THE SYNTAX OF SPOKEN ARABIC

A COMPARATIVE STUDY OF MOROCCAN, EGYPTIAN, SYRIAN, AND KUWAITI DIALECTS

KRIStEN E. BRUSTAD

GEOGETOWN UNIVERSITY PRESS
WASHINGTON, D.C.

(2000)
1 THE DEFINITENESS CONTINUUM

1.0 Introduction

It is increasingly recognized that, in natural language, many grammatical features such as definiteness, number and animacy interact with each other, and behave more like continua than like strictly delimited categories (Comrie 1981, Givón 1979a, Croft 1990). In Arabic, nouns are said to be either definite or indefinite, but this dichotomy imperfectly represents the real world, in which entities can be more or less definite and specific. Natural language data from spoken Arabic in all four dialect areas contain entities that are neither wholly definite nor wholly indefinite, but rather lie somewhere in between, in an area that may be called ‘indefinite-specific.’ In this chapter, I will argue that speakers of Arabic exercise a degree of control in manipulating the rules of syntax to try to approximate these shades of meaning. This control allows speakers the flexibility they need to more closely describe the continuum of definiteness that the natural world presents.

1.1 Definite and Indefinite Markers

Definiteness in Arabic seems, at first glance, to be quite straightforward. Prescriptive and descriptive grammars alike describe the system of definiteness and indefiniteness as dichotomous: nouns are either definite or indefinite, and proper nouns are definite whether or not they are marked by the definite article. Nouns can be made definite with the addition of the definite article /i/-, or by the specification of a noun by the addition of another substantive to it to form a possessive construct (/i-dāfa/). In all of the dialects examined here, the definite article is /i/- or a phonetically determined variant. The following examples demonstrate:

<table>
<thead>
<tr>
<th>Indefinite Nouns</th>
<th>Definite Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>اين ضيف</td>
<td>في الباب</td>
</tr>
<tr>
<td>córadum dīf</td>
<td>fi l-bāt</td>
</tr>
<tr>
<td>came-he-to-them guest</td>
<td>in the house</td>
</tr>
<tr>
<td>A guest visited them</td>
<td>in the house</td>
</tr>
<tr>
<td>فستان يجنبن</td>
<td>fi l-bēt</td>
</tr>
<tr>
<td>dress it-make-crazy</td>
<td>in the house</td>
</tr>
<tr>
<td>a stunning dress</td>
<td>in the house</td>
</tr>
<tr>
<td>تاخدی شرفة؟</td>
<td>fāt ‘a l-maṭam</td>
</tr>
<tr>
<td>you-take piece</td>
<td>went-he into the-restaurant</td>
</tr>
<tr>
<td>Will you take a piece?</td>
<td>He went into the restaurant.</td>
</tr>
<tr>
<td>خوش مكان</td>
<td>‘a l-bahr</td>
</tr>
<tr>
<td>good place</td>
<td>on the beach</td>
</tr>
<tr>
<td>a good place</td>
<td>on the beach</td>
</tr>
</tbody>
</table>

In addition to the definite article, several of these dialects make use of other articles. Moroccan and Syrian employ the particle /ši/ some, and all four dialects permit limited use of the article /wāhid/ (Moroccan /wāḥid l-/) one (f /wāḥda/).

In Moroccan, Harrell (1962) notes two “indefinite” articles: a “concretizing” article /wāḥid l-/ one, and a “potential” article /ši/ some (1962:147, 189). Examples of these articles from my data include:

<table>
<thead>
<tr>
<th>M1</th>
<th>كايح واحد الحاجة</th>
</tr>
</thead>
<tbody>
<tr>
<td>kān wāḥad l-hāja</td>
<td>there-is one thing</td>
</tr>
<tr>
<td>there is something</td>
<td></td>
</tr>
<tr>
<td>M10</td>
<td>كييفي كيقول شيء كلمة غبية</td>
</tr>
<tr>
<td>kaybqa kayqul ši kolma qbiḥa</td>
<td></td>
</tr>
<tr>
<td>indic-he-keeps indic-he-says some word ugly</td>
<td></td>
</tr>
<tr>
<td>he keeps saying a some bad word</td>
<td></td>
</tr>
</tbody>
</table>

Syrian Arabic shares with Moroccan the article /ši/, although Cowell classifies Syrian /ši/ as a partitive noun rather than an article (1964:467). Syrian speakers also use /wāhid/ (f /wūdē/ (Grotzfeld 1965:76), primarily with human nouns. Examples of both /ši/ and /wāhid/ from my data include:
1.2 Definiteness, Indefiniteness, and Specification

Chafe gives a good description of 'definite' status: "I think you know and can identify the thing I have in mind" (1976:39). In other words, in order for a nominal phrase to assume definite status in discourse, it must meet one of several conditions: (a) it must have been previously mentioned in the discourse; or (b) it must be a member of a universal set of entities, such as the sun, that can be assumed to be known and identifiable without further specification, or (c) the speaker must have good reason to think that the entity is retrievable by the listener through knowledge shared by the interlocutors (Chafe 1976). Pragmatically, a definite noun usually represents given information, or information that has already been established in the discourse or can be assumed by the speaker to be present or active in the mind of his or her interlocutor. An indefinite noun, then, does not meet any of the above conditions, and represents an undefined, not retrievable entity.

