

Copyright  
by  
Michael Lee Turner  
2013

**The Thesis Committee for Michael Lee Turner  
Certifies that this is the approved version of the following thesis:**

**Definiteness Marking in Moroccan Arabic:  
Contact, Divergence, and Semantic Change**

**APPROVED BY  
SUPERVISING COMMITTEE:**

**Supervisor:**

---

Kristen E. Brustad

---

Patience L. Epps

**Definiteness Marking in Moroccan Arabic:  
Contact, Divergence, and Semantic Change**

**by**

**Michael Lee Turner, BA**

**Thesis**

Presented to the Faculty of the Graduate School of  
The University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**Master of Arts**

**The University of Texas at Austin**

**August 2013**

## **Abstract**

### **Definiteness Marking in Moroccan Arabic: Contact, Divergence, and Semantic Change**

Michael Lee Turner, MA

The University of Texas at Austin, 2013

Supervisor: Kristen E. Brustad

The aim of the present study is to cast new light on the nature of definiteness marking in Moroccan Arabic (MA). Previous work on the dialect group has described its definiteness system as similar to that of other Arabic varieties, where indefinite entities are unmarked and a "definite article" /l-/ modifies nouns to convey a definite meaning. Such descriptions, however, do not fully account for the behavior of MA nouns in spontaneous natural speech, as found in the small self-collected corpus that informs the study: on one hand, /l-/ can and regularly does co-occur with indefinite meanings; on the other, a number of nouns can exhibit definiteness even in the absence of /l-/. In response to these challenges, the study puts forth an alternate synchronic description of the system, arguing that the historical definite article \*/l-/ has in fact lost its association with definiteness and has instead become lexicalized into an unmarked form of the noun that can appear in any number of semantic contexts. Relatedly, the study argues that the historically indefinite form \*Ø has come under heavy syntactic constraints and can best be described as derived from the new unmarked form via a process of phonologically conditioned disfixation, represented {- /l/}. At the same time, MA has also apparently retained an older particle *šī* and developed an article *wahəd*, both of which can be used to express different types of indefinite meanings. To support the plausibility of this new description, the study turns to the linguistic history of definiteness in MA, describing how

a combination of internal and external impetuses for change likely pushed the dialect toward article loss, a development upon which semantic reanalysis and syntactic restructuring of other forms then followed. If the claim that MA no longer overtly marks definiteness is indeed correct, the study could have a significant impact on work that used previous MA descriptions to make grammaticality judgments, as well as be of value to future work on processes of grammaticalization and language contact.

## Table of Contents

List of Figures .....	ix
1. Introduction.....	1
1.1. Aims of the Study .....	2
1.2. The Oral Texts .....	3
1.3. Transcription System .....	5
2. Theoretical Background.....	8
2.1. Moroccan Arabic .....	8
2.1.1. Historical Dimensions.....	8
2.1.2. MA in Contact.....	10
2.2. Definiteness, Specificity, Referentiality, and Related Concepts .....	11
2.2.1. The Givenness Hierarchy.....	12
2.2.2. Generic Entities and the Wheel Diagram.....	17
2.2.3. Individuation and First Mention .....	19
2.3. Definiteness in Arabic Varieties .....	21
2.3.1. Ideological Dimensions .....	21
2.3.2. Definiteness in Classical Arabic .....	23
2.3.3. Definiteness in Modern Arabic Dialects.....	25
3. Previous Descriptions of Definiteness Marking in MA.....	28
3.1. The Proposed System.....	28
3.1.1. Proposed Unmarked Form Ø .....	29
3.1.2. Proposed Definite Article /l-/.....	31
3.1.3. Proposed Indefinite Article <i>wahəd</i> (/l-/) .....	34
3.1.4. Proposed Indefinite Article <i>šī</i> .....	36
3.1.5. Overview of Previously Proposed System.....	38
3.2. Challenges to Previous Descriptions .....	42
3.2.1. Presence of /l-/ in Indefinite Contexts .....	43
3.2.1.1. /l-/ with ‘Specific’ Statuses.....	43

3.2.1.2. /l-/ with ‘Referential’ Statuses .....	45
3.2.1.3. /l-/ with ‘Type Identifiable’ Statuses .....	46
3.2.2. Absence of /l-/ in Definite Contexts .....	48
4. A New Model for Definiteness in MA .....	56
4.1. Synchronic Description.....	56
4.1.1. Unmarked Form Ø (etymological */l-/).....	60
4.1.2. Article <i>waḥad</i> Ø .....	66
4.1.3. Disfixed Form {- /l-/} (etymological *Ø).....	69
4.1.4. Particle <i>ši</i> {- /l-/} .....	81
4.2 Diachronic Dimensions.....	83
4.2.1. Unmarked Form Ø (etymological */l-/).....	84
4.2.2. Article <i>waḥad</i> .....	90
4.2.3. Disfixed Form {- /l-/} (etymological *Ø).....	94
4.2.4. Particle <i>ši</i> .....	96
5. Conclusions.....	99
5.1. Implications for Previous Work.....	100
5.2. Future Directions .....	104
5.3. Closing Remarks .....	109
Appendices.....	111
A. Givón’s (1976) Wheel Diagram for Relevant Languages.....	111
B. In-Text Occurrences of Various MA Nominal Forms .....	112
C. Moroccan Arabic Texts.....	114
Text 1: Carpet-Making in Taznakht.....	114
Text 2: Getting Married .....	116
Text 3: Harsha Recipe.....	117
Text 4: How to Make Tanjia.....	117
Text 5: Making Tea.....	119
Text 6: Moroccan Tea and Coffee .....	119
Text 7: Ramadan in South Morocco .....	120

Text 8: The Baccalaureate Exam .....	123
Text 9: Weddings in Asfi.....	125
References.....	129



## List of Figures

Figure 1: MA Consonantal Phonemes .....	6
Figure 2: Gundel et al.'s (1993) Givenness Hierarchy .....	12
Figure 3: Referential Status as a Function of Unique Identifiability .....	15
Figure 4: Modified Portion of Givenness Hierarchy .....	16
Figure 5: Givón's (1978) Wheel Diagram for English .....	18
Figure 6: Terminological Overlap with Brustad (2000) .....	19
Figure 7: Wheel Diagram for Classical Arabic.....	24
Figure 8: Definiteness Marking Strategies in Various Arabic Dialects.....	26
Figure 9: Proposed Semantic Statuses of MA Forms .....	38
Figure 10: Proposed MA Forms for Givenness Hierarchy .....	39
Figure 11: Common Nouns that Avoid /l-/ (Harrell, 1962) .....	49
Figure 12: Ungrammatical Results for Proposed Forms.....	50
Figure 13: Reanalyzed MA Forms & Givenness Implications.....	58
Figure 14: Wheel Diagram for Moroccan Arabic.....	59
Figure 15: Contested Unmarked Forms .....	62
Figure 16: Deriving the Disfixed Form from the Unmarked Noun .....	71
Figure 17: Heine's (2012) Grammaticalization Stages for the Definite Article.....	86
Figure 18: Heine's (2012) Grammaticalization Stages for the Indefinite Article..	91

## **1. Introduction**

This study is predicated on the idea that language is always in the process of change. Although descriptive grammars may attempt to assign a given form a particular meaning, the reality of spoken language often betrays the grammarian's well-defined categories; over the course of time, a form that once implied one semantic connotation may well come to signal another. This unstable relationship between form and meaning can at times manifest itself in apparent departures from the codified rules of the language, with proficient native speakers producing forms that run counter to linguists' expectations for their semantic contexts. In such cases of discrepancy it is not the speakers' command of their own language that should be called into question, but rather the comprehensiveness and theoretical adequacy of contemporary grammatical treatises and descriptions.

In few linguistic domains is the tenuous relationship between meaning and form so clear as it is in that of definiteness. Definiteness, which Chafe (1976) summarily describes as a question of "whether I think you already know and can identify the particular referent I have in mind," is ultimately a function of the speaker's cognitive perception of a nominal entity's status in the discourse, a concept that one can expect to apply universally to human language. Although the semantic notion of definiteness is thus quite stable cross-linguistically, the means for expressing it – or, more appropriately, the grammatical forms – vary quite widely. Languages such as English enjoy an "overt surface marking definite status" via the definite article 'the,' which contrasts with an indefinite 'a' or zero-marking (Chafe, 1976); others, like Turkish, do not overtly mark either state and rely on other means for distinguishing them (Tura, 1986). That different languages can grammatically express the same semantic notion in so many possible ways only underscores the susceptibility of definiteness marking to variation across time and space.

### 1.1. AIMS OF THE STUDY

A first glance at Moroccan Arabic (henceforth MA), which is the focus of this study, might suggest an easily discernible paradigm for marking definiteness. Like the Classical Arabic and modern Arabic dialects to which it is related, most nouns in MA can take one of two primary morphological shapes. One of these forms corresponds with the Old Arabic indefinite state (as in *wald* ‘boy’), which gave the most basic meaning of the noun; the other contrasts with it in that it is prefixed with Arabic’s historical definite article \*/l-/ (yielding *l-weld*), the precise phonological realization of which has been described as determined by the first consonant of the indefinite form. The fact that Moroccan has these two nominal forms, both of which have been well-established in other Arabic varieties as corresponding with indefinite and definite meanings, respectively, could easily lead one to conclude that they play the same roles in MA before even considering evidence otherwise. To do so, however, would be to tie meaning to form in a way that belies either’s very real propensity for diachronic change and variation, particularly in a dialect group that has been developing largely on its own terms since the arrival of Arabs in North Africa in the late seventh century.

The present study’s primary goal is to highlight precisely this capacity for diachronic change in the MA definiteness system by showing that the historical morphophonological material that once served as a marker of definiteness no longer is strictly associated with that semantic sense, and that a new paradigm for marking definiteness, specificity and referentiality has come into play in place of the old arrangement. As part of this goal, it critiques current descriptions of definiteness marking in MA by arguing that although many scholars have had made valuable observations on the semantic roles of various MA forms, none have yet provided a fully adequate account of the syntax of definiteness in the dialect. In turn, it provides a solution through which the semantic behavior of all nouns in MA – including those that have previously been treated as anomalous – can be explained through a single set morphosyntactic rules. Key to this model are the cross-linguistically relevant concepts of markedness, grammaticalization, and contact-induced convergence, all of which are used to explain

the processes behind the restructuring and reanalysis that has allowed for the current system. In this way, the study hopes to provide not only a refined synchronic description of definiteness marking in MA but also a plausible diachronic explanation for how it may have come about.

## **1.2. THE ORAL TEXTS**

The theoretical claims found in this study regarding synchronic patterns of definiteness marking in MA are primarily substantiated through instances of actual language use as it occurs in a small corpus of oral texts that I collected while doing fieldwork in Morocco. All of the nine included texts – which are fully transcribed and glossed in Appendix C – were originally captured as high-quality digital recordings in relaxed, informal settings (typically the speakers' homes). These texts were explicitly selected for their value in elucidating the grammatical structures in question within the framework of a known discourse, meaning that the reader can refer to the original texts in order to fully investigate the narrative context surrounding any given noun and its morphosyntactic expression. Unlike for many other collections of MA texts (see the appendix of Harrell, 1962 for an example), there have been no revisions or edits made to the speech and it is represented as originally produced, with pauses, false starts, and occasional mispronunciations. The transcriptions and glosses are solely my own and I take responsibility for any possible inaccuracies.

In total, the nine texts are sourced from eight different speakers – five females and three males – ranging from 16 to 44 years of age. The majority of speakers were in their twenties at the time of the research, and all but the oldest had completed at least some high school education. In this regard the profiled speakers are not necessarily a representative cross-section of Moroccan society on the whole, but do nonetheless provide a valuable portrait of MA as spoken by Morocco's youth, who are themselves leading linguistic changes that we will likely witness in greater force in the future. In terms of geographic origin, five of the participants were natives of the Taznakht area in Morocco's southern Ouarzazate province, two had grown up in the Asfi region before

moving to Taznakht, and one was a lifelong resident of Sefrou (near Fes). Ethnically, five of them identified as Arab whereas three identified as Berber; all, that said, are fluent speakers of Moroccan Arabic who have been exposed to the language since early childhood. Although the language of those speakers hailing from Taznakht – a market town in Berber-dominant area where Arabic has only recently gained ground – cannot be said to fit into a traditional Arabic dialectology model, I follow Maas & Procházka (2012) in the belief that such locales' speech patterns are particularly valuable as a reflection of the emerging national koiné, which serves as a linguistic model in lieu of locally precedent Arabic varieties.

As I am a non-native speaker of Moroccan Arabic who was involved in fieldwork of which the participants had direct knowledge, it is not at all unreasonable to ponder the impact my presence, not to mention the act of recording, might have had on the language used in the oral texts, and in turn the validity of my conclusions. I see it fit to address this possibility directly from the outset. Ever since Labov's (1972) exposition on the 'Observer's Paradox' the notion that a fieldworker's mere presence can effect changes in linguistic production has gained wide recognition in the linguistics community. While it is doubtful that one can ever fully counter these effects, there is nonetheless evidence (briefly summarized in Cukor-Avila, 2000) that they can at least be mitigated if certain conditions are met. On this count, I maintain that the research conditions were nearly as ideal as they can be for a researcher of my background: to begin, I had recently been a long-term resident and member of the community in which all but one of the texts (#4) were recorded, and had likewise known all but one speaker (in #9) for at least a year. In addition, all but one of the texts (#8) were recorded with other native MA speakers in the immediate vicinity of the speaker, meaning there was less impetus to tailor speech for the sole benefit of the non-native listener.

A final defense is that even in the case that my role did uniquely effect morphosyntactic changes in the recorded speech that informs the analysis, they would have likely been in the direction of the prescriptively preferred Modern Standard Arabic (MSA), which would detract from the argument I am otherwise making in favor of MA's

system of definiteness marking being quite distinct from it. As these differences still appear quite salient, I stand by my conviction that the selected texts are adequate (though not flawless, if there is indeed an ideal) representations of contemporary Moroccan Arabic, an assertion that can be corroborated by examining other lexical and grammatical features that likewise appear in them. Still, where possible I have tried to support my analyses with additional evidence from other authors' work, itself an indication that the theoretical claims made herein can stand independent of a specific set of supporting data.

### 1.3. TRANSCRIPTION SYSTEM

The texts and examples in this study make use of a phonemic transcription system that has been selected as a means of interfacing the work with previous treatments of Moroccan Arabic. The system is borrowed from Heath (1987, 1989, 1997, 2002) and maintains a number of conventions familiar to Arabists, such as representing pharyngealized consonants with a sub-letter dot and palato-alveolars with a caron. Other phonemic representations, such as those of the uvular, pharyngeal, and glottal consonants, overlap with those described in the International Phonetic Association's (1999) handbook. Ultimately, however, and in lieu of a standardized orthographic convention for writing Moroccan Arabic, the choice of symbol for many MA phonemes is arbitrary; the decision to represent them as they appear here is made simply to establish continuity with other work. In keeping with the same principle, all examples cited from other authors have been re-transcribed as necessary to use the same phonemic representation outlined here. The consonantal phonemes are given in Figure 1.

In addition to the consonants, the transcriptions represent three full vowels, two ultra-short or epenthetic vowels, and one secondary articulatory feature with vocalic implications. The three medial-length full vowels that represent the "stable core of the vocalic system," again following Heath 1997, are written /a/, /i/, and /u/. The two ultra-short or epenthetic vowels are given as /ə/ and /ũ/ where they occur, though these are susceptible to syncope and may not manifest in the same way for all speakers (some will thus produce the same indefinite article variously as *waḥəd* or *waḥd*). Similarly, the

### Primary Consonantal Phonemes in Transcriptions (based on Heath, 1997)

	(Bi-)labial	(Denti-) alveolar	Pharyngealized (denti-)alveolar	Palato-alveolar	Velar	Uvular	Pharyngeal	Glottal
<b>Voiceless stops</b>		t	ṭ		k	q		
<b>Voiced stops</b>	b	d	ḏ		g			
<b>Voiceless fricatives</b>	f	s	ṣ	š		x	ħ	h
<b>Voiced fricatives</b>		z	ẓ	ž*		ɣ	ʕ	ʔ
<b>Nasals</b>	m	n						
<b>Laterals</b>		l	ɭ					
<b>Rhotics</b>		r	ɾ					
<b>Semivowels</b>	w		y					

*Dotted line demarcates consonants for which the historical article \*/l-/ assimilated*

Figure 1: MA Consonantal Phonemes

secondary articulatory feature of labialization is represented as a consonantal feature /C<sup>w</sup>/ where it has been detected. There remains some debate (refer to Aguadé, 2010; Heath, 2002) about whether /ũ/ in some MA dialects is truly a full vowel as opposed to an allophone of /ə/ in the vicinity of labials; as this debate has little effect on the present study it has not been engaged. For the most part, in fact, the representations of ultra-short vowels and labialization are morphologically dispensable and their wholesale removal would do little to affect the analysis. They have been maintained only as an aid for the reader.

A few other adjustments have been made to the transcription system to accommodate additional phonological features. Since at least one foreign loan *l-bak* ‘the baccalaureate [exam]’ uses a pharyngealized /b/ to lower the following vowel and

approximate the French lexical equivalent, the consonant has been accordingly represented. Where MSA loans appear in Moroccan speech, they are represented with their given Moroccan vocalic phonology; items from MSA and other Arabic varieties given in isolation, however, display the short/long vowel distinctions /a/, /ā/, /i/, /ī/, and /u/, /ū/. True code-switches into European languages are italicized and represented in their own orthographies. Code-switches and loans from Moroccan Berber languages, which have a high degree of phonological similarity to MA (see Kossmann & Stroomer, 1997), on the other hand, are represented using the same transcriptions as for MA.

Finally, throughout the transcribed texts I have made liberal use of dashes to separate morphologically significant elements of words, among them the aspectual particles *ka-* and *ta-*, the abbreviated future marker *ya-*, the negative prefix/suffix *ma-* and *-š*, attached subject pronouns, and the etymological Arabic definite article \*/l-/. In the case of the latter, the presence of the dash does not imply – as the study goes on to discuss – that the element is necessarily a definite article in the synchronic sense. It is instead represented as morphologically analyzable first to encode the historical morphology of the word for the benefit of the reader, and second to indicate its involvement in synchronically productive processes of disfixation, an explanation of which is more fully rendered in what follows.



## **2. Theoretical Background**

As a prelude to the central concerns of the study, a few topics in particular are worthy of discussion. I present these below, beginning with a brief overview of Moroccan Arabic and its history in North Africa. Next I turn to the concept of ‘definiteness,’ looking more closely at what the term actually implies and giving a number of useful terms and models for conceptualizing the semantic connotations it embodies and evokes. Finally, I briefly discuss how these semantic concepts apply to Arabic on the whole as a means of bridging into the discussion of Moroccan Arabic specifically.

### **2.1. MOROCCAN ARABIC**

Moroccan Arabic, as it is used in this study, refers to a group of related Arabic dialects spoken in and around a geographic area roughly corresponding with the modern-day state of Morocco, making it the first or second language of nearly 30 million people (Maas & Procházka, 2012). The borders of this group are not, of course, as clearly drawn as the political boundaries; eastern Moroccan dialects, for example, share a number of traits with Algerian dialects and vice versa. Nonetheless the concept is useful because it allows us to speak of a collection of dialects that have a significant degree of internal variation but still share some important major features. Among these shared features, particularly in the case of the central-type or koinized varieties with which we are primarily concerned, is the system of definiteness marking that is the focus of this study. While questions of inter-dialectal variation thus remain important, the analysis given herein is apt to apply to a majority of modern urban and rural dialects in Morocco. Still, for the better purpose of contextualizing the study, some brief background on the larger MA dialect group follows.

#### **2.1.1. Historical Dimensions**

Like other Arabic varieties, Moroccan Arabic is a member of the Central branch of Semitic languages, along with Aramaic, Ugaritic and Canaanite languages such as

Hebrew and Phoenician. The modern-day descendent of dialects originally spoken in the Arabian Peninsula, it was brought to North Africa during a series of expansions that followed the advent of Islam in early seventh century. Heath (2002) subdivides Moroccan Arabic into three major dialect types, each of which corresponds with a period of historical Arab migration into the area and is today characterized by an internal set of isoglosses and grammatical features that tend to distinguish it from the other two types. The first of these, the 'Northern' type, reflects the earliest Arab settlement in what is now Morocco, probably beginning at the very end of the seventh century. Also often called the 'Pre-Hilalian' dialects (as in Pereira, 2011), they now refer to a set of dialects still extant in the north of the country, but appear to have been typical of Moroccan urban centers before receding to their current, more limited range. Heath's second reference is to the 'Saharan-type' dialects, which represent the linguistic legacy of the Beni Ma'qil bedouin who are said to have migrated into North Africa in the late twelfth century; such dialects are typically found in the oases of pre-Saharan southern Morocco, but are not spoken by a significant proportion of Moroccan speakers today.

It is instead Heath's third type, or the 'Central' dialects whose genesis is often associated with the arrival of the Beni Hilal bedouin starting in the eleventh century, that encompasses the majority of contemporary Arabic dialects in Morocco. Heath speculates that these dialects emerged out a union between "a small core" of incoming nomads, already-established sedentary Arab groups, and contingents of "Arabized Berbers," all of whom may have intermixed elsewhere in North Africa and then migrated into Morocco, where their dialects crystallized. Historically, the resulting Arabic varieties have been spoken primarily in rural areas along a large strip running from the Middle Atlas Mountains to the Atlantic coast. Starting with the increasing urbanization and industrialization that accompanied the European colonial period, however, large-scale migration from rural areas to new urban centers has boosted the status of these dialects to that of the principal basis for an emerging pan-Moroccan koiné. This koiné, associated with the respective economic and political influence of Casablanca and Rabat (see Hachimi, 2007 for more on the sociolinguistic implications of the koiné), has already had

a far-reaching impact as a result of internal migration; with it now instantly accessible in millions of homes via satellite TV, this impact can only have been growing in recent years. Most otherwise unspecified descriptions of ‘Moroccan Arabic’ refer to a variety that approximates this national koiné. It is likewise the primary dialect type under scrutiny in this study, which is meant to be a small contribution to its description.

### **2.1.2. MA in Contact**

Contact with other languages has undoubtedly played a key role in Moroccan Arabic’s history. Heath (2002) suggests that the earliest wave of Arab migrants likely set up garrisons in preexisting urban centers, where many then intermarried with remaining speakers of North African Late Latin; it indeed seems plausible that the role of this Late Latin substrate in Morocco, as well as elsewhere in North Africa, has been underestimated, though few besides Heath have championed it. Relatively uncontested, on the other hand, is the notion that the major language group with which MA has been in contact is Berber (Maas, 2000). Also constituting a branch of Afroasiatic, Berber languages have been archaeologically attested in North Africa from as early as 800 B.C. (Kossmann, 2012), and were likely the dominant vernacular in all but a few urban communities when Arabs migrants first arrived. With a large minority of Moroccans still speaking Berber today (estimates vary significantly, but see Haut-Commissariat au Plan du Royaume du Maroc, 2004 for a low estimate of around 25%, or Stroemer, 2008 for a high of 45%), it can be assumed that the spread of Arabic in North Africa was on the whole a process of Berber speakers shifting to Arabic, though this likely happened neither rapidly nor exclusively. Contact with Berber has been pinpointed as the source of a number of proposed lexical and grammatical innovations in Moroccan Arabic (Chtatou, 1997; El Aissati, 2006; Maas, 2000; Tilmatine, 2011). At least one of these – the use of the numeral *waḥad* ‘one’ as a presentative marker or indefinite-specific article – relates to definiteness marking, but to my knowledge no previous studies have looked closely at contact with Berber as relevant to the entire system.

In addition to Berber and possibly Late Latin, MA has also been in relatively more recent contact with European languages. Chief among these are Spanish and French, both of which were became rooted in Morocco during a period of colonial rule and remain in common use today among some social classes. Spanish has had a long history of contact with Moroccan Arabic, dating back to at least the fall of the Iberian caliphate in 1492 when thousands of Spanish-speaking Jews fled to northern Morocco; the presence of the language only grew over the next few hundred years as Spain enacted its colonial ambitions first in the north, and then in the far southwest, of the country (Sayahi, 2011). French entered the picture in the early twentieth century alongside the establishment of the French “Protectorate” throughout most of Morocco, continued to gain prominence even after independence was declared in 1952, and is still the preferred language in a number of educational and technical fields (Youssi, 1995). Spanish and French – along with other European languages, including English – have contributed hundreds of loanwords to MA (Heath, 1989), a sign of their impact on the vernacular; nonetheless their relationship with MA, which can be described as having borrowed from them in a scenario of language ‘maintenance,’ contrasts with the substrate role of Berber as a language group from which speakers were actually shifting (Thomason & Kaufman, 1988). As this would lead one to expect European languages’ impact on MA to be more lexical than grammatical, they are not likely to have played much of a role in the evolution of MA’s definiteness marking system, but are still useful for the analysis because the behavior of some foreign loans can cast a unique light on the relationship between nominal form and meaning.

## **2.2. DEFINITENESS, SPECIFICITY, REFERENTIALITY, AND RELATED CONCEPTS**

Up to this point, ‘definiteness’ has been referenced as a sort of utilitarian cover term that refers to the state of a noun that is ‘known’ or ‘identifiable’ by the participants in a discourse; ‘definiteness marking,’ in turn, implies the way in which this semantic notion can be grammatically expressed using a particular morphological or syntactic arrangement. A comprehensive treatment of the subject, however, cannot operate on a

simple binary distinction between ‘definite’ and ‘indefinite.’ It is instead more appropriate to envision definiteness and indefiniteness as distinct ranges of a semantic continuum, the whole of which can be further subdivided as a function of how and to which degree an entity is known.

### 2.2.1. The Givenness Hierarchy

A useful starting point for this discussion is Gundel, Hedberg & Zacharski’s (1993) notion of a ‘Givenness Hierarchy,’ which allows us to situate our discussion within the broader context of how nominal entities in a discourse are assigned cognitive statuses. The authors’ original hierarchy, along with the English forms they associated with each status, is given in Figure 2.

#### The Givenness Hierarchy (Gundel et al., 1993, p. 275)

<b>in focus</b>	<b>&gt;</b>	<b>activated</b>	<b>&gt;</b>	<b>familiar</b>	<b>&gt;</b>	<b>uniquely identifiable</b>	<b>&gt;</b>	<b>referential</b>	<b>&gt;</b>	<b>type identifiable</b>
<i>it</i>		<i>that</i> <i>this</i> <i>this N</i>		<i>that N</i>		<i>the N</i>		<i>indefinite this</i> <i>N</i>		<i>a N</i>

Figure 2: Gundel et al.’s (1993) Givenness Hierarchy

In Gundel et al.’s model, nominals are assigned one of six statuses as a function of how ‘given’ they are, givenness being a measure of the speaker’s assumptions about how available (or relevant) knowledge of the entity is for those involved in the discourse. Each of these statuses reflects the accumulation of an additional type of knowledge about the entity, differentiating it from other statuses, so that as one moves from right to left the object becomes increasingly given and thus contextually retrievable on the part of both discourse participants. Similarly, the hierarchy implies that any active status will subsume the meanings of whatever statuses lie to the right of it. Thus an entity that is ‘uniquely identifiable’ can also be assumed to be ‘referential’ and ‘type identifiable’ (but not necessarily ‘familiar’ or ‘activated,’ whereas one that is ‘in focus’ can be understood

to display all of the other statuses' qualities. Inversely, an entity that is assigned the status 'type identifiable' displays only that quality and none of the others.

As it is the rightmost portion of this hierarchy – that spanning the statuses from 'uniquely identifiable' to 'type identifiable' – that is of particular relevance to our discussion, we will now look more closely at how the concepts in this range can be both defined and slightly refined for our purposes. We begin with the first and most basic status, '**type identifiable**.' Entities that are assigned the status of 'type identifiable' assume only that the addressee can "access a representation of the object described by the expression" (Gundel et al.), which is by definition a requirement for representing any nominal entity. A speaker who uses a noun of this status implies that the object of discussion could be any one or more of its type – meaning it, as an individuated entity, is not relevant to the discourse – and thus he or she has no means (or reason) to further specify it. The authors give 'a dog' in the context below as an example of a noun with this status:

- (1) *I couldn't sleep last night. A dog (next door) kept me awake.* (Gundel et al., 1993, p. 276)

For nominals of the status '**referential**,' Gundel et al. stress the speaker's intent "to refer to a particular object or objects," but not any that the addressee can uniquely identify. In this sense, the status primarily implies an element of individuation, in that the speaker conceptualizes the entity as distinct from others of its type. The same notion appears in Givón's (1978, p. 294) definition of referentiality, in which he suggests that a speaker using a referential nominal must "commit him/herself to the existence of [a] specific individual member of that genus" within the "universe of discourse." This would contrast with entities that are only 'type identifiable' – or what Givón simply calls 'non-referential' – and that could, again, be any of a number of members of the type. Gundel et al. identify the semantic notion expressed by indefinite 'this' in colloquial English, below, as an example of referentiality:

(2) *I couldn't sleep last night. This dog (next door) kept me awake.*  
(Gundel et al., 1993, p. 277)

'**Uniquely identifiable**' nominals are, on the other hand, those that can be assumed to be identifiable for both the speaker and the addressee. This knowledge, according to Gundel et al., may be present either because the entity has already been represented in the immediate discourse and is thus retrievable from the addressee's memory, or alternatively because the present discourse encodes enough information to identify the individual entity in the nominal itself. It may also be available because the entity itself is a unique referent that is unlikely to be confused with any other, such as the sun (Chafe, 1976). The nominal 'the dog' in the following example can be described using this status:

(3) *I couldn't sleep last night. The dog (next door) kept me awake.* (Gundel et al., 1993, p. 277)

Gundel et al. identify the status of 'uniquely identifiable' as essentially equivalent to 'definite,' which is the term that both Chafe and Givón use for the same semantic concept. I too hereby adopt the term 'definite' rather than 'uniquely identifiable,' first because it is better known and second to avoid confusion, because indefinite (referential) entities can in fact be uniquely identifiable in some cases, but only for the speaker. This, in turn, gives us three working terms for the tiers of a semantic hierarchy: '**definite**,' '**referential**,' and '**type identifiable**' (Givón's 'non-referential').

There is a single key refinement to the Givenness Hierarchy that I feel should be made for the purposes of this study. Put shortly, this is the subdivision of the Gundel et al.'s 'referential' tier into two independent tiers: one that remains simply referential, and a new one that I will call '**specific**.' As we have seen, referentiality is primarily a measure of the speaker's intent to refer to an individuated entity that exists within the discursive universe. As it stands, the Givenness Hierarchy implies that a referential nominal must also be uniquely identifiable by the speaker, but what this model seems unable to account for is the distinctive case of an entity that is both individuated and relevant to the

discourse (i.e. referential) but not identifiable for either the addressee *or* the speaker. To re-appropriate Gundel et al.’s model sentence for my own argument, I offer the below example (4), using the English form ‘some’ to approximate this semantic status:

(4) *I couldn’t sleep last night. Some dog (next door) kept me awake.*

In this context, ‘some dog’ can clearly be differentiated from ‘a dog’ (as in example 1) in that the speaker indeed does intend to refer to an individuated, discourse-relevant object that can be distinguished from others of its type – meaning the entity is referential – but does not have the requisite level of knowledge to uniquely identify it. At the same time, even though ‘some’ implies referentiality, it is not equivalent to indefinite ‘this’ in example 2 because the latter indicates that the speaker (but still not the addressee) has explicit knowledge of the entity and can uniquely identify it. Following Ionin’s (2006) work on the use of English’s indefinite ‘this’ in the same semantic contexts, I hereby refer to the trait that distinguishes these two semantic senses as ‘specificity’ and insert it into the hierarchy between ‘definite’ and ‘referential.’<sup>1</sup> The key element of the distinction between the statuses referential, specific, and definite, all of which refer to nominals with referential meaning, thus hinges upon whether or not the entity is uniquely identifiable to both participants in the discourse, only one, or neither (Figure 3).

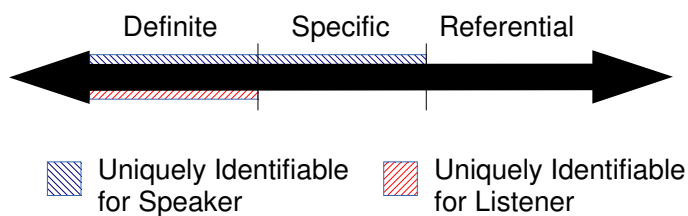


Figure 3: Referential Status as a Function of Unique Identifiability

As with other statuses in the hierarchy, a ‘specific’ nominal subsumes the statuses to the right of it so that it is also referential and type identifiable; likewise, a ‘definite’

<sup>1</sup> Using Givón’s terms, ‘specific’ and ‘referential’ as used here would equate to ‘referential-indefinite’ and ‘referential-nondefinite,’ respectively.



entity is necessarily specific, referential, and type identifiable. In the same way, an object that is only ‘type identifiable’ cannot, by definition, be specific.

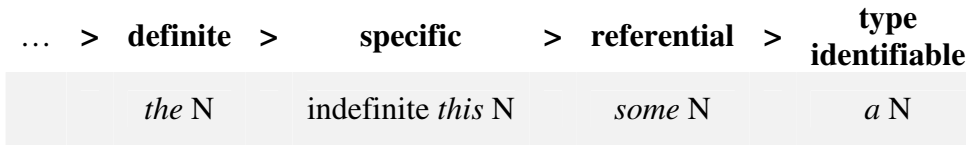


Figure 4: Modified Portion of Givenness Hierarchy

Figure 4 gives the in-focus range of the Givenness Hierarchy with the adjustments I have described above taken into account, as well as the English forms that correspond with each status. The forms are, of course, only for elucidation. I stand by the notion, expressed eloquently in Belyayeva (1997, p. 48), that “because definiteness is partly a semantic property, its extra-linguistic representation is relatively stable.” Thus while the semantic concepts relevant to the hierarchy remain applicable cross-linguistically, it is their representation via grammatical forms that can vary from language to language. The same Gundel et al. study goes on to highlight such marking strategies in a number of languages, showing that there need not be one-to-one correspondence between semantic status and expressive form. Russian and Japanese, for example, both leave definite, referential, and type identifiable unmarked, meaning the same form encodes three different statuses in the hierarchy (or four, if we are to include ‘specific’). Spanish, on the other hand, distinguishes definite entities with an article *el* but uses the same two forms *un* and  $\emptyset$  to mark both referential and type identifiable nominals. The same principle of variation in form can even apply across registers of a single language. A case in point of this would be typical absence of indefinite ‘this’ from many varieties of written English, in which the otherwise type identifiable article ‘a’ expresses the same semantic status via context.

### 2.2.2. Generic Entities and the Wheel Diagram

The (slightly modified) hierarchy proposed by Gundel et al. is sufficient for the bulk of the analysis of MA definiteness marking, but a few other concepts remain worthy of discussion. One of these is that of **generic** entities. We have already looked at the concept of referentiality and described it as the speaker's intention to refer to an individuated entity that is relevant to the world of the discourse; we have likewise seen how nominals that display this quality contrast with those that are simply 'type identifiable.' To a large degree, generic nominals overlap with the 'type identifiable' category because in using them, a speaker does not intend to refer to a specific individual of the type. They are not, however, always marked identically. Consider, alongside example (1), the following:

(5) *He wants to get a dog.*

(6) *I think the animal you saw was a dog.*

(7) *The dog descended from its ancestor, the wolf.*

In none of the three instances here does the interlocutor intend to identify a unique individual. Instead, the speaker is simply referring to the nominal entity's genus or type. Although the examples are similar in this respect, they can nonetheless be marked differently, as example (7) reveals. Givón answers this behavior by turning to the grammatical role of the non-referential item in a sentence. Accordingly, example (5) would be classed as a 'non-referential object,' (6) as a 'generic predicate,' and (7) as a 'generic subject.' These three categories, in addition to another three referential categories that Givón recognizes, together constitute the six panes of a semantic 'wheel' model (Figure 5). Givón sees the semantic wheel (again, a cross-linguistically stable outcome of human cognition) as a sort of frame around which discrete grammatical forms are arranged, typically contiguously. Cases of diachronic semantic change, he suggests, can often be modeled as the semantic extension of one or more preexisting forms to a wider range of the wheel, as in the case of *a(n)* in English, which can now represent a majority of semantic statuses, even if not exclusively.

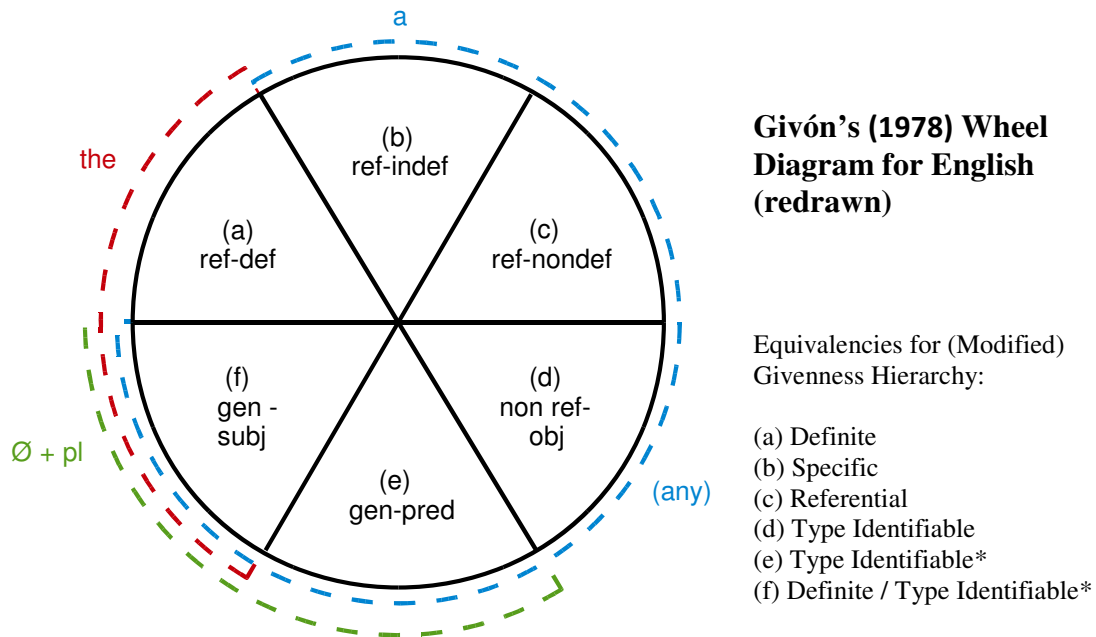


Figure 5: Givón's (1978) Wheel Diagram for English

It is possible to tentatively interface Givón's model with the modified Givenness Hierarchy we have already discussed, thus conferring the advantage of better being able to classify generic entities. Some of these equivalencies are fairly easy to establish and are given as items (a)-(d) in Figure 5. The fully generic statuses are harder to identify. Generic predicates can probably best be sub-classified as a sort of type identifiable nominal, though at least some languages (such as Hawaiian, in Givón) mark them distinctly. Generic subjects, on the other hand, are very regularly grouped with definite nominals cross-linguistically, a fact that Givón (1978, p. 298) attributes to "the quite universal overlapping of the notions of 'subject' and 'topic.'" In this light, generic subjects can probably be classified either alongside 'definite' or 'type identifiable' meanings, depending on the semantic assumptions in operation in a given language. Where such entities occur, it is nonetheless helpful to make note of their generic semantic status even if they behave syntactically as if they are definite.

### 2.2.3. Individuation and First Mention

Finally, and as a means of bridging the concepts discussed above into previous work on definiteness in Arabic varieties, I turn to a couple of key ideas that appear in Brustad's (2000) comparative dialectological work on colloquial Arabic. While Brustad does not cite Gundel et al.'s Givenness Hierarchy specifically, she touches on a number of concepts that parallel its assumptions and deserve integration into this discussion. Among the most fundamental of these is the notion that definiteness, rather than being a binary opposition, operates as a semantic continuum along which an entity may occupy various points. This concept is, naturally, crucial to the operation of the hierarchy. Relatedly, Brustad identifies the range between 'definite' and 'indefinite' as 'indefinite-specific.' Although the terminology differs, we can establish equivalencies between Brustad's categories and the given statuses we have discussed with ease: 'definite,' remains the same, her '(fully) indefinite' is our 'type identifiable,' and her 'indefinite-specific' encapsulates both 'specific' and 'referential' statuses. Figure 6 shows how these terms align.

definite	<	indefinite-specific	<	(fully) indefinite	(Brustad, 2000)		
definite	<	specific	<	referential	<	type identifiable	(Current Study)

Figure 6: Terminological Overlap with Brustad (2000)

The major factor that Brustad, building on Khan (1984), identifies as relevant to an entity's placement along this continuum or hierarchy is a concept she refers to as 'individuation.' Individuation is presented as the sum of an array of various traits – among them agency/animacy, definiteness, specificity, textual/physical prominence, qualification via descriptive language, and quantification – that together influence a speaker to syntactically mark a noun in a certain way. I hold that individuation, as Brustad describes it, is essentially another term for what we have already defined as referentiality: a measure of the speaker's commitment to the existence of a particular

object in the discursive universe. In particular, the fact that definiteness and specificity correlate so clearly with this semantic status is unsurprising when one recalls that, at least in our modified Givenness Hierarchy, referentiality is a prerequisite for both. The other traits Brustad lists (agency, prominence, qualifiability, and quantifiability), which appear to be intrinsic qualities of the entity as envisioned by the speaker rather than statuses it is accordingly assigned, that said, may aptly be called ‘the ingredients of referentiality’ – traits that differentiate an object from others of its kind and require that it be marked appropriately, if discrete grammatical structures for this sense are available.

The other relevant concept we can identify in Brustad is that of ‘first mention,’ a discursive context with unique semantic implications. When a speaker first introduces an entity into the discourse, that entity may coincide with any number of cognitive statuses on the Givenness Hierarchy. If the speaker assumes it is unique enough to be immediately recognizable, for example, it will be definite; conversely, an entity for which individuation is unneeded or irrelevant will be ‘type identifiable.’ What Brustad points out, however, is that there is a particular range of the definiteness continuum that is directly associated with – or perhaps even reserved for – the first mention of prominent entities, specifically those that constitute new topics. While she only identifies this status as “near the indefinite-specific” range, we can use our current working model of givenness to more explicitly identify it as what we have called ‘specific.’ That entities of this status are generally restricted to first mention can be explained by the fact that they are initially known to the speaker but not the addressee (see 2.2.1). Once the topic has been introduced, however, it can be assumed to be known to both participants in the discourse, and will thereafter be assigned the status of definite. For this reason, where a language has a unique form that encodes this specific tier of the hierarchy, it will coincide with the role of new topic marker, but this is not necessarily the case for all languages.

## 2.3. DEFINITENESS IN ARABIC VARIETIES

### 2.3.1. Ideological Dimensions

Before moving on to Moroccan Arabic in particular, a quick general overview of definiteness marking in Arabic is in order. In an ideal descriptive scenario, this overview would be unnecessary and Moroccan Arabic could be treated exclusively on its own terms. As I am of the opinion, however, that previous descriptions of MA have operated largely through the lens of theoretical assumptions inherited from a longstanding Arabic grammatical tradition, it seems apropos to at present what that lens is. To begin, few, if any, authors who have put forth descriptive work on MA have done so without having first engaged the linguistic conventions and ideologies of Classical Arabic (CA) and its contemporary incarnation, Modern Standard Arabic (MSA). In doing so they have inevitably been exposed to what Milroy (2001) defines as ‘standard language ideology,’ or a set of beliefs about what constitutes correct or proper language. Milroy cautions that in the presence of a standardized language, well-meaning linguists often unconsciously absorb the assumptions that surround it, which in turn inform their outlook on non-standard varieties.

In few places can this issue be seen so clearly as in the assumptions surrounding Classical Arabic (the codification of which is a topic worthy of its own discussion), which throughout its history has been regarded in intellectual circles as a “pure” or “uncorrupted” form of Arabic, contrasting with the vernacular speech that is believed to have descended from it but since become distorted. A representative quote is from Matar (1966, p. 35), who in referring to “improper speech” means that which departs from codified CA:

فإن السبب في [اللحن بمعنى الخطأ في اللغة] راجع الى أن العرب لم يلتقوا الى الخطأ في اللغة إلا حين اخلطوا ...  
...بغيرهم من أهل البلاد المفتوحة

“The reasons for [improper speech] go back to the fact that Arabs didn’t make linguistic errors until they began to mix with the non-Arab peoples of conquered territories ...”

Matar goes on to describe the history of an entire canon of Arabic literature that exists around a topic known as *laḥn al-ʿāmma*, or ‘commoners’ improper speech,’ the bulk of which was written as a guide to purging Classical Arabic of linguistic “errors” that had allegedly entered via the colloquial dialects that developed out of contact with foreigners. This ideological tradition has, in turn, undoubtedly informed the popular view of language in Arab societies, and outside of only a few academic circles there is widespread belief not only that Arabic vernaculars are the direct descendants of CA, but also that they are unworthy of study precisely because they are corruptions of it.

Linguists may have evaded the proscription against study of Arabic dialects and rejected the prescriptive use of terms such as “pure” and “corrupt” language, but – if one is to believe Milroy’s admonishments – they are still not necessarily free of these ideological assumptions and their implications. On one hand, some of the tradition’s key tenets have been internalized, as seen in the astounding number of sources that still uncritically accept modern Arabic dialects as the “daughters” of CA (see Kusters, 2003 for an example). Even where Arabic linguists have consciously sought to reject such assumptions, one can claim that their work still remains situated within an ideological battleground: in the face of a long native tradition that has considered spoken vernaculars unworthy of scholarly interest, except for where they can shed light on the classical language, framing descriptions of them in terms borrowed from CA’s grammatical tradition is a (potentially subconscious) way of highlighting similarities and legitimizing the attention given to the colloquial. Perhaps for this reason, while an Arabist may speak of the distinction between the ‘analytic genitive’ and ‘synthetic genitive’ (as in Owens, 2006), the latter term is often used interchangeably with the familiar CA term *iḍāfa*, of a similar meaning. Such co-identifications are not problematic if made discerningly; resorting to them as a simple matter of course, however, may obscure very real structural differences between Arabic varieties. I maintain that this hidden ideological dimension has played a role in previous descriptions of MA’s definiteness system, which have tended to interpret forms similar to those present in CA as semantically equivalent across varieties.

### 2.3.2. Definiteness in Classical Arabic

Like any human language, Classical Arabic can be assumed to be sufficient for expressing all of the semantic categories of givenness that we have previously discussed. Its means of explicitly marking these categories via nominal morphology, that said, is relatively uncomplicated and can be best described as making a binary distinction between indefinite and definite meanings (and, by extension, often grouping a number of distinct semantic statuses together under the domain of one form). In the case of indefinite meanings in CA, nominal stems are marked as such through a process commonly known in the literature as ‘nunation,’ where a phoneme /n/ (having the Arabic name *nūn*, from which the term ‘nunation’ is derived) is suffixed to the final case vowel in the nominal. To express definiteness, on the other hand, CA uses a prefix /(a)l-<sup>2</sup> (Wright, 1896, p. 247):

(8) *kalb-u* ‘dog’ (nominative) > *kalb-un* ‘a dog’

(9) *kalb-u* ‘dog’ (nominative) > *al-kalb-u* ‘the dog’

In keeping with the cross-linguistic tendency mentioned by Givón (2.2.2), Classical Arabic uses the latter of these forms (9) to express both referential definite entities and generic subjects, which are unsurprisingly grouped together. For other semantic statuses, the indefinite form (8) prevails. Representing CA via the wheel diagram (Figure 7) gives us a sense of the respective semantic ground covered by these two morphological possibilities.

---

<sup>2</sup> For the prefix /(a)l-/, the phoneme /l/ assimilates for the majority of coronal consonants, resulting in gemination of the first letter. One can compare *al-kalb-u* ‘the dog’ to *aš-šams-u* ‘the sun’ (both nominative). With some slight variation, the same process is true of /(V)l-/ in colloquial Arabic varieties. Refer to Figure 1 for the assimilating consonants in MA.



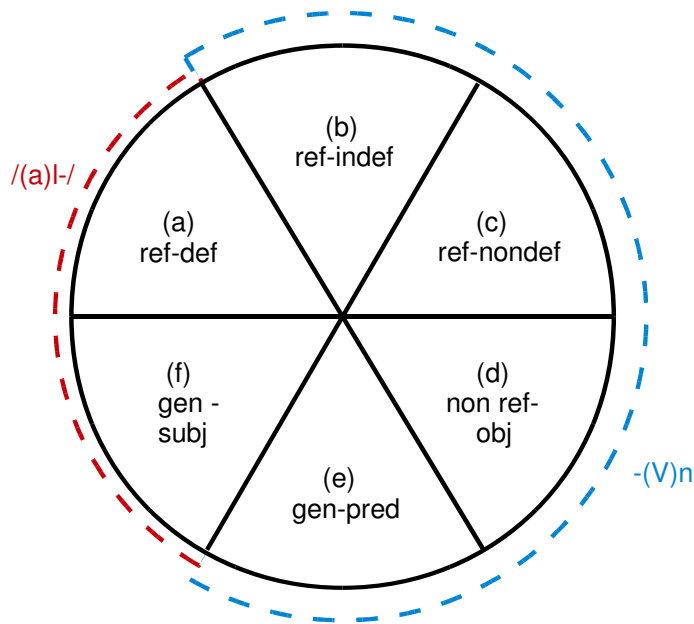


Figure 7: Wheel Diagram for Classical Arabic

To reiterate a key point, that CA does not have distinct forms for marking certain cognitive statuses does not at all imply that it is incapable of expressing their respective semantic notions. The distinction between a ‘type identifiable’ and ‘specific’ entity, for example, can still be made, but it is more dependent on discursive context and qualification of the noun via adjectives and relative clauses than might otherwise be the case. In any case, however, what is most important to our discussion is the overall relative distribution of the representative forms. Significantly, the diagram in Figure 7 makes it clear that it is the prefixless form  $-(V)n$  that has the widest semantic distribution and could easily be called the unmarked form; meanwhile the prefixed  $/(a)l-/$  is very exclusively associated with the definite/generic subject group. In this sense it is quite proper to see CA’s  $/(a)l-/$  as a true definite article.

To what degree CA coincides in this respect with the ancestors of modern Arabic dialects is an open question. There is good reason to reject the initial assumption that CA is necessarily representative of these historical varieties, with Watson (2011, p. 861) citing a growing consensus among researchers that “Classical Arabic almost certainly never reflected the linguistic system of the ancient dialects.” In regard to definiteness

marking specifically, however, CA is at least useful as a snapshot of such a system in one historically extant register. Likewise, while we cannot know precisely what other ancient varieties looked like, that a majority of modern Arabic dialects show similar semantic patterning for some forms with CA suggests that, at least when it comes to definiteness marking, CA may at least approximate the predominant historical scheme.

### 2.3.3. Definiteness in Modern Arabic Dialects

The most highly visible point of congruence between CA and modern Arabic dialects lies in the prefix /*(V)l-*/, an analog of which is present in all major Arabic dialects today.<sup>3</sup> Disregarding Moroccan for the time being, in all of these dialects /*(V)l-*/ is the corresponding form for the same semantic statuses that it indicates in CA, namely those of definite referential entities and generic subjects. On the most basic level, /*(V)l-*/ thus serves as a fully definite article, providing a semantic contrast with the indefinite meanings of prefixless nominals. To cite an example from Syrian Arabic, *il-kalb* would accordingly represent ‘the dog’ while contrasting with *kalb* ‘a dog.’ For non-referential indefinite meanings, nunation of the sort found in Classical Arabic is not present in most modern dialects, a fact that may or may not be related to today’s dialects not marking grammatical case. This means that in many varieties the central distinction between definite and indefinite nouns (that are not otherwise definite via a possessive pronoun or genitive construction) is based exclusively on the presence or absence of a prefixed /*(V)l-*/.

If we refer back to Brustad’s (2000) concept of a continuum that spans a range from fully definite to fully indefinite (or from ‘definite’ to ‘type identifiable,’ in our working terms), then, we can consider the continuum’s poles to be /*(V)l-*/ and Ø, respectively. This is a pattern that would seem to hold true across Arabic dialects, in keeping with Brustad’s comment that the definite article /*(V)l-*/ is one of the “basic morphological and syntactic properties” that “spoken and written registers of Arabic all

---

<sup>3</sup> The exceptions to this rule would be high-contact varieties such as Central Asian Arabic types (Jastrow, 2005; Ratcliffe, 2005) or Arabic-lexifier creoles such as Ki-Nubi in eastern Africa (Owens, 1985).

share” (p. 14). Where there is variation from dialect to dialect, that said, it has been described as within the range between these two values. Brustad shows there to be a number of linguistic strategies in place across modern Arabic dialects for marking what we can now recognize as the ‘specific’ and ‘referential’ tiers of the adjusted Givenness Hierarchy. For marking specificity in particular, a common strategy in urban dialects is to use a grammaticalized number *wāḥid* ‘one’ to present animates; in bedouin dialects a suffix *-in* that appears etymologically related to the nunation of CA performs a similar function. Figure 8 gives my provisional placement for some of these forms, including those mentioned by other authors (even if not specifically as markers of givenness), on the hierarchy.

	<b>definite</b>	<b>&gt;</b>	<b>specific</b>	<b>&gt;</b>	<b>referential</b>	<b>&gt;</b>	<b>type identifiable</b>	
Syrian	/il-/ N		<i>wāḥid</i> N*		ši N		Ø N	(Brustad, 2000)
Egyptian	/il-/ N		<i>wāḥid</i> N*		Ø N		Ø N	(Brustad, 2000)
Kuwaiti	/il-/ N		<i>wāḥid</i> N*		Ø N		Ø N	(Brustad, 2000)
Najdi	/al-/ N		N-in		Ø N		Ø N	(Ingham, 1994 via Brustad)
Andalusi	/al-/ N		N-an		Ø N		Ø N	(Corriente, 1977 via Brustad)
Iraqi	/il-/ N		<i>fadd</i> N		<i>fadd</i> N		Ø N	(Erwin, 1963)
Tunisian	/il-/ N		<i>fad</i> N		<i>fad</i> N		Ø N	(Watson, 2011)

\* with human animates

Figure 8: Definiteness Marking Strategies in Various Arabic Dialects

Despite a significant amount of cross-dialectical variation in the referential-specific range of the hierarchy, what is clear from this quick survey is that the fundamental dichotomy of form that we saw in CA remains: the presence of /*(V)l-*/ is associated with definite meanings, whereas its absence indicates indefiniteness. Even where there are unique markers for ‘specific’ and ‘referential’ nominals, we can still posit the indefinite zero-marker as a component of the underlying form (i.e. *wāḥid* N = *wāḥid* Ø N). In addition, some dialects have animacy restrictions for when a specific (or what we could perhaps call ‘presentative’) marker can be used, meaning the standalone prefixless form Ø still has full span of the indefinite range of the hierarchy.

What all of this suggests is that scholars with background in other Arabic varieties – meaning essentially everyone who has written on Moroccan Arabic thus far – already have a strong *a priori* reason to believe that the same associations between form and meaning will be present in MA even before scrutinizing the dialect. If the assumption that /-(V)l-/ is obligatorily a definite article is already inherited from the CA grammatical tradition, then it can only be reinforced by the observation that all of the other major dialects exclusively associate the same form with definiteness. Under the same logic, Ø can be preemptively envisioned as the primary underlying form for indefiniteness. What variation does exist in other Arabic varieties, in turn, would suggest that the most promising semantic range in which to search for other marking strategies would be in the referential-specific portions of the Givenness Hierarchy. In the section to follow, I argue that researchers have found in MA precisely what they have been looking for – sometimes in spite of the evidence otherwise.

### 3. Previous Descriptions of Definiteness Marking in MA

A number of descriptive grammars, monographs, and scholarly articles have outlined the system of definiteness marking in Moroccan Arabic (Brustad, 2000; Caubet, 1983; Harrell, 1962; Maas, 2011; Marçais, 1977; Moscoso García, 2003; Youssi, 1992). Browsing these sources, perhaps what is most striking is the degree of conformity they exhibit in doing so. On all major counts, descriptions agree: MA nouns can typically be expressed via four distinct forms, each corresponding with a different degree of definiteness. The most basic of these is represented by the “prefixless” or “zero-marked” noun (as in *kəlb* ‘a dog’), which corresponds with fully indefinite meanings and is a clear analog to the prefixless nominal form in other Arabic varieties. Similarly, and on the other side of the semantic continuum, MA is described as having a “definite article” /l-/ (as in *l-kəlb* ‘the dog’), likewise identifiable with phonologically similar forms in other dialects. Finally, and as one might expect in the light of data from other Arabic dialects, MA is given two “indefinite articles” – *waħəd* (l-) and *ši* – that indicate meanings that lie somewhere on the semantic spectrum between the zero-marked, fully indefinite Ø and the presumed definite prefix /l-/.

#### 3.1. THE PROPOSED SYSTEM

We begin with an overview of each of these proposed forms as it has been described in the previous literature. Here my intent is to extract and present the key theoretical assumptions at play for the majority of previous authors, even if they have not necessarily expressed these with the same terminology. I subsequently make an effort to align these descriptions with our working model for givenness and its subsidiary semantic statuses as described in section 2.2. These generalizations are afterwards subjected to critique and adopted as a point of comparison for my own reanalysis of the forms and their relationship with definiteness both synchronically and diachronically.

### 3.1.1. Proposed Unmarked Form Ø

Central to the general outlook of previous MA descriptions, and in keeping with the patterns we have seen in CA and other Arabic varieties, is the presumably axiomatic distinction between a prefixless indefinite form and a /l-/prefixed definite form. This view is perhaps best summarized by Marçais (1977, p. 160), who makes explicit reference to the perceived similarities between CA and MA:

Tous les parlers maghrébins connaissent, dans l'emploi du nom, l'opposition indétermination-détermination, comme il en est en arabe classique. L'indétermination est caractérisée par la représentation du nom à l'état nu, au degré zéro, c'est à-dire sans article ni complément déterminatif. La détermination est caractérisée par la préfixation au nom d'un article, ou par l'annexion au nom d'un complément déterminatif.

“All Moroccan dialects realize, in the use of the noun, the interdeterminate-determinate opposition as it exists in Classical Arabic. Indetermination is characterized by the representation of the noun in its bare state, zero-marked, i.e. without either article or determiner. Determination is characterized by the prefixation of the noun with an article, or by its annexation by a noun that has a determiner.”

Marçais goes on to claim that the prefixless form represents the noun in its “most general sense,” and that “it compares with nunation in Classical Arabic,” giving the examples *raʒəl* ‘man,’ *mʁa* ‘woman,’ and *šəy* ‘thing.’ One can interpret this description as meaning that the prefixless noun is the default (or unmarked) nominal form from which others are derived.

Consistent with this description and the expectations to which it gives voice, other works invariably describe the prefixless MA noun as “indefinite” and maintain the assumption that it is the most basic nominal form. Youssi’s analysis agrees with Marçais’s, stating unequivocally that the prefixless form he gives as Ø refers to the unspecified or unmarked state of the noun, is associated with “absolute indetermination,” and is equivalent to an indefinite article. He gives such examples as:

(10) ...*rtakəb ʒarima* ‘he committed a crime’ (Youssi, 1992, p. 144)

- (11) *ka-tfəkkru f n-nsa... bħal ila kanu fwakih* ‘you all think of women as if they were fruit[s]’ (ibid.)

Harrell likewise highlights the indefinite nature of the prefixless noun by giving its English translation equivalent as “a” or “some.” He identifies some of the primary contexts in which the form occurs as that of a predicate complement, in “negative expressions,” and with “the specific numerical meaning of one,” giving, among others, the examples:

- (12) *huwa qadi* ‘he is a judge’ (Harrell, 1962, p. 187)

- (13) *dəwwzət ʕəndu ʕam* ‘she spent a year with him’ (ibid.)

- (14) *w ma-kayn-š lli iqdər iqul lha hətta kəlma* ‘and there’s no one who can say even a single word to her’ (ibid.)

Finally, more recent work by Maas (2011) builds on these assumptions, more exactly stating that “zero-marking defines a term as non-referential,”<sup>4</sup> a semantic status that would coincide with ‘type identifiable’ in our modified version of the Givenness Hierarchy. Maas gives as evidence:

- (15) *küll nhar ka-ntiyyəb hawli* ‘every day I cook a sheep.’ (Maas, 2011, p. 154)

The interpretation of the prefixless form as non-referential or ‘type identifiable’ corresponds with Caubet’s (1983) claim that “one can assign it the metalinguistic value of ‘any one X,’”<sup>5</sup> itself another way of saying that it is unimportant that the nominal entity be individuated within the discourse (or, put shortly, be referential). On the whole, then, previous descriptions seem to be in full accord on the role of the prefixless form as unmarked, indefinite, and semantically analogous to prefixless forms in Classical Arabic and other varieties. Where the distinction is made, it is also identified as the primary means of marking nouns exhibiting a ‘type identifiable’ semantic status.

---

<sup>4</sup> “Die Nicht-Markierung (Ø) definiert einen Term als nicht-referenziell.”

<sup>5</sup> “...on peut lui attribuer la valeur métalinguistique de «UN X QUELCONQUE».”

A point of interest that remains, however, is the fact that at least a few authors have voiced an opinion that the Ø-marked form does not in fact occur all that often in actual MA speech. Even while Harrell gives common contexts for its use, for example, he states that “prefixless nouns are of relatively restricted occurrence,” though he does not clarify exactly where Ø fails to occur when we might otherwise expect it to. This feeling is perhaps reflected in Chafik’s (1999) comment that “If you hear ‘*šūft ražal*’ [‘I saw a man’ (without /l-/)] you can know that a speaker is either of Arab [Bedouin] origin and has grown up exclusively around them, or that he is a graduate of some sort of Arabic college.”<sup>6</sup> What both comments imply is provisional recognition that prefixless nouns of the type described here cannot, in the speech of most Moroccans (among them those whose speech approximates the koiné), simply occur in any position, but little work has been done to identify precisely what these restrictions entail and why they might occur.

### 3.1.2. Proposed Definite Article /l-/

If the indefinite and prefixless nominal form Ø represents, as Marçais and other authors imply, one side of an indeterminate-determinate opposition in MA that resembles that of CA, it follows that the other side would be represented by nouns prefixed with a determiner /l-/. On the nature and function of this form, descriptions are again in accord, unanimously calling /l-/ a “definite article.” Harrell (1962) sees /l-/ as a “modification” of the noun that is used to indicate definiteness and is often translatable as the English article *the*, as in *l-bab* ‘the door.’ This impression is echoed by other scholars. Youssi (1992, p. 141), for example, likewise considers /l-/ to be an article that is appended to the unmarked form, and exclusively translates it as French *lalle*, as in *l-məbʕut* ‘the envoy’ (‘l’envoyé’) and *l-bənt* ‘the girl’ (‘la fille’). Moscoso García’s (2003) Spanish renderings of what he calls “el artículo definido” are similar, with *ella* consistently given as equivalents in examples such as *l-ħlib* ‘the milk’ (‘la leche’) and *l-ɣyal* ‘the boys’ (‘los niños’). All of these representations corroborate Caubet’s (1983) assertion that /l-/

---

<sup>6</sup> "وإن تسمع «شفت رجل» فاعلم أن المتكلم إما عربي الأصل والمنشأ (بين الأعراب خاصة)، وإما هو خريج مدرسة عربية ما."



is an article “associated with the type of determination wherein an item is distinguished”<sup>7</sup> (presumably meaning ‘uniquely identifiable’).

No sources of which I am aware stray far from this general depiction of /l-/ as a definite article, corresponding not only with the phonologically similar form /(a)l-/ in CA but definite articles in European languages as well. Titles such as Heath (2002, p. 252) largely forgo discussion of the form’s semantics, which they seem to take for granted, simply calling it a “definite prefix” and moving directly into discussions of its phonological realizations without further investigation of its meaning. A few works, however, have at least hinted at a more complicated picture, even while they maintain the same terminology of “definite article” to refer to /l-/. Harrell, for example, does mention that “the Moroccan definite article has a much wider range of use than the English definite article” (Harrell, 1962, p. 190) and shows that it can be used to refer to, in addition to previously known entities, “abstractions” and “categories as a whole.” He gives the examples:

(16) *s-siba hadi!* ‘this is anarchy!’ (Harrