-level predicates, “highly temporary states or events” (Carlson 2005:16), and the use of proper nouns or existentially calculated NP subjects\(^{53}\). The sentence in (a), according to Carlson’s account of S-level/I-level predicate distinction, reports an isolated fact about Laila being angry at a particular time and location. This property is ephemeral, and thus applies to a spatio-temporal slice of the individual Laila. Similarly, (b) reports an accidental fact about an existentially unique sparrow as being in a specific tree on a specific time. The sentences in (c&d) give similar episodic interpretations\(^{54}\) in that both sentences refer to particular, existentially calculated individuals on specific situations. Therefore, it seems plausible to claim that the insertion of S-level predicates enforces unequivocal episodic reading of a sentence.

\(^{53}\) Individual-level predicates are treated as part of the family of generics. However, at least three different views on the distinction of I-level and S-level predicates are argued for in the literature. For Carlson (1977, 2005), genericity of I-level predicates is triggered by their interaction with the NP subjects, whethere they are existential or generic. Kratzer (1995), however, accounts for the distinction by positing that stage-level (but not individual-level) predicates have an extra eventuality/occasion argument. A third view is argued for in Cherchia (1995). Cherchia’s view is that all predicates, stage-level and individual-level, have eventuality argument, but what makes an individual-level predicate individual-level is that the argument must be bound by a generic operator. For Cherchia, all sentences that make use of individual-level predicates are generics of a type, hence his term inherently generics.

\(^{54}\) It is worth noting that verbless sentences in MSA disallow bare indefinite nouns in copula subject position unless modified.
However, MSA entertains present tense verbless sentences which do not tolerate an overt copula \( ya-\text{kuun} \), and still receive ambiguous readings. In fact, incorporating I-level predicates brings about ambiguous sentences. In particular, a verbless sentence with I-level predicate is ambiguous between a more prominent generic reading and a less salient episodic reading. The context gives the required clues to tease these two readings apart based on the correct denotation of the \( al-N \) subject, definite or indefinite (see chapter 3, section 5). The examples in (13) illustrate this point further.

(13) a. \( al-\text{layl-u} \quad \text{haalik} \)
\( \text{the-night-Nom} \quad \text{dark} \)
\( \text{(The night is dark)} \)

b. \( al-\text{faakihat-u} \quad \text{mufiidat-u-n} \quad li-s\\’ihhat-i \quad l-\text{ʔinsaan} \)
\( \text{the-fruits-Nom} \quad \text{useful-Nom-N} \quad \text{for-health-Gen} \quad \text{the-human} \)
\( \text{(Fruit is good for human health)} \)

c. \( a\,f-\text{fitaa?-u} \quad \text{baarid} \)
\( \text{the-winter-Nom} \quad \text{cold} \)
\( \text{(Winter is cold)} \)

d. \( al-\text{buum-u} \quad \text{fu\?m} \)
\( \text{the-owls-Nom} \quad \text{ominous} \)
\( \text{(Owls are ominous)} \)

The sentences in (13) do not accept an overt copula \( ya-\text{kuun} \), and any attempt to insert it in any of these sentences will render the sentence ungrammatical. The asymmetry that these sentences create is that not tolerating a copula \( ya-\text{kuun} \) notwithstanding, these sentences, on their more salient readings, do not express temporal episodic propositions. They all express atemporal regularities about groups of individuals or objects. These sentences are dubbed lexical characterizing sentences in Krifka et al. (1995); who emphasize that lexical sentences do not generalize over events, but rather over “characterizing properties of individuals” (p.17); and are dubbed inherently-generic sentences in Chierchia (1995).

Each sentence in (13) is ambiguous between generic/nongeneric interpretations, though the generic reading is more natural and salient. As has been discussed in chapter 4, the ambiguity of sentences incorporating I-level predicates and \( al-N \) subjects emerge from the alternate denotation of the \( al-N \) subject. If the \( al-N \) subject is semantically definite, the sentence expresses an episodic,
isolated fact about that particular referent of the subject NP. If the *al-N*, however, is shown to be indefinite, the sentence is generic. The context helps determining the in/definiteness of the *al-N* based on the definiteness criteria. If the sentence is uttered out of the blue, as the case with the sentences in (13), the episodic interpretation is highly unnatural and requires complex pragmatic machinery to be validated. This is because the *al-N* subjects in these sentences are more naturally interpreted to denote any individual, rather than a particular, definite individual, which satisfies the descriptive content of the noun. This ambiguity is likely due also to the lack of explicit morpho-syntactic generic markers. It is worth mentioning that in the absence of a generic marker; every generic sentence has an episodic interpretation that can be gleaned based on some pragmatic machinery. More accurately, every generic sentence, in the absence of a generic marker, is derived from an episodic sentence that can be promoted contextually, and based on other pragmatic tools. In brief, in MSA, and many other languages, every generic sentence that lacks a generic marker can be interpreted episodically, but no episodic sentence entertains an alternative generic interpretation. This explains why most generic sentences in MSA are ambiguous between generic/nongeneric readings in the absence of an explicit context which can limit the denotation of the *al-N* subject to a particular individual, or to any individual to which the descriptive context of the noun applies.

Careful scrutiny of the above sentences shows that the predicates used are all essential I-level predicates, rather than temporal S-level predicates. In fact, these sentences, on their generic readings, express laws or regulations that are not gleaned statistically as the inductive approach devises, but rather are seen as reflecting a causal rule or regulation that represents a certain structure out in the world (Carlson 1995). Such inherently generic sentences are better explained by the rules-and-regulations approach as some scholars have observed (Greenberg 1998, Cohen 2002, to name but two). The sentence in (a), for example, provides two alternative readings. On one reading, which is more prominent, the sentence reports a property of nights in general being dark; it says that the property of ‘be dark’ is an inherent characteristic of night. This sentence can be plausibly restated as follows: To be night is to be dark. This is the generic reading. The other reading is episodic; it reports
an isolated fact about an existentially unique night salient in the context domain as being ‘dark’.

Similarly, (b) can be interpreted either generically as reporting usefulness-to-human-health property of fruit in general, or as existentially, on which particular fruit is described as being useful to humans. The sentences in (c&d) can be interpreted in the same fashion. The context and other pragmatic machinery determine the exact denotation of the al-N subject, hence the interpretation of the verbless sentence with I-level predicate\(^{55}\).

However, this does not explain the obligatory absence of copula ya-kuun in such sentences. A plausible explanation for this asymmetry is attested in Carlson’s (1995) claim that sentences that establish their genericity on a rule or regulation out in the world do not entertain a quantificational semantic structure. The truth and falsity of these statements are not statistical or quantificational.

According to Krifka et al. (1995), these lexically characterizing sentences entertain the generic operator GEN, however, GEN does not generalize over situations, but rather over individuals to which the predicated properties are attributed. Copula ya-kuun, however, either allows for or grammatically marks genericity in sentences that express generalizations over events, situations, or cases. It can only surface in habituals, and is deemed absent in lexically characterizing sentences, see (3.2) below. In other words, copula ya-kuun functions as a morphosyntactic marker of habituality in MSA. This analysis also explains why ya-kuun is absent in verbless sentences with episodic readings. There is no generalization over situations in these sentences; rather there is reference to a particular situation.

As the case with verbal sentences that incorporate K-level predicates (see chapter 4, section 5), verbless sentences incorporating K-level predicates promote reference to a kind genericity as well. The sentences in (14) flesh this out.

(14) a. \(\text{o-namir-u \qquad \text{mina} \qquad \text{\theta-\text{\textit{adiyyaa}}}\)\]
\[\text{the-tiger-Nom \quad from \quad the-mammals}\]
(The tiger is a mammal)

\(^{55}\)See chapter 4, section 4.3, for a discussion of the semantics of sentences with I-level predicates, for which verbless sentences with the same predicate level is no different.
b. al-fiīl-u ḥayawāa-nu-نتشر salaa wa fak-i l-ingiraadʕ
the-elephant-Nom animal-Nom-N on verge-Gen the-extinction
(The elephant is on the verge of extinction)

c. ḥawz-u l-hind-i muntasīr-u-n fi l-manaatʕiq l-ʔistiwaaʔiyya
nut-Nom the-India-Gen common-Nom-N in the-areas the-tropical
(The cocoanut is common in tropical areas)

d. al-ʕaqīr-u mina l-qawaaridʕ
the-mouse from the-rodents
(The mouse is a rodent)

e. al-fiyyalat-u ḥayawaanaa-nu-نتشر ʕaafiba
the-elephants-Nom animal.Pl-Nom-N herbivore
(Elephant are herbivores)

The sentences in (14) represent reference to kind generics. In (a), for example, the K-level predicate mina ʔ-thadīyyaat ‘be a mammal’ can only select a generic NP an-namir ‘the tiger’ which denotes the species itself. Similarly, (b) uses another K-level predicate salaa wa fak-i l-ingiraadʕ ‘on the verge of extinction’, which can only apply to a kind or subkind; no single elephant or a group of elephants can be on the verge of extinction, only the kind itself can hold this property. The sentences in (c-e) receive the same reference to a kind interpretation, and this is clear from the K-level predicates used.

In sum, the semantics of the verbless sentence with obligatorily null copula is in general determined by the predicate level and the noun phrase incorporated. According to this analysis, there are unequivocal correlations between S-level predicates and episodic interpretation, I-level predicates and ambiguous generic/nongeneric interpretation, and K-level predicates and reference to a kind interpretation. All the examples discussed in this section disallow copula ya-kuun to surface. In fact, inserting the copula in any of the examples will render it ungrammatical. Copula ya-kuun cannot surface in both episodic sentences and lexically characterizing sentences that represent laws, rules, or regulations.

However, some verbless sentences of a certain structure optionally allow copula ya-kuun to surface, creating a semantically more complex structure that will be investigated in (3.2) below. It will be shown that ya-kuun can only mark genericity in quantificationally generic sentences, habituals in particular.
3.2 Optionally-Present Copula

Sentences where copula ya-kuun can optionally surface are slightly more complicated than those discussed in (3.1) above, where ya-kuun is deemed phonologically null for those sentences to be acceptable. In these sentences, however, copula ya-kuun can be either realized overtly, where only a generic interpretation could be obtained, or remain unpronounced, rendering the sentence ambiguous between a generic/nongeneric interpretation depending on the predicate, NP subject, and other pragmatic considerations. Some of the sentences in (9), repeated in (15) for convenience, lay this point out.

(15) a. al-djaww-u baarid-u-n fi f-fitaaʔ?
   the-weather-Nom cold-Nom-N in the-winter
   (The weather is cold in winter)

   a'. ya-kuunu l-djaww-u baarid-a-n fi f-fitaaʔ?
   Pres-be the-weather-Nom cold-Acc-N in the-winter
   (The weather is cold in winter)

   b. al-muslimuuna kuramaʔ-u fii ramad’aan
   the-muslims.Nom generous.Pl-Nom in Ramadan
   (Muslims are generous in (the month of) Ramadan)

   b'. ya-kuunu l-muslimuuna kuramaʔ-a fii ramad’aan
   Pres-be the-muslims.Nom generous.Pl-Acc in Ramadan
   (Muslims are generous in (the month of) Ramadan)

   c. at-tanaqul-u s’ās-b-u-n fi l-mudun
   the-moving-Nom hard-Nom-N in the-cities
   (Commuting is hard in cities)

   c'. ya-kuunu at-tanaqul-u s’ās-b-a-n fi s’sābaah
   Pres-be the-moving-Nom hard-Acc-N in the-morning
   (Commuting is hard in the morning)

The sentences in (15)\textsuperscript{56} have additional ingredient, time/place reference adverbials, which might function as a situations provider\textsuperscript{57}. Therefore, these sentences do not quantify over individuals,

\textsuperscript{56} Astute readers must have noticed that the case of the predicate is nominative in the absence of copula and accusative when expressed. This syntactic asymmetry has no semantic bearing on the ultimate interpretation of the verbless sentence, hence will not be probed here. This issue has been studied widely in the literature of the verbless sentence in MSA (e.g. Bakir 1980; Fassi Fehri 1993; Benmamoun 2000).
but rather over situations provided contextually by the time/place reference adverbials. In other words, these are true habituals, rather than lexically characterizing sentences. The first pair (a&a’) gives a generalization about winter situations, rather than about weather individuals. Similarly, the pair in (b&b’) generalizes over Ramadan situations, not over Muslim individuals. The pair in (c&c’) express a generalization about morning situations which have commuting events saying that these situations are such that commuting is hard in them. Before delving in the semantic intricacies of these problematic structures, I briefly investigate the semantics of habitual sentences, focusing on the approach that will be adopted to account for the semantic intricacies of these structures.

3.2.1 Habituals: Two Contrasting Views

Habituals are sentences that give generalizations about events or situations. The truth conditions of a habitual sentence are tied to the regularity of occurrence that is asserted in the sentence, rather than to the occurrence of an event at a particular place and time. This regularity is not established iteratively since in habituals the event does not need to occur several times in succession to be true. This is a condition for another class of sentences known as iteratives in the literature (Carlson 2006). Iteratives refer to a closed sequence of events, whereas habituals denote a potentially unlimited set of events (Lenci 1995). The major difference that sets habituals apart from other characterizing sentences, dispositional sentences for examples, is that for a habitual sentence to be true there must be some instance events that have already occurred or observed. The sentence ‘Mary takes the bus to school’ would be false if no single instance of Mary-taking-the-bus-to-school has occurred. Dispositional sentences, however, do not entertain such a provision; they do not require a single instantiation of the eventuality to be true. A dispositional sentence like ‘This printer prints 100

57 By situation, I mean some event-like structure. Events, situations, and eventualities are not differentiated here. Treating the adverbial adjuncts as situations provider is one way of accounting for habituality meaning in verbless sentences modified by adverbials. Another way is to take the copula as the one that contributes a set of eventualities for the habitual operator to quantify over.

58 In the absence of spatio-temporal adverbials, these sentences, with expressed yakkun in particular, are incompatible with adverbs of quantification like ‘always’. Those with unexpressed ya-kuun, however, sound odd with such adverbs, and need complex pragmatic machinery to validate them.
copies per minute’ would be unequivocally true even if that printer has not been used yet. Mari, Beyssade, and Del Prete (2013b:56) emphasize that “Generics - at least their habitual subvariety… seem to involve reference to plural events…to have already occurred and another plurality of such events to be expected to occur in the future.” Habituals in general share two basic properties: an extensional property realized through recurrence of events, and an intensional property of generalization associated with stativization (Vogeleer 2012).

In the literature, two competing approaches have been elaborated to account for habituals: the quantificational approach (Krifka et al. 1995), which explicitly considers habituals a subvariety of generics, and the non-quantificational analysis proposed in (Boneh & Doron 2008, 2010, 2013; Kratzer 2007; Lenci 1995), which treats habituals, at least a certain class of habituals, as a distinct linguistic phenomenon.

The quantificational analysis assumes that habituals are a subvariety of generics, in the sense that they share the same tripartite semantic structure, which incorporates an unexpressed modal generic operator, but instead of quantifying over a set of individuals, GEN in habituals quantifies over a set of situations provided by context. This operator has quasi-universal force, something like usually (Krifka et al. 1995). Habituality is taken as a subvariety of genericity based on its rule-like nature. This analysis accounts for the exception tolerance of habituals, a benchmark characteristic which they share with other generic types. In other words, the predicate does not need to hold in all situations for the habitual sentence to be true.

A quantificational habitual is argued for, Vogeleer (2012) claims, when the sentence provides the ingredients of a tripartite quantificational structure. She claims that in quantificational habituals, the tripartite structure is triggered by an adverbial adjunct: a manner adverbial, a locative PP, or an instrumental PP, which she dubbed a phrase with cyclic iteration like after dinner, in summer, on Thursdays, etc. Most proponents of the quantificational analysis of habituals believe that the modal quantifier in quantificational habituals is GEN itself, the generic operator. Therefore, a quantificational habitual like (16) can be formally represented as follows:
(16) Aziz eats an apple after dinner.
(17) GEN $s \{\text{after dinner(s)} \& \text{Aziz in } s; \exists x \\{\text{apple } (x) \& \text{eat } (s, \text{Aziz, } x)\}\}$

The sentence in (16) is felicitous because it provides a restrictor for the generic operator to quantify over. When this restrictor is not explicitly provided, the sentence is rendered ungrammatical. Consider (18).

(18) *Aziz eats an apple.
    GEN $s \{-------- \& \text{Aziz in } s; \exists x \\{\text{apple } (x) \& \text{eat } (s, \text{Aziz, } x)\}\}$

(16) entertains a tripartite structure where GEN restricts to the situation, and the nuclear scope can fall under GEN. Therefore, (16) is felicitous since GEN restricts to Aziz-after-dinner situations, and an apple takes a narrow scope, hence falling under GEN. This allows for a different apple per Aziz-after-dinner-eating situation. However, (18) is infelicitous since no restrictor for GEN is provided. This sentence can be only interpreted with the existential quantifier having scope over the generic operator, which is infelicitous because the same apple cannot be eaten habitually. However, if we replace the singular apple with the plural form apples, the sentence can be felicitous since apples can be eaten habitually.

The non-quantificational analysis, however, assumes that the basic ingredient of habituality, plurality of events, is not provided by a covert quantifier, but rather by the lexical verbs themselves, which “are born as plurals” (Kratzer 2007: 269). The verb fall, Kratzer claims, could relate plural individuals to plural events. On this view, verbs denote sums of events. Vogeleer (2012) argues that unlike quantificational sentences, which provide a restrictor for a quantifier over situations, non-quantificational habituals lack such restrictor. This is bolstered by the fact the in quantificational sentences, the reading does not change drastically if the covert quantifier is made overt: ‘Mary eats breakfast with a fork= Mary usually eats breakfast with a fork’. The only change is that the latter loses the episodic reading. However, with non-quantificational sentences like ‘*Mary eats an apple vs. Mary usually eats an apple’, adding an overt adverbial cancels the ban on telic predicates like eat, and renders the sentence acceptable.

59 This claim has first been suggested in Krifka, M. (1992).
Lenci (1995) claims that quantificational and non-quantificational habituals are not one homogenous class. He presents four contexts that reveal the asymmetric behavior of the two types, and proposes based on these asymmetries that “these classes are really two well-distinguished semantic types, that act very differently in some relevant contexts” (P. 144). These contexts include:

a. **Perceptual reports:**

Unlike quantificational habituals (19), non-quantificational habituals (20) lose their habitual meaning when embedded in a perceptual report. Consider the examples below which are based on Lenci’s (1995) examples.

(19)

a. John often smokes.
   b. The lawyer saw John often smokes.

(20)

a. John smokes.
   b. The lawyer saw John smokes.

(19.b) preserves its habitual interpretation, whereas (20.b) can only receive an episodic interpretation. It is noteworthy that the argument is somewhat complicated by the fact that the eventuality requires a change in Tense/Aspect morphology. ‘The lawyer saw John smokes’ does not report the direct perception of smoking event in which John took part. Rather, it can be paraphrased as ‘The lawyer saw that John smokes’, where the lawyer perceived something that indicates that John is indeed a smoker.

b. **Singular indefinite object NPs:**

Quantificational habituals with indefinite object NPs entertain two alternative readings: one with an indefinite NP taking wide scope over the quantifier, and another with a narrow scope. Consider (21):

(21)

a. Mike often repairs a car.
   b. Mike repairs a car.

The sentence in (a) is ambiguous between two readings. On the first reading, the indefinite NP is non-specific. This indicates that there is a different car repaired by Mike. The second reading, the wide scope reading, the same car is repaired. However, non-quantificational sentences with indefinite NP
objects are not ambiguous. The only reading that (b) can receive is episodic; it is anomalous on a habitual reading.

c. **Impossible habitu**als:

With respect to the predicate level, quantificational habitu**als** are possible with any predicate, whereas non-quantificational habitu**als** are not. If all habitu**als** are one class, in which a generic operator splits the sentence into two parts, which is often associated with tense-aspectual markers, Lenci argues, why some sentences with exactly the same structure do not promote a habitual reading. This is illustrated in (22).

(22) a. Mike sings.
    b. Mike sleeps.
    c. Mike wakes up.

While (a) is perfectly acceptable as a non-quantificational habitual, it is not clear why (b&c) are not. Notice that providing (b&c) with an adverbial renders them correct on a habitual reading.

(23) a. Mike sleeps after dinner.
    b. Mike wakes up early.

d. **Negation**:

Quantificational habitu**als** give rise to two alternative readings under negation. This is due to the scope effect of negation on the quantifier.

(24) a. Mike does not often smoke.
    b. Mike does not smoke.

Sentence (a) is ambiguous between two readings: (1) ‘It is not true that John smokes most of the time’, (wide scope of negation), (2) ‘It is often the case that John is not smoking’, (narrow scope of negation). However, scope ambiguity is not attested in non-quantificational sentences; (b) simply means that ‘John is not a smoker’.

Proponents of the non-quantificational account reject analyzing habitu**ality** as a subtype of genericity, and instead propose treating it by means of Hab operator, which is distinct from the generic operator GEN. Hab is an intensional summation of events. It is “a modalized existential quantifier over sums of events” (Boneh and Doron 2013:177). However, Non-quantificational
approach advocates do not treat all habituals in the same way; they seem to distinguish between habitual sentences based on the syntactic ingredients incorporated, and whether a situations provider ingredient is available or not. In habituals which place explicit restrictions over events, GEN can be used; whereas, habituals that do not have explicit restrictor, another operator, Hab, is involved (Boneh and Doron 2013). Vogeleer (2012) claims that neither approach is successful in accounting for both habitual types. She proposes that labor be divided between the two approaches instead. Quantificational approach deals with quantificational habituals, and non-quantificational approach deals with non-quantificational habituals or “verbal-plural” habituals.

Since the verbless structure we are dealing with is quantificational, in the sense that it involves a phrase with cyclic iteration that provides the required restrictor for a tripartite-structure analysis, the semantic account that will be provided will adopt the quantificational account presented in Krifka et al. (1995).

3.2.2 The Semantics of Verbless Sentences with Optionally Null Copula

Verbless sentences with optionally unexpressed copula (25) are ambiguous between habitually generic reading and episodic reading based on the context and other pragmatic clues which determine the definiteness status of the al-N subjects and objects. These sentences provide the time/place reference ingredient functioning as a restrictor in a tripartite quantificational structure. In MSA, this additional ingredient can take different syntactic forms, which share one significant characteristic in that they provide situations.

(25) a. al-ʔardˤ-u zaahiyat-u-n fi r-rabiiʕ
   the-land-Nom bright-f-Nom-N in the-spring
   (The land is bright in spring)

   b. al-ʤaww-u baarid-u-n fi j-fitaaʔ
      the-weather-Nom cold-Nom-N in the-winter
      (The weather is cold in winter)

   c. al-ʤaww-u haar-u-n Ŧinda δˤ-ʔahiira
      the-weather-Nom hot-Nom-N at the-noon
      (The weather is hot at noon)

   d. al-layl-u fadiidu δˤ-ʔulmat-i fi j-fitaaʔ
The sentences in (25) are ambiguous. On one reading, the sentence in (a), for instance, reports a habitual generalization, stating that all relevant spring situations are such that the land is bright in them. This reading can be formally represented using a tripartite structure as in (26). The other reading, though less salient and requires complex pragmatic machinery to be validated, is episodic. On this reading, the sentence reports an isolated fact about a certain land being bright at a particular spring situation.

(26) GEN s [spring (s) ∃x [land (x) & bright (s, x)]

Informally, (26) says that usually, whenever it is spring, the land is bright. Notice that this formula allows different lands on par with different spring situations. This is because the generic operator takes scope over the existential variable. This structure represents the canonical quantificational habitual sentence. As for (b), the sentence also entertains two readings. The first reading, which sounds more plausible and neutral to pick by native speakers unless some pragmatic considerations push towards ignoring it, renders the sentence generically habitual. It expresses a generalization over winter situations pertaining to how the weather regularly looks in these situations. The second interpretation allows for an episodic reading of the proposition where that sentence refers to a situation, rather than generalizes over situations. It says that a particular contextually salient individual weather, in a particular contextually prominent winter situation, is cold. Likewise, the sentence in (c) quantifies over noon situations, rather than over weather individuals. On its habitual reading, the sentence reports a generalization over noon situations, stating that in these situations the weather is hot. The second interpretation allows for an episodic reading of the sentence, where that
sentence refers to a particular temporal situation. This less salient reading reports an accidental fact about a specific noon situation, where the weather is hot. (d) is not a lexically characterizing sentence about nights in general, but rather a habitual sentence that generalizes over winter situations, summing up these situations as involving nights which are habitually extremely dark. The other episodic reading of (d) is no different from those of (a-c). The sentences in (e&f) receive the same ambiguous interpretations between habitual/episodic readings.

An interesting question pertaining to the classification of these sentences as habituals may arise. More precisely, why do we consider these sentences habituals, rather than lexically characterizing sentences? Is the presence of reference to time adjuncts sufficient to treat such sentences as habituals? The answer to this question is twofold. First, as mentioned in chapter 4, section 5, only reference to kind generics and habituals are compatible with S-level predicates. Other characterizing sentences are compatible with I-level predicates only. Close scrutiny of the predicates used in (25) reveals that all the predicates are S-level; bright, cold, hot, extremely dark, in demand, and necessary are S-level predicates, and nonetheless the sentences entertain generic readings. In fact, replacing these predicates with I-level predicates renders verbless sentences with exactly the same structure ungrammatical. This is illustrated in (27).

(27) a. *az-zaraafat-u t’awiil-at-u-n fi f-fitaaʔ
the-giraffe-Nom-N tall-Nom-N in the-winter
(The/A giraffe is tall in winter)

b. *an-namir-u hayawaan-u-n muxat’i’at’u-n fi l-mudun
the-tiger-Nom animal-Nom-N striped-Nom-N in the-cities
(The tiger is striped in cities)

c. *alkalb-u δakiyy-u-n ʕinda δ-i’-d’aiira
the-dog-Nom intelligent-Nom-N at the-noon
(The dog is intelligent at noon)

d. *al-fyalaat-u hayawaanaat-u-n ʕaafibat-u-n fi s-safar
the-elephants-Nom animal.Pl-Nom-N herbivore-Nom-N in the-travel
(elephants are herbivores while traveling)

What the sentences in (27) show is that I-level and K-level predicates are incompatible with this structure where quantification is not over individuals, but over situations. Being tall, striped,
herbivore, and intelligent are incompatible with habitual readings, where sentences generalize over sums of situations. These predicates denote inherent properties essential for the characterization of the individuals/kinds they describe; they cannot hold only in certain situations. They hold in all situations that have the individuals to which these properties are attributed. Since the sentences in (27) use S-level predicates, and nonetheless entertain a generic reading, we can plausibly argue that these sentences are habituals, rather than lexically characterizing sentences.

Another piece of evidence comes from deleting the adjuncts in both (25&27). Deleting the adverbial adjuncts in (27) not only changes the sentences’ grammatical status rendering them all acceptable, but also affects their semantics, in that they become characterizing sentences. ‘The giraffe is tall’, ‘The tiger is striped’, ‘The dog is intelligent’, and ‘Elephants are herbivores’ are all true characterizing sentences in both MSA and English. However, deleting the adjuncts in (25) brings about drastic change in meaning. Consider some of the sentences in (25), repeated below without adjuncts in (28).

(28) a. al-ʔardˤ-ʔu zaahiya
    the-earth-Nom bright
    (The land is bright)

    b. al-djaww-u haar
        the-weather-Nom hot
        (The weather is hot)

    c. al-haqaaʔib-u d’arurriyya
        the-bags-Nom necessary
        (Bags are necessary)

    d. al-muθalladgaat-u matˤluuba
        the-ice cream demanded
        (Ice cream is in demand)

The sentences in (28) are good on episodic readings only. The generic reading is not available. (a) reports an accidental fact about a certain piece of land being bright at a particular location and time. The sentence no longer entertains a generic reading as its minimal counterpart with time reference adjunct does. Similarly, (b) gives an isolated fact about the weather being hot at a particular place and time. (c) expresses a fact about particular bags being necessary. The sentence in (d) reports an
accidental fact about particular ice cream being in high demand at specific time and location. It is clear that deleting the situations provider adjuncts in such construction brings about significant change in meaning, in the sense that the habitual reading is blocked and only the episodic interpretation is available. This is due to the S-level predicates incorporated which are compatible with an episodic interpretation only, see (3.1) above. This indicates that the sentences in (25), on their more salient reading, are habituals that require restrictors to be felicitous. It is noteworthy that the sentences in (28) are compatible with adverbs that force an existential interpretation like haadíhi l-.Chain ‘these days’.

In sum, verbless sentences with optionally null copula are ambiguous between a habitual reading, which is more intuitive and salient, and an episodic reading, which requires complicated pragmatic machinery to be validated. The adverbial adjuncts provide restrictors for such quantificational habituals, without which the semantics of these sentences show dramatic change in acceptability and meaning.

3.2.3 The Semantics of Verbless Sentences with Optionally Overt Copula

Verbless sentences in MSA have counterparts that optionally incorporate a copular verb. However, the semantics of verbless sentences with overt copula is different from that of verbless sentences with unexpressed copula. This explains why MSA entertains two different structures to express what inaccurately appears to be the same meaning. If both structures express the same meaning, then economy would rather give rise to one structure over the other, which by the course of time tends to disappear, or change its function. It is noteworthy that although the two structures are distinct in the type of semantics they express, verbless sentences with phonologically expressed copula are less frequent than those with phonologically silent copula. This conforms to Curnow (2000) observation that if there are two or more constructions encoding relatively the same information and speakers feel that one adds more information not found in the other, the one adding additional meaning is considered less basic and the other most basic. To investigate the type of information conveyed by verbless sentences with overt copula, and how different it is from that
expressed by their minimally contrasting sentences with null copula, the sentences in (25) are repeated below in (29) with expressed copula.

(29) a. ta-kuunu  al-ʔardʕu  zaahiya-t-a-n  fi  r-rabiiʕ
Pres.f-be  the-land-Nom  bright-f-Acc-N  in  the-spring
(The land is bright in spring)

b. ya-kuunu  al-ʤaywʔ-u  baarid-a-n  fi  f-fitaaʔ?
Pres-be  the-weather-Nom  cold-Acc-N  in  the-winter
(The weather is cold in winter)

c. ya-kuunu  al-ʤaywʔ-u  haar-a-n  ʕinda  δʕ-δʔahiira
Pres-be  the-weather-Nom  hot-Acc-N  at  the-noon
(The weather is hot at noon)

d. ya-kuunu  al-ʕayʔ-u  fadiid-a  δʕ-δʔulmat-i  fi  f-fitaaʔ?
Pres-be  the-night-Nom  extremely-Acc  the-dark-Gen  in  the-winter
(The night is very dark in winter)

e. ta-kuunu  al-muθallaḍqaat-u  maʔluubat-a-n  fi  l-harr
Pres.f-be  the-ice cream  demanded-Acc-N  in  the-heat
(In hot (weather) ice cream is in demand)

f. ta-kuunu  al-ʔaqqaʔib-u  dʔaruuriyyat-a-n  fi  s-safar
Pres.f-be  the-bags-Nom  necessary-Acc-N  in  the-travel
(Bags are necessary while traveling)

The main semantic difference the sentences in (29) entertain is that the episodic reading is completely blocked. The only reading available is generic, habitual in particular. This indicates that copula ya-kuun functions as a formal, habitual marker, in the sense that it can only live in the shade of a habitual sentence. More precisely, its overt presence gives rise to a habitual reading of the sentence, and simultaneously excludes the episodic reading that the verbless sentence with unexpressed copula has along with the habitual reading. The sentence in (a), for example, can only be interpreted as generalizing over sums of groups of spring situations, not reporting an isolated fact about a particular spring situation. Similarly, (b) does not entertain an episodic interpretation, where an accidental fact about a particular noon situation is expressed, but rather quantifies over noon situations in general. Similarly, (c) abstracts away from expressing an episodic interpretation, and avails the generically habitual reading as the only reading possible. The sentences (d-f) are analyzed in the same fashion.
In sum, verbless sentences with time/place reference ingredient seem more complex than those lacking such an ingredient. In the former, this ingredient provides the required situation restrictor for a tripartite, quantificational structure, which gives rise to a habitual reading of the sentence, rather than a lexically characteristic reading, which the latter entertains. However, verbless sentences with time reference ingredient are not completely homogeneous. This is because two minimal counterparts, with/without overt copula, are used in the language to encode roughly the same information. Nonetheless, these two versions appear not identical, but rather one structure- the one with null copula- is pragmatically ambiguous between habitual/episodic interpretations. Whereas, the one with overt copula does not show this ambiguity and gives rise to an unequivocal habitual reading at the expense of the episodic reading, which is excluded. This explicit exclusion of the episodic reading of the sentence is the additional semantic ingredient that a verbless sentence with overt copula obtains over its minimally contrasting counterpart with unexpressed copula. Put differently, when copula ya-kuun is overtly realized, it functions as a habitual marker that restricts the interpretations of the sentence to a generic one, just like other genericity markers that have been observed in other languages (e.g. Carlson 1989; Krifka et al. 1995; Greenberg 2002).

3.3 The Semantics of the Verbless Sentence with Obligatorily Overt Copula

Some verbless sentences in MSA require an overt copula to be acceptable (30.a-d), or to be able to express habituality (30.e&f). Copula ya-kuun in these sentences is obligatorily expressed. Consider the sentences in (30)\(^60\).

\[(30) \text{a. ya-kuunu}/\emptyset \quad xaalid-u-n \quad naa?im-a-n \quad min \quad l-Sifaa?-i \quad ?ilaal \quad l-fadgr} \]

Pres-be/Ø Khalid-Nom-N asleep-Acc-N from the-evening-Gen to the-dawn
(Khalid is asleep from evening to dawn)

b. ya-kuun/\emptyset \quad s'adiiq-ii \quad fi \quad bayt-i-hi \quad ?ayyaam-a \quad l-fut'al}

Pres-be/Ø friend-my in house-Gen-his days-Acc the-holidays
(My friend is at his house on holidays)

c. ta-kuunu/\emptyset \quad ?ummi \quad fi \quad l-Samal-i \quad mina \quad s-saabiñat-i}

Pres.f-be/Ø mother-my at the-work-gen from the-seven-Gen

\(^60\) The question mark (?) indicates that the sentence without copula is odd on habitual reading, and needs complex pragmatic machinery to validate its episodic reading.
(My mother is at work from seven to three p.m.)

d. ya-kuunu/*Ø ya-zan-u masruur-a-n ʔinda intihaa?-i s-sanat-i
   Pre-be/Ø  Yazan-Nom  delighted-Nom-N  at  end-Gen  the-year-Gen
d-diraa-asiyya
   the-school
   (Yazan is delighted at the end of a school year)

e. ya-kuunu/?Ø l-ʕiidayl-u dʕaxm-a-n ʔi ʔa-ʔaaniyat-i min ʕumr-ih
   Pres.be/Ø  the-calf-Nom  large-Acc-N  in  the-second-Gen  of  age-its
   (A calf is large on its second year)

f. ya-kuunu/?Ø l-fariis-u muhradg-a-n ʔi ʕiaylat-i z-zafaaf
   Pres.be  the-groom-Nom  embarrassed-Acc-N  in  night-Gen  the-wedding
   (The groom is embarrassed on the wedding night)

The sentences in (30) are true habitual sentences in that they report generalizations about situations. In (a), for instance, a generalization is reported about from-evening-to-dawn situations stating that these situations are such that Khalid is asleep in them. The sentences in (b&c) are habituals in that they express generalizations about holiday and seven-to-three situations, respectively. In the former, it is reported that these situations are such that the speaker’s friend is usually home, and in the latter, the speaker’s mother is usually at work. Similarly, the sentence in (d) expresses a habitual generalization about end of school year situations; these situations are such that Yazan is delighted in them. The sentence in (e) expresses a generalization about calves’ two-year-old situations, saying that usually calves are large in such situations. The sentence in (f) quantifies over wedding night situations, reporting that in general a groom feels embarrassed in these situations.

Therefore, it seems that the overt presence of copula ya-kuun unequivocally marks habituality in verbless sentences. This function of copula ya-kuun is not different from its function discussed in (3.2.3) above. The only major difference is that the overt presence of copula ya-kuun there is optional, i.e., verbless sentences without ya-kuun are grammatically acceptable and semantically interpretable. The sentences in (30), however, do not easily allow null copula ya-kuun, and when ya-kuun is null a sentence is rendered somewhat odd and can express episodicity only. Deleting the copula in (30.a-d), marked with an asterisk (*), renders them semantically uninterpretable, and somewhat
ungrammatical. For those in (e&f), marked with a question mark (?), deleting the copula excludes the habitual reading and provides an episodic reading only. This indicates that the copula in these sentences has to be overt for such sentences to express habituality meaning. This adds to the distinction between these sentences and those with optional expressed/unexpressed copula.

This is an interesting asymmetry. Precisely, what is it that renders ya-kuun optional in the sentences in (29) discussed in (3.2.3) and obligatory in the sentences in (30)? An answer to this question can draw on the predicates used in both verbless structures. Put differently, the predicates used in the sentences in (29) are different from those used in (30), (a-d) in particular. This answer, however, is unsatisfactory because both structures express habituality, and hence both invariably use S-level predicates; I-level predicates are incompatible with habituals. Scrutiny of the time reference adverbials and their interaction with the predicates, however, can be helpful in accounting for this asymmetry. The time reference adverbials used in (30.a-d) are incompatible with the S-level predicates used in the absence of copula ya-kuun. Put differently, in these sentences, episodic readings are hard to derive because the time reference adverbials are incompatible with the ephemeral predicates used. In (a), for example, the time reference adverbial min l-$\text{ʕifaa}\$?-i $\text{ʔiilaa l-fadgr}$ ‘from evening to dawn’ is incompatible with the S-level predicate $\text{naaʔim}$ ‘asleep’ on episodic reading. More accurately, it is infelicitous to report an episodic fact about Khalid being asleep from evening to dawn unless the speaker has been watching Khalid sleeping the whole period. In this case, a verbless sentence in the past tense would be used. Expressing an episodic sentence about Khalid being ephemerally asleep at the time of expression does not require giving a time reference, unless the speaker intends to report that Khalid usually sleeps from evening to dawn. In this case, copula ya-kuun has to be phonologically expressed to mark this reading. The absence of copula ya-kuun gives the impression that this sentence can be interpreted episodically, and this contradicts the interaction between the predicate and the
time reference adverbial which does not promote an episodic reading. In the present tense, the sentence in (30.a) does not entertain an episodic reading in which an accidental fact is reported about Khalid being asleep on a particular from-evening-to-dawn situation. Notice that the episodic reading of the sentence would be completely felicitous if the past or future copulas kaan/sa-yakuun ‘was/will-be’ were used. It is worth noticing that in such structure, deleting the time reference adverbial (31.a) renders the sentence without copula acceptable, and its counterpart with overt copula ungrammatical.

(31) a. xaalid-u-n naaʔim
    Khalid-Nom-N asleep
    (Khalid is asleep)

    b. *ya-kuumul xaalid-u-n naaʔim
    Pres-be/∅ Khalid-Nom-N asleep
    (Khalid is asleep)

The asymmetry in (31) is explicable. The sentence in (a) is felicitous as it reports an accidental fact about Khalid being asleep on a particular time and location. The sentence in (b), however, is ungrammatical because notwithstanding it lacks time reference ingredient necessary for providing situations, it incorporates copula ya-kuum which is a habitual marker and requires situations to quantify over.

The sentence in (30.b) requires an obligatorily overt copula for the same reason discussed above. The time reference adverbial ?ayyaam-a l-ʕutʔal ‘on holidays’ is incompatible with the S-level predicate fi bayt-i-hi ‘in his house’ on an episodic reading. The truth conditions of an episodic reading of this sentence require that the speaker’s friend be home on all time slices or sub-situation intervals of the particular holiday situation for this sentence to be true. In other words, the sentence on an episodic reading is false if the speaker’s friend is not in his house for any period on that holiday. The verbless sentence in the present tense with a time reference adverbial cannot express this episodic reading because the sentence can only be true of a particular slice of the situation, not the situation as a whole. However, the habitual reading of the sentence does not entertain this condition, and can be true even though a not-being-home-on-holiday occurs sometimes. Therefore, the absence of copula ya-kuum creates an asymmetry in that the structure of the sentence gives rise to an episodic reading
along with the habitual reading, but the incompatibility of the time reference and the S-level predicate does not entertain an episodic reading. This asymmetry renders the sentence odd and unnatural, hence copula *ya-kuun* is deemed obligatory to exclude the unwanted episodic reading, and promotes the habitual reading as the only available reading. The same analysis is applicable to the other sentences in (30.c&d). Deleting the copula in (30.e&f), however, changes the meaning of these sentences drastically, changing them from habitual sentences to episodic ones, hence the obligatory presence of the copula.

Therefore, it seems that the overt presence of copula *ya-kuun* gives an unequivocal habitual interpretation of the verbless sentence in MSA. The absence of copula *ya-kuun*, however, gives rise to a variety of readings summarized in the generalization in (11) above.

A slightly similar generalization pertaining to the distinction in the phonological status of copula in the verbless sentence, and the kind of interpretation expressed, generic/episodic, holds in Hebrew. Shlonsky (1997) claims that there is a correlation between the presence of copula and the exclusively generic interpretation on the one hand, and the absence of copula and the prominence of an episodic interpretation on the other hand. He argues that a sentence with absent pronominal copula gives rise to an episodic interpretation, while its minimally contrasting counterpart with an expressed pronominal copula forces a generic interpretation only. This is demonstrated, according to Shlonsky, by the compatibility of null copula with S-level predicates, and overt copula with I-level predicates in the verbless sentences in Hebrew.

The analysis proposed above grounded in generic/non-generic distinction can be drawn to account for another syntactic asymmetry in the negative verbless sentence in MSA. This asymmetry is discussed below in section 4.

## 4. Manifestations of Genericity in the Negative Verbless Sentence in MSA

Genericity in the verbless sentence in the present tense is not only relevant to the presence/absence of copula *ya-kuun* in the affirmative sentence; it also manifests itself in the negation
of this construction. In MSA two negative particles are used to negate a verbless sentence, \textit{laa} and \textit{laysa} (both roughly mean not). These particles are in complementary distribution. The data in (32) flesh this out.

\begin{align*}
\text{(32) a. } & \text{Muhammad-}u \quad \text{t'aalib-}u-n \quad \text{mu\d{g}tahid} \\
& \text{Muhammad-Nom-N student-Nom-N diligent} \\
& \text{(Muhammad is a diligent student)} \\
\text{a'. Muhammad-}u \quad \text{laysa/laa} \quad \text{t'aalib-}a-n \quad \text{mu\d{g}tahid} \\
& \text{Muhammad-Nom-N NEG student-Acc-N diligent} \\
& \text{(Muhammad is not a diligent student)} \\
\text{b. } & \text{al-bint-}u \quad \text{fii} \quad \text{l-matbax} \\
& \text{the-girl-Nom in the-kitchen} \\
& \text{(The girl is in the kitchen)} \\
\text{b'. al-bint-}u \quad \text{laysa-t/laa} \quad \text{fii} \quad \text{l-matbax} \\
& \text{the-girl-Nom NEG -f in the-kitchen} \\
& \text{(The girl is not in the kitchen)} \\
\text{c. } & \text{al-mu	extsc{u}allim-}u \quad \text{yad}\text{\d{b}aan} \\
& \text{the-teacher-Nom angry} \\
& \text{(The teacher is angry)} \\
\text{c'. al-mu	extsc{u}allim-}u \quad \text{laysa/laa} \quad \text{yad}\text{\d{b}aan} \\
& \text{the-teacher-Nom NEG angry} \\
& \text{(The teacher is not angry)} \\
\end{align*}

The data in (32) show that the verbless sentence can combine with the negative particle \textit{laysa}, but incorporating \textit{laa} renders the sentence ungrammatical. It is noteworthy that the sentences in (a-c) and their negative counterparts (a'-c') are episodic sentences. This is clear from the predicates used; \textit{diligent}, \textit{in the kitchen}, and \textit{angry} are S-level predicates. I turn to investigate negative verbless sentences with I-level and K-level predicates, and see whether the same fact holds. This is explained in (33).

\begin{align*}
\text{(33) a. } & \text{al-kalb-}u \quad \text{laysa/laa} \quad \text{yabiyy} \\
& \text{the-dog-Nom NEG stupid} \\
& \text{(The dog is not stupid)} \\
\text{b. } & \text{ar-rad\textsc{gul}-}u \quad \text{laysa/laa} \quad \text{?un\d{b}aa} \\
& \text{the-man-Nom NEG female} \\
& \text{(A man is not female)} \\
\text{c. } & \text{an-namir-}u \quad \text{laysa/laa} \quad \text{mina} \quad \text{z-zawaahif} \\
& \end{align*}
the-tiger-Nom NEG from the-reptiles
(The tiger is not a reptile)

d. al-fii-l-u laysa/*laa mungaridˤ
the-elephant-Nom NEG extinct
(The elephant is not extinct)

The sentences in (33) incorporate both I-level and K-level predicates. It is worth noticing that these sentences do not tolerate copula *ya-kuun* because they lack time reference ingredients necessary for habitual marker *ya-kuun* to quantify over. Therefore, *ya-kuun* is obligatory absent in these sentences. The data in (33) clearly show that laysa, not laa, is compatible with negative verbless sentences with obligatorily absent copula. This fact does not observe the predicate’s level; (a-b) incorporate I-level predicates, and (c-d) use K-level predicates.

An interesting question to be addressed here is relevant to the syntactic behavior of negative verbless sentences with time reference adjuncts. More precisely, does the verbless sentence with optionally expressed/unexpressed copular verb behave syntactically in the same fashion under negation, and what are the semantic reflexes? The data in (34&35) lay this down.

(34) a. al-djaww-u baarid-u-n fi f-fitaaʔ?
the-weather-Nom cold-Nom-N in the-winter
(The weather is cold in winter)

a'. al-djaww-u laysa/*laa baarid-a-n fi sˤ-sˤaif
the-weather-Nom NEG cold-Acc-N in the-summer
(The weather is not cold in summer)

b. al-ʔardˤ-u zaahiya-t-u-n fi r-rabiiʕ
the-land-Nom bright-f-Nom-N in the-spring
(The land is bright in spring)

b'. al-ʔardˤ-u laysa-t/*laa zaahiya-t-a-n fi f-fitaaʔ?
the-land-Nom NEG -f bright-f-Acc-N in the-winter
(The land is not bright in winter)

c. al-muθallađaqat-u matˤluubat-u-n fi l-harr
the-ice cream-Nom demanded-Nom-N in the-heat
(In hot (weather) ice cream is in demand)

c'. al-muθallađaqat-u laysa-t/*laa matˤluubat-a-n fi l-bard
the-ice cream NEG -f demanded-Acc-N in the-cold
(In cold (weather) ice cream is not in demand)
The data in (34) represent the verbless sentence with optional absence of copula *ya-kuun*. As mentioned above, in a verbless sentence with an optionally null copula, both readings, generic and episodic, are available. The negative particle *laysa*, rather than *laa*, is compatible with this structure. This indicates that *laysa* is the unmarked negative particle in MSA, which is typically used in the absence of copula *ya-kuun*, a structure taken as more frequent and to a certain extent unmarked.

The data in (35) show that when copula *ya-kuun* is phonologically overt, and thus only a habitual interpretation is available, the negative particle *laa* can be used; whereas, the use of the negative particle *laysa* renders the sentence ungrammatical. This same analysis is applicable to sentences where *ya-kuun* is obligatorily expressed. This indicates that the two particles are in complementary distribution in terms of the habitual/episodic readings of the verbless sentence when the sentence does not express an ambiguity. Put differently, MSA utilizes a specific negative particle to mark unequivocal habitual reading of a sentence, and another unmarked particle to mark episodicity and generic/nongeneric ambiguity in the verbless sentence construction. However, *laa* negative particle and *ya-kuun* habitual marker are used in other constructions not relevant to
genericity like prohibitive sentences, and modal sentences, respectively. In other words, the use of laa and ya-kuun in MSA conforms to a typological classification of linguistic expressions which are used mainly to mark genericity and meanwhile can be used to express other functions in different contexts (Dahl 1995). Concisely, both laa and ya-kuun are necessary but not sufficient habituality markers in the verbless sentence in MSA.

In sum, the negative verbless sentence in MSA is expressed in three different constructions with only two interpretations. A construction without time reference adjunct expresses generic/nongeneric interpretation based on the predicate used and the semantic definiteness status of the subject NP, (32&33); it can only combine with the negative particle laysa. A construction with copula ya-kuun unexpressed, (34), avails the two generic/nongeneric interpretations. In this case, negative particle laysa is used to mark the ambiguous interpretation, while negative particle laa is incompatible. A construction with copula ya-kuun overtly realized (35), which can receive a generic interpretation only, conforms to the exclusive use of the generic negative particle laa

5. **Conclusion**

In this chapter, the semantics of the present tense verbless sentence has been probed. The claim this chapter made is that there is a semantic explanation for the syntactic asymmetry represented by the obligatory absence of copula ya-kuun in some sentences as opposed to the optional/obligatory presence of the copula that other sentences exhibit. It has been shown that this syntactic asymmetry can be plausibly explained semantically by positing a generic vs. nongeneric interpretation. It has been argued that sentences which do not incorporate adverbial adjuncts of time or place do not allow overt ya-kuun. Since these sentences lack adverbial adjuncts, an essential ingredient for expressing quantificational habituality, copula ya-kuun is incompatible with these sentences, hence deemed obligatorily absent. The exact interpretation of these sentences depends on the predicate and the NP subject. Put differently, S-level predicates give rise to an episodic reading of the sentences. I-level predicates give rise to an ambiguous generic/non-generic reading based on the definiteness status of the subject NP; a semantically definite NP subject promotes a non-generic
reading, and an indefinite subject is compatible with a generic reading. K-predicates are compatible with reference to a kind genericity. In other words, the semantics of verbless sentences lacking time reference adjuncts is similar to their verbal counterparts discussed in chapter 4.

Sentences in which copula ya-kuun can optionally surface in their PF have been shown to be ambiguous between generic and nongeneric readings when ya-kuun is phonologically null. Pragmatic considerations along with contextual factors play a pivotal role in favoring one reading over the other. However, the overt presence of copula ya-kuun in these sentences and in sentences where it is obligatorily overt militates against an episodic reading and admits a generic interpretation only. More precisely, copula ya-kuun has been shown to function as a formal habitual marker that gives rise to a quantificationally habitual reading only when overt, and to two alternative readings, habitual and episodic, when optionally null. However, since this morphosyntactic marker is used in the language to encode other functions like its use in modal sentences; this marker is taken to be necessary but not sufficient to mark habituality in the verbless sentence in MSA.

In the negative verbless sentence, the two available negative particles, laa and laysa, have been shown to be in complementary distribution. While the negative particle laa can only be used in a sentence with an overt copula ya-kuun, negative particle laysa can occur in all other environments. This indicates that laysa is the unmarked negative particle in the verbless sentences in MSA, which is compatible with a variety of interpretations, episodic, characterizing, and habitual. Negative particle laa, however, is only compatible with sentences that admit a habitual reading only. This shows that laa encodes habituality in the negative verbless sentence in MSA. However, laa cannot be taken as a necessary and sufficient habituality marker in the language, simply because it is used in other syntactic structures to encode other functions like its use in prohibitive sentences.

A related issue that has been tentatively addressed in this chapter is the exact semantic meaning of the copula from a compositional viewpoint. As has been emphasized, copula ya-kuun is compatible with verbless sentences modified by adverbial adjuncts only. An insightful point pertaining to the semantic contribution of copula ya-kuun to such a structure arises. From a
compositional viewpoint, when standardly co-occurent pieces of morphology are found in one structure, the division of labor between the two becomes hard to decide. In MSA, the co-occurrence of copula ya-kuun and spatio/temporal adverbials in habitual verbless sentences makes it hard to tell whether habituality meaning is contributed by the adverbials or the copula. More precisely, which expression provides a set even-like structures over which a generic or habitual quantifier quantifies? I suggest that the copula, not the adverbials, contributes such a set of eventualities. This suggestion needs more investigation when building formal models of genericity manifestations in the verbless sentence in MSA. However, this suggestion is bolstered by two observations. First, some verbless sentences are modified by adverbial adjuncts (5.a&12.a,b); nonetheless, these sentences do not tolerate copula ya-kuun. These sentences can express episodic reading only because they lack situation variables necessary for a habitual operator to quantify over. The adverbial adjuncts in these sentences do not denote variable situations, but rather refer to particular situations. If eventualities were provided by adverbial adjuncts, these sentences would express habituality even if ya-kuun is obligatorily absent, but they do not. Second, some verbless sentences are modified by adverbial adjuncts; nonetheless they require overt copula to express habituality (30). This indicates that without the copula, which supposedly contributes event-like structures, habituality cannot be expressed.

Suggesting that copula yakkun contributes these eventualities gives another insight regarding the hot debate among modern Arab linguists whether ya-kuun exists in the underlying structure of the verbless sentence, the one modified with adverbial adjuncts in particular, or not. This suggestion indicates treating ya-kuun as a covert copula in sentences in which its presence is optional. However, the exact semantic contribution of the copula to habituality remains to be investigated in a more formal approach which will not be pursued in this project.
CHAPTER 6: CONCLUDING REMARKS

This dissertation has examined the manifestations of genericity and definiteness in Modern Standard Arabic, one of the least studied languages with respect to genericity phenomena. This and similar studies are intended to open a window for building reliable formal models that capture both the distributional facts and semantic structures of generics not only in MSA, but also in other studied and yet to be studied languages. By accurately capturing the semantic structure of generics, such models make possible their cross-linguistic study and provide arguments for and against the universality of various meaning-making mechanisms. Many semanticists, for example, argue for a universal semantic structure of characterizing sentences (e.g. Cohen 1996; Greenberg 2003, 2012; Krifka et al. 1995; Pelletier 2010b). They argue for a tripartite quantificational structure with null generic operator. The viability of this structure as a common, if not universal, semantic apparatus remains an open question if its efficiency is only tested against limited facts coming from a small set of languages like English, Italian, Dutch, and German. This study and similar studies of genericity in lesser-studied languages pertaining to this pervasive phenomenon constitute a crucial step towards testing such claims.

The interaction between genericity and definiteness is an important start point for investigating genericity in natural languages. This interaction is explicated by the observation made by many semanticists (e.g. Krifka 1987; Krifka et a. 1995; Link 1995; Pelletier 2010b) that the distinction between the two sub-types of generics, reference to kind generics and characterizing sentences is established on the semantic definiteness of the NP subject upon which a generalization is made. Based on data from a number of languages, the former appear to entertain definite NP subjects; whereas, the later require indefinite NP subjects. Considering this distinction between generics and interaction between genericity and definiteness, this dissertation has argued for
departure point from Arab traditional grammarians’ view of identifying definiteness in MSA. It argues that determining definites in MSA should not be done based on morphology alone as the traditional view claims. Definiteness, as a semantic phenomenon, needs to be investigated semantically based on the semantic concepts of familiarity, identifiability, and uniqueness which first triggered its morphological grammaticizations. In MSA, although the definite article *al* sometimes marks the noun it is prefixed to definite, it does not do so invariably. More often than not, *al* is prefixed to a noun whose denotation is not a familiar, identifiable, or unique individual, but the opposite. The definite article *al* has been shown to be inconsistent in its behavior in terms of encoding semantic definiteness.

A direct repercussion of adopting a semantic cluster for determining definites in MSA is rejecting the idea that characterizing sentences are either nonexistent in the language, or that their semantic structure is peculiar in that it incorporates a definite NP subject, tied to a particular individual, to be bound by GEN operator. This dissertation has argued that characterizing sentences in MSA entertain the same semantic structure of those in studied languages. Although characterizing sentences tolerate *al*-Ns only in subject position for grammatical reasons not related to genericity, these *al*-Ns are semantically indefinite. Semantically indefinite NPs are required in characterizing sentences as they contribute variable individuals over which the GEN operator quantifies. Accurately capturing the distributional facts about characterizing sentences in MSA contributes to the crosslinguistic endeavor initiated in Carlson (1989, 1995) and Dahl (1995). This project aims at investigating whether and how generics are formally marked in natural languages. In other words, it aims to investigate how genericity is encoded in natural languages.

In this dissertation, a semantic analysis for the syntactic asymmetry of copula *ya-kuun* in present tense verbless sentences has been proposed. This analysis has explained the asymmetric distribution of copula *ya-kuun* based on a generic/nongeneric distinction. Sentences where the copula is obligatorily unexpressed have been shown to entertain generic or episodic meanings based on the predicate and the *al-N* subject used. Such sentences lack adverbial adjuncts argued to license copula

202
ya-kuun in this structure. It has been argued that sentences in which the copula is obligatorily expressed avail a generic reading only. Those with optionally un/expressed copula, however, have been shown to differ in the type of reading they avail on par with the phonological status of the copula. Sentences with expressed copula avail a generic reading only, while those with unexpressed ya-kuun are ambiguous between generic/nongeneric readings. The context and the predicate used help teasing the two readings apart. Although this dissertation has not investigated whether or not verbless sentences with optionally unexpressed copula have the copula in their underlying structure. It has alluded to how interesting this issue is. The analysis proposed in the dissertation has suggested a meaning difference between the two sentence types, those with ya-kuun and those without. This, in turn, might be developed further. The sentence types differ by more than just phonology. If they only differed by phonology, we would expect the same range of meaning. Since the meanings are different, it might indicate that there is no hidden ya-kuun, and that ya-kuun is making a semantic contribution of some sort. This insight needs a thorough investigation in the future to be done by the writer of this project, or it might inspire other linguists interested in the verbless sentence in MSA.

The overall goal of this dissertation has been to investigate genericity phenomenon in MSA by grammatically describing and semantically analyzing its manifestations in the language. The overarching goal for the future is to build upon what has been done in this project, and exploit the distributional facts outlined herein to build formal models that capture the semantics of generics in the language. The semantic contribution of copula ya-kuun, for instance, needs to be investigated further to show whether it actually functions as a habitual marker or not. If yes, the copula might contribute eventualities for the GEN operator to quantify over. However, this analysis must reconcile the fact that the copula is licensed by adverbial adjuncts. The contribution of such adjuncts and their interaction with the copula in verbless sentences need to be investigated and articulated clearly.

Finally, this dissertation is intended to contribute to a better understanding of generics in MSA and in other natural languages by using the results and reasoning of modern semantic theories and approaches to investigating the phenomenon.
BIBLIOGRAPHY