However, not all indefinite nouns are created equal. Medieval Arab grammarians called the grammatical specification of nouns /taxšiš/ specifying. Wright notes that /taxšiš/ includes modification of an indefinite noun by adjectival phrases and annexation (1898ii:198). A Lebanese encyclopedia on Arabic grammar defines it thus:

الtaşšiš هو تلبية الاشتكاك العامل في الكلمات، يكون بالوصف أو الإضافة. نحو: أي رجل علم، فإضافة رجل، إلى علم، خففت من تكرره. لأنه إذا فلنا، أي علم، كان مشاعم، أما إذا فلنا، أي رجل علم فإضافة تكون قد أزننا عن بعض الشيم.

Specification is the lessening of the commonality that occurs in indefinite entities, and [this] may occur through modification or the genitive (fīdāfī), such as: He is a man of learning, as the addition of 'man' to 'learning' has lessened the indefiniteness of [the former], because if we say, He is a man, [the statement] is general, but if we say, He is a man of learning, we will have eliminated from it some of that generality (al-Tūnji and al-'Asmar 1993:156).
The principle of partial specification of nouns is thus not new to the description of Arabic. Nor is it new to language typology. Croft calls it referentiality, and formalizes this definiteness hierarchy (1990:116):

Hierarchy of definiteness
  definite
  referential indefinite
  nonreferential indefinite

Croft's term referential indefinite describes al-Tūnji and al-Asmar's phrase a man of learning, and his nonreferential indefinite describes their nonspecific noun a man.

Khan's work on Semitic provides a good model to use in the analysis of nominal syntax in spoken Arabic. Khan (1984) has adapted the work of Chafe (1976), Timberlake (1977), Hopper and Thompson (1980), and others to show that one or more of the features listed in Table 1-1 may operate to attract object marking and agreement pronouns in Semitic, such that the more individuated the noun, the greater the tendency of these markers to occur. He groups these qualities together under a rubric he calls individuation or salience, which he illustrates as follows (1984:470):

Table 1-1: Khan's Hierarchies of Individuation

<table>
<thead>
<tr>
<th>Individuated/Salient</th>
<th>Non-individuated/Non-salient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Define</td>
<td>Indefinite</td>
</tr>
<tr>
<td>2. Non-reflexive</td>
<td>Reflexive component</td>
</tr>
<tr>
<td>3. Specific</td>
<td>Generic</td>
</tr>
<tr>
<td>4. Concrete</td>
<td>Abstract</td>
</tr>
<tr>
<td>5. Qualified</td>
<td>Unqualified</td>
</tr>
<tr>
<td>6. Proper</td>
<td>Common</td>
</tr>
<tr>
<td>7. 1st person &gt; 2nd &gt; 3rd &gt; Human</td>
<td>Inanimate</td>
</tr>
<tr>
<td>8. Textually prominent</td>
<td>Incidental</td>
</tr>
</tbody>
</table>

The notion of individuation provides great explanatory power for the syntactic behavior of nouns in spoken Arabic. However, I will modify Khan's framework slightly to include those features that most influence the syntactic behavior of nouns in spoken Arabic. Reflexivity, while relevant to the individuation of nouns, is often expressed in Arabic through verb morphology. And since abstract nouns in Arabic normally take the definite article, concreteness appears to have less central a role than specification and qualification in the syntactic marking of nouns. Parallel to Khan's textual prominence is physical prominence: nouns tend to be marked as more salient when they are present in the immediate environment (see further 4.2). Three other modifications are inspired by and adapted from the work of Cowell and Janda.

First, I will add to the list the feature quantification. Cowell contrasts agreement patterns of nouns denoting "collectivity or generality" with those denoting "heterogeneity or particularity" (1964:423). Belnap's study of number agreement in Cairene Arabic lends further support to this analysis (1991:68-72). Quantification involving numbers ten and lower also appears to have some relevance to the marking of new discourse topics (see 1.5).

Cowell also contrasts concepts of identification and classification, which play a role in the expression of possessive constructions. Identificatory annexion refers to the assigning of an entity to a specific possessor, while classificatory annexion assigns an entity to a set or group (1964:458). This notion of a proper identification contrasted to a generic one parallels Khan's hierarchy of specificity vs. genericness, and Khan's terms specific and generic are defined here to include Cowell's insight.

Finally, I will add to Khan's animacy hierarchy the concept of agency, defined here as the degree to which an individual or entity has the ability to act independently. Agency may be viewed as a kind of sociolinguistic parallel to textual prominence. The concept finds precedence in the work of linguists who have argued for the role of social status and power in language variation and change, among them Janda, who argues for the centrality of a feature she calls "virility" to salience in Slavic languages, which helps explain both synchronic patterns and diachronic changes in Slavic case and number agreement.1 While virility represents an appealing concept in the age of feminist theory, the concept of agency applies more generally. Factors that contribute to agency include animacy and social status; thus for humans,

---

1Laura A. Janda, UNC Chapel Hill, public lecture at Emory University, 10/16/1997, "Virility in Slavic: A Conspiracy of Factors Over Time and Space," see also Janda 1999.
gender and age play a role as well. Children have less agency than adults, animals have less agency than humans, and inanimate objects are unlikely to be perceived as having agency at all. I will argue here and in Chapter 2 that agency provides a possible explanation for certain patterns of definite marking and agreement in some dialects.

The major objection to hierarchies of individuation or salience lies in their vagueness. The features that comprise such hierarchies need detailed investigation and definition involving a large corpus of naturally occurring data. Pending such a study, attention will focus here on those features which seem to have the greatest relevance to or explanatory power for the syntax of spoken Arabic. The list in Table 1-2, revised from Khan’s (Table 1-1), includes syntactic, semantic, pragmatic, and sociolinguistic features, covering a range of levels that interact in the production of language. These features appear to play a role in pushing a noun toward higher or lower individuation, affecting the speaker’s choice of syntactic marking of nouns in spoken Arabic when a choice of marking presents itself.

Table 1-2: Features Affecting Individuation

1. Agency: includes humanness/animacy, social status or power, perhaps gender and age groups as well.
2. Definiteness: syntactic marking or semantic status (e.g., proper nouns).
3. Specificity vs. genericity: the extent to which a speaker has a specific entity in mind.
4. Textual or physical prominence: the extent to which a noun plays a role important to the discourse, or is physically present and prominent.
5. Qualification: modification of a noun with adjectives and other descriptives.
6. Quantification vs. collectivity: the extent to which a noun is specifically quantified, especially with numbers from 2-10.

Taken as a group, it is clear that these features affect the syntactic behavior of nouns, but in a way that is not (yet) possible to describe formally. It is more convenient to formalize them using a continuum, which has the additional advantage of privileging speaker control in determining the marking of a given noun. In Figure 1, the features on the right tend to reflect (or, from the speaker’s point of view, attract) higher individuation or salience, the ones on the left, less:

1.3 Definiteness and Individuation

![Image of a table with features and their effects on individuation]

The importance of this continuum lies less in predicting the marking on the noun itself than in explaining the syntactic behavior of the noun in the larger sentential context, as I will show. In other words, the higher the individuation of a noun in the mind of the speaker, the more that noun will attract certain kinds of syntactic marking.

1.3 Definiteness and Individuation

The correspondence between individuation and definiteness in Arabic is not direct, because syntactic rules also affect definite marking. For example, abstract and generic nouns in Arabic are marked with the definite article. Thus an unspecified noun referring to any member of the class may be marked with /i-l/, while a specified human noun may not be marked definite. In the following introduction to a joke, the unnamed but somewhat specified bedouin is marked indefinite with the indefinite article /wâhid/, while the generic restaurant and waiter are both marked definite.

\[\text{ف} \text{ه} \text{و} \text{ا} \text{ح} \text{د} \text{ب} \text{د} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{i} \text{د} \text{ع} \text{ن} \text{س} \text{م} \text{ش} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{wâhid/}, while the generic restaurant and waiter are both marked definite.\(^2\)]

\[\text{S2} \text{ف} \text{ه} \text{و} \text{ا} \text{ح} \text{د} \text{ب} \text{د} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{i} \text{د} \text{ع} \text{ن} \text{س} \text{م} \text{ش} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{و} \text{i} \text{ذ} \text{ا} \text{م} \text{ك} \text{ل} \text{ل} \text{لا} \text{م} \text{س} \text{ي} \text{ذ} \text{ا} \text{م} \text{ش} \text{أ} \text{ر} \text{ب} \text{wâhid/}, while the generic restaurant and waiter are both marked definite.\(^2\)]

\[\text{There's this bedouin who went into a restaurant. He said to the waiter, give me ice cream.}\]

Grammar views definiteness as a dichotomy: nouns are either definite or not. But the real world that a speaker knows and desires to represent is far from black and white: some nouns may be somewhat defined or specified but not entirely, either from the point of view of

\(^2\)The joke is cited in full, and its nominal markings further discussed, in 1.5.
The Definiteness Continuum

the speaker or in the assumptions the speaker makes about the listener's knowledge. Definiteness is a grammatical category in which the speaker has a degree of control, and speakers of Arabic need to represent a range of undefined, partially defined, and fully defined entities. Definite and indefinite represent the black and white ends of a definiteness continuum, while the grey area of partially defined or specified entities falls in between. I will call this range indefinite-specific, after Wald (1983). It includes nouns that are syntactically indefinite, but carry a degree of specific reference that may be reflected in a number of possible syntactic constructions. In other words, an indefinite noun that carries a degree of individuation or specificity may attract a degree of definite or specifying syntactic marking. Or, as described from a different perspective, a speaker may be influenced by these features to mark a noun in a certain way. The rest of the chapter will explore the use of definite and indefinite markers in the dialects, and show how spoken Arabic uses various combinations of syntactic markings to indicate a range of indefinite-specific status.

1.4 Indefinite-Specific Marking

The semantic range from highly individuated (e.g. definite, specific, or animate) nouns to non-individuated (e.g. indefinite, non-specific, or inanimate) nouns includes varying degrees of definiteness and specificity. The more specific the reference of the noun, the greater the probability that the noun will be syntactically marked with some kind of article or specifying construction. This section will examine two kinds of indefinite-specific marking, the article /ši/ in Moroccan and Syrian, and the indefinite suffix /-in/, found in some Gulf regions.

1.4.1 Indefinite-Specific Article /ši/

Both Syrian and Moroccan speakers often identify a noun in the indefinite-specific range with the referential indefinite article /ši/ some (kind of). In the following sentence, /ši/ lends a degree of specificity to the noun /muqaddime/ introduction.

\[\text{S2: } \text{lazim na'mil-lu ši muqaddime la-hatta mā yinšidim necessary we-make-for-him some introduction so-that neg he-be-shocked We must arrange some kind of preparation for him so that he won't be shocked}
\]

Analogous Moroccan examples include the following /ši kalma qbiha/ some nasty word, /ši mas'ā'īl qbiha/ some nasty things, and /ši nās diğfān/ some guests:

\[\text{M10: } \text{kaybqa yqūl ši kalma qbiha, kayqūl ši mas'ā'īl qbiha indica-he-keeps he-say some word ugly, indica-he-says some things ugly He keeps saying some nasty word, he says some nasty things}
\]

\[\text{M2: } \text{wa 'arna 'ndi ši nās diğfān and I at-me some people guests While I had some guests}
\]

Here, as in the Syrian example, the article /ši/ indicates the partial specificity of the nouns it modifies. It is worth noting that all of the nouns in the Moroccan examples are modified, reflecting the relevance of qualification to the individuation of a noun. The Moroccan data thus provide evidence of correspondence among qualification, specificity, and syntactic marking on nouns. All these examples demonstrate that speakers use /ši/ to indicate that they have a particular type of entity in mind.

1.4.2 Nuation as Indefinite-Specific Marking

In formal Arabic, nuation or /tanwīn/ refers to the endings /-un/, /-an/, and /-in/ that function as indefinite case markers on nouns and adjectives. Nuation thus represents part of the case-marking system of formal Arabic. However, spoken Arabic has no case-marking system. In a number of dialects, nuation, in the form of /-an/ (/tanwIn fatha/) occurs in certain fixed adverbial expressions, such as /'abada/ ever, at
all, and /dāyman/ always. Adverbial /-an/ is of higher productivity in the educated registers of spoken Arabic.

Another type of nunciation, usually realized as /-in/ and occurring on indefinite nouns in a number of bedouin dialects, especially in their poetic register, are considered to be vestiges of the case-marking system of formal Arabic. Yet this suffix provides no case information. Holes mentions "the vestige of" a suffix /-in/, found in "the speech of some less educated Gulf speakers, and in dialect poetry," which marks indefinite nouns only when occurring in "Noun + Attributive Adjective" phrases (1990:115). In his study of Najdi Arabic, Ingham includes the indefinite marker /-in/ as part of noun phrase structure (1994:47).

<table>
<thead>
<tr>
<th>indefinite</th>
<th>bêt</th>
<th>'house' or 'a house'</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite</td>
<td>al-bêt</td>
<td>'the house'</td>
</tr>
<tr>
<td>possessed</td>
<td>bêt-i/-ik</td>
<td>'my/your house etc.'</td>
</tr>
<tr>
<td>indefinite (marked)</td>
<td>bêt-in</td>
<td>'a (particular) house'</td>
</tr>
</tbody>
</table>

This schema suggests that /-in/ functions as an indefinite-specific marker, and Ingham's description of the usage patterns of this ending further confirms this analysis (1994:49; IND = indefinite):

(i) where a nominal follows another nominal as in bêt-in kibir (house-IND large) 'a large house;' (ii) where a modifying prepositional phrase follows a nominal as in wathd-in min ar-rabû (one-IND from the group) 'one of the group' (i.e. one of my friends) ... (iii) where a noun is followed by a modifying clause as in kalnit-in gâl-ð-hâ-li (word-IND said-they-it-to-me) 'a word which they said to me.'

Nunciation in Najdi Arabic occurs on indefinite nouns modified by an adjective or relative clause, in other words, on specified indefinite nouns. The fact that this ending occurs on indefinite nouns that are modified in some way is significant, for it indicates that such a noun is not purely indefinite, but has a degree of specificity. Urban Kuwaiti seems to have lost the ending except in highly formalized contexts such as poetry; most published material on this phenomenon does not distinguish between poetic use of /-in/, which belongs to a special artistic register, and examples taken from naturally-occurring speech. Al-Ma'tūq's study of the tribal dialect of the 'Ajmān in Kuwait reports


1.4 Indefinite-Specific Marking

This function of nunciation appears to be quite old. Evidence for the use of nunciation with /-an/ as an indefinite-specific marker may be found in early Spanish Arabic texts. Corriente reports that this kind of nunciation has "the function of linking constituents," these constituents being an indefinite noun with a following adjectival or relative clause (1977:121-2). Corriente's examples clearly show contexts in which an indefinite noun is qualified or partly specified, resulting in an indefinite-specific noun phrase marked with /-an/. These phrases include (121-2):

- kalban abyad 'ala qalban kâfir
  - a white dog with an unfaithful heart
- bi-xâ'rîn yattaqa al-mîl al-nâr
  - with a mind as bright as fire

Shumaker's (1981) study of the indefinite suffix /-an/ in Galland's fourteenth or fifteenth century manuscript of Alf Layla wa Layla (The Thousand and One Nights, ed. Mahdi 1984) establishes patterns of /-an/ as a syntactic marker of textually prominent entities that appear to correspond to indefinite-specific patterns. Her conclusions suggest that /-an/ functions in the text as a type of indefinite-specific article. Examples from the text support this analysis; in the following passages, indefinite nouns that are specified with adjectival or relative phrases end in /-an/.

From the "Story of the Porter and the Three Girls" (Arabic text from Mahdi 1984i:126; transcription and translation mine):

... وقفت عليه امرأة مائعة في أيزاب موصل مشعر بحرير ...

[i'd waqafat 'alayhi imra'a-at-an multaffa fi 'izâr musîli muša'ar bi-harîr ...]

when she stood over-him woman-an wrapped in shawl Mosuli fringed with-silk

When suddenly there stood before him a woman wrapped in a Mosuli shawl fringed with silk

The suffix /-an/ on /imra'a-at-an/ woman is marked in the Arabic text, and the noun is partly specified by the following phrase. Another part of the story contains an instance of nunciation marking an indefinite noun modified by a relative clause (Mahdi 1984i:138):
1.5 Definiteness and First Mention: New Topic

On the continuum of definiteness, near the indefinite-specific range, there appears to be an area reserved for the first mention in discourse of 'textually prominent' entities (one of Khan’s features, see Table 1-1). The first mention of a nominal entity in discourse may be indefinite, indefinite-specific, or, in some cases (and in Moroccan particular), definite in marking. Not all first-mention nouns are of equal importance to the discourse; some nouns play a more prominent role than others. Discourse analysts distinguish between figure and ground, or central and marginal entities (and events) in discourse respectively (Hopper and Thompson 1980). An entity that plays an important role in a text is likely to be marked in a way that reflects that status, whereas an entity that is part of the background is less likely to be so marked.

In the following joke, the specified and textually prominent figure *bedu*in is marked with an 'indefinite' article /wābih/, while the inanimate, unspecified, even generic nouns /l-ma‘ām/ *restaurant* and /l-garsōn/ *waiter* are marked with the definite article, and inanimate /būza/ *ice cream* has no article in its first occurrence:
1.5 Definiteness and First Mention: New Topic

also syntactically restricted to human nouns:

E2  ... كان فيها واحدة ست... 
was-it there-is one woman

There was a/this woman...

In this Kuwaiti example (repeated from 1.1.1 above), /wāḥid/ also modifies a human noun, /mṭāwwa/ 'learned religious man.'

K3 راحت حك واحد مطوع
rahāt ḥagg wāḥid mṭāwwa‘
went-she to one religious-man

She went to a learned religious man

Evidence thus suggests that “new topic” status as marked syntactically by the article /wāḥid/ is restricted in Egyptian, Syrian, and Kuwaiti dialects to textually prominent, highly individuated, and specifically human, nouns.

In contrast to the restricted use of /wāḥid/ in the other dialects, the Moroccan article /wāḥad l-/ is not restricted to human or even animate nouns, but extends to inanimates as well:

M10 كاين واحد النوع آخر د الحوت
kān wāḥad nūr ʾāxur d l-hūt
there-is-ms one the-kind other gen the-fish

There's this other kind of fish...

In Moroccan, /wāḥad/ is obligatorily followed by the definite article /l-/ /wāḥad l-/, and it is used with a much broader range of entities, not only human but also inanimate entities if they are textually prominent. The following Moroccan passage introduces two characters in similar fashion, and a third differently (all marked in boldface):

M11 حاجيت لك هادا واحد الراحيل ما عنده الولاد، عنده غير
hājit-lk, hāda wāḥad ʾr-rāžl mà ʾandūs l-wlād, ʾandū gīr l-mrā w marūt ʾammrha mà wldat. nāḍat wldat wāḥad l-bant
told-story-I-to-you, this one man neg-at-him the-children, at-him only the-wife and wife-his life-her neg bore-children-she. arose-she bore-she one girl

---

5I have adapted the term 'new topic' from Wald (1983).
I tell you a story, this is a man who doesn't have children. He has only the wife, and his wife never bore children. Then she had a daughter.

This story opens with the introduction of a main character identified by the article /wāḥad 1-/ a. The next characters introduced are /l-wlād/ children, marked definite because of their high animacy, except that their existence is negated, and so they do not constitute a topic. The wife /l-mra/ is not marked as a new topic, however, presumably because her existence may be inferred, as most men are married, and she thus constitutes a retrievable entity that does not need to be singled out for introduction. The next topic introduced is /l-bant/ girl, marked with /wāḥad 1-/, as a new topic and important to the story. The function of /wāḥad 1-/ as a new topic marker is substantiated by the fact that the man and the girl are the only two characters who reappear in the story.

Two entities are introduced in the next passage. The first, /wāl d l-malik/[aj king's son is human, and also has an important role to play in the story, and so the speaker takes care to establish textual prominence for him by marking the phrase with /wāḥad/. The second identity, /šaržam/ window is neither animate nor topical, and thus is not marked by the new-topic article.

M2
nāḍ a-lalla wāḥad wald l-malik -- w-rah 'andha šaržam l-ḥāl hāk"a. tallat m š-šaržam
arose-he O lady one son of the-king -- and see-now at her window like like-this. leaned-out-she from the-window
So, my lady, this prince up and -- She had a window like that. She leaned out of the window.

In the third passage, the repetition of /mra x'ra/ another wife with the addition of the 'new-topic' article /wāḥad 1-/, may be due to the necessity of marking this second wife as a significant character in the story. After omitting the article in her first mention of the woman, the speaker may have subconsciously felt that the character needed more of an introduction:

M11
nāḍ tāwweẓ mra x'ra tāwweẓ wāḥad l-mra x'ra
arose-he married-he woman other married-he one the-woman other
He up and married another woman, married this other woman.

Further evidence that /wāḥid/ functions as a new topic marker lies in the use of /ši/ without /wāḥad/ in both Moroccan and Syrian. While partially individuated, textually prominent persons may be introduced with /ši wāḥ(i)d/ someone, speakers of both dialects can use /ši/ to mean someone when the reference is non-individuated, non-specific, not textually prominent, and of low social status. An example from Caubet's Moroccan texts (1993ii:7):
M
šī kikmī šī saβaši, šī kayboddal žwānāt, šī ...
some indic-he-smokes the-hashish-pipe, some indic-he-exchanges joints, some ...
Someone is smoking the hashish pipe, someone is exchanging joints, someone ...

From my Syrian data:
S2
īlla ya'ni 'alam 'aswa' min[h]a bo-kitr kitr ya'ni še 'ilon mādī kitr mšāraḥ minšan yiqbalu yāxdu ya'ni ha l-bināt
... except that-less worse and-she by-a-lot a-lot that-is thing to-them past alot sordid in-order-to they-accept they-take that-is this the-girl
... except for people who are much worse than her, that is, people who have a very sordid past, to accept to marry the girl.

The marking of an indefinite noun with the new-topic article /wāḥid (1-)/ appears to be motivated by the high degree of textual prominence played by that noun. The derivation of the article itself, /wāḥid/ one, further supports the notion that quantification, individuation, and textual prominence are all related to each other as well as to the syntactic marking of nouns.
1.6 Definite Marking in Moroccan

Moroccan speech in particular is characterized by the occurrence of unexpectedly "definite" nouns that fulfill none of the conditions specified by Chafe (see 1.2). First-mention nouns are normally indefinite in Arabic, as in most languages that mark definiteness. But in the following passage (repeated from above), even though children is semantically indefinite, since the man in question has none, and his wife is introduced for the first time, both /l-/awlad/ the children and /l-mra/the wife both have definite marking in this passage:

M11 حجاجيت لك، هاذا واحد الراجل معندوش الوالد، عندنْ غير الموالد
hāžīt-lk, hāda wāḥad ar-rāžal ma'andāš l-awlad, 'andu gī[ru] l-mra
told-I-to you this one the-man neg at-him the-children, at-him
only the-wife
I'll tell you a story, this is a man who has no children. He has only a wife

The next passage contains the first mention of /bint/ room in the story; therefore, it would be expected to be indefinite. On the other hand, the noun /bint/ has highly individuated reference in this case, which 'attracts' the definite article here.

M11 باها ملك سلطان وناض دار لها البيت ديانها يوحدها
bbʰَا-āha malik saṭān w nāḍ dār-lha b-int dālha bʰʾuhdha
father-hers king sultan and arose-he made-he for-her the-suite
gen-hers by-herself
Her father [was] a king, a sultan, and he up and made her a
suite of her own

Similarly, /l-wald/ the son in the following refers to a nonexistent child. The importance and (future) agency of the desired son attract definite marking to a noun that should otherwise be indefinite:

M11 ما ندهاش الوالد. ناخد كان لها انا خصصني الوالد
mā'nāḥaš l-wald. nāḏ gāl-lha 'ana xasṣīni l-wald
neg-at-her the-son arose-he said-he to-her I needed-for-me the-son
She didn't have [a] son. He up and told her, I need [a] son

There is no previous mention of /tūr/ bull in the passage from which the next example is taken, nor does it appear to have specific or generic reference, and yet it is marked definite. In this case, either animacy has influenced the 'indefiniteness' of the noun and pushed it towards the definite end of the continuum, or the speaker assumes that a slaughtered bull is an expected and retrievable part of celebrating a son's birth:

M11 دبع التور، عرض عنة الناس
dbah t-tūr, 'rad la n-nāš
slaughtered-he the-bull, invited-he the-people
He slaughtered a bull, invited people

Finally, even though the reference of /l-hū/fish in the following sentence is semantically indefinite, generic, and nonspecific, it carries definite marking. What attracts definite marking here may be the textual importance of fish to the passage on the depressed fishing economy of Larache from which this sentence is taken:

M10 ماشي تصور بعينك تقول راه كاين الموت?
māšī tšōf b-ink tqāl rāh kāyn l-hūt
fut you-see with-eye-your you-say see-here there-is the-fish
You'll see with your own eyes and say there are fish

The examples just cited may be contrasted to the following passage, in which an indefinite noun does occur. Here true indefinite nouns are not marked with the so-called "indefinite" articles: /blād/ a place is without marking, due to its indefinite, unspecific, and inanimate status in the story; in other words, it is not individuated.

M11 والنهار اللي ولدت أمك هربوا عليها جعلها بالمكر وحروبا بلاد
w-n-nahr lli wald mmʃ k harbu t-'liha xallāwha b-l-kaʃ w xwāw
blād w 'ammur blād
and-day that bore-she mother-your deserted-they on-her left-they
her with-the-belly and emptied-they place and settled-they place
The day that your mother gave birth they deserted her. They left
her pregnant and moved to another town.

See Cabet (1983) for a more detailed, formalized approach to definiteness in Moroccan Arabic.
1.6 Definite Marking in Moroccan

The discrepancy in marking between the indefinite girl and definite son may be attributable to the social importance of the male child, giving him a higher degree of individuation. In any case, it is clear that these features operate as semantic continua, leaving a fair degree of control to the speaker.

Another feature that seems to attract definite marking to Moroccan nouns is that of inalienable possession. Diem (1986) investigates the concepts of alienability and inalienability in Semitic, and suggests that this broad distinction is useful in determining the grammaticality of sentences containing pseudo-verbs *and/ and /li-/, both meaning to have, in spoken Arabic. These concepts are indeed important, and in fact they are more broadly applicable than Diem suggests, especially to Moroccan. Diem cites two contrasting examples from a Fez informant, the one on the right judged to be grammatical, and the one on the left ungrammatical (1986:278):

\[
\begin{align*}
\text{M} & \quad \text{دى nào دار كبيرة} \\
& \quad \text{Er hat ein grosses Haus} \\
\text{M} & \quad \text{عذبة بآ مشهور} \\
& \quad \text{He has a famous father} \\
\text{K} & \quad \text{عذبة دار كبيرة} \\
& \quad \text{He has a famous father} \\
\text{K} & \quad \text{عذبة بآ مشهور} \\
& \quad \text{Er hat ein grosses Haus}
\end{align*}
\]

The ungrammaticality of the second example, Diem claims, is due to the inalienable nature of father. A more precise explanation is that this inalienable nature interferes not with the construction to have, but with the reference of the noun itself: /bba/ father is interpreted by my informants as my father.
1.6 Definite Marking in Moroccan

Moroccan speakers thus consistently avoid using terms for male relatives in the indefinite. Diem's concept of inalienability, applied to the individuation continuum, explains this pattern. In turn, it is worth noting that these persons have a high degree of agency, which also may operate to attract specific, definite marking. The individuation hierarchy explains why certain kinds of nouns cannot remain "unassigned" or unspecified. The higher the individuation of a given noun or noun phrase, the more likely it is to receive definite syntactic marking. In the case of Moroccan, certain nouns with high animacy indicating close familial relations seem to take definite marking even in contexts low in definiteness.

The continuum of individuation also explains another feature particular to Moroccan speech. In most varieties of Arabic, both members of a definite noun-adjective phrase must agree in definiteness, such that both will carry the definite article, as the following Egyptian and Syrian examples show:

**E1**

it-tayy\r il-maw\n
the-outfit the-mauve

**S2**

il-mu\tama\t s-suri\n
the-society the-Syrian

the mauve outfit

Syrian society

However, my Moroccan data include examples of asymmetrically definite constructions such as the following:

**M10**

... kayt\b\u f l-h\han\u t \as\ri\n
... indic-they-are-sold in-the-store modern

... they are sold in a modern store

The phrase /l-h\han\u t \as\ri/ at the modern store consists of a definite noun (/l-h\han\u t/) modified by an indefinite adjective (/\as\ri/), a construction not permitted under the syntactic rules of Arabic. If it were an isolated occurrence, it could be dismissed as a performance error; however, several such examples occur in my data. Moreover, Harrell notes the occurrence of asymmetrically definite noun-adjective phrases in Moroccan, and notes that the adjective tends to remain unmarked unless the reference is quite specific (1962:166). In fact,
specificity may not be the only factor involved. The following example contains a noun-adjective phrase in which the noun is specified, but the adjective is indefinite:

\[ \text{M1: } \begin{align*} &\text{'andha dūqa xāṣṣ} \\
&\text{at-her taste-her special} \\
&\text{She has her own taste or She has a special taste of her own} \end{align*} \]

The asymmetric definiteness of the phrase /dūqa xāṣṣ/ her own special taste may be explained by the low animacy and high abstractness of the noun /dūqa/ taste, rendering the noun relatively low in individuation, and attracting the indefinite form of the adjective. A Moroccan informant confirms that the adjective is marked definite only in highly individuated contexts, such as the following:

\[ \text{M3: } \begin{align*} &\text{'andha d-där l-mxayyara f ʃ-ʃara} \\
&\text{at-her the-house the-choice in the-street} \\
&\text{She has the best house on the street (elicited)} \end{align*} \]

These Moroccan examples show that indefinite marking can interact with definite marking when the noun in question is not highly individuated. The reverse can also happen: at times definite marking can interact with indefinite to give higher specificity or individuation to an otherwise indefinite noun. Chapter 3 will show that this latter phenomenon occurs as well in relative clause structures in all four dialects.

1.7 Summary

The data cited in this chapter show clearly the inadequacy of traditional categories of definite and indefinite, and suggest that definite and indefinite marking represent two ends of a continuum of definiteness, which includes an indefinite-specific range that may be expressed syntactically by one of several strategies.

The articles /šī/ and /wāḥid/ do not mark true indefinite nouns, but rather nouns that lie somewhere between definite and indefinite, in an indefinite-specific range. This range is represented in the syntax of all four dialects, in Moroccan and Syrian by the article /šī/, in all four dialects by the article /wāḥid (ı-)/, which functions as a "new-topic" article, and in Syrian, Egyptian and Kuwaiti by the interaction of the semantics with other definite and indefinite markers, such as the use of definite relative pronouns with morphologically indefinite nouns.\(^7\)

The geographical distribution of indefinite articles in Moroccan and Syrian, both of which share /šī/, the related Egyptian article /tšī/, and the apparent indefinite-specific function of the suffix /-in/ in the Arabian Peninsula—even though it is disappearing in urban dialects—are facts that invite further diachronic research. The fact that Moroccan and Syrian dialects share the article /šī/ and Tunisian and Iraqi share a similar article /fard/ suggests that these articles have fulfilled this function for a very long time. Likewise, evidence that nation plays a role as specifying marker in Andalusian texts, a medieval Thousand and One Nights manuscript, and the bedouin dialects of the Gulf, suggests that this function has long been part of spoken Arabic. The absence in Egyptian and urban Kuwaiti of an indefinite-specific article suggests that these dialects may have undergone a linguistic levelling process which reduced the number of syntactic markers of definiteness and specificity.

Overall, Moroccan and Syrian dialects show greater richness and variation in nominal syntactic marking, leading to speculation of a degree of cross-dialect hierarchy in which Moroccan shows highest definite and indefinite-specific marking, while Egyptian shows the least. The fact that Cairene Arabic falls on the lower end of the spectrum of dialects in the range of specifying articles is paralleled also by its relative paucity of demonstrative forms, as Chapter 4 will show.

\(^7\) Other dialects appear to have indefinite-specific articles as well: Tunisian and Iraqi appear to share a specifying article /fard/one (see Marçais 1977 for a description of Tunisian and Erwin 1963 for Iraqi).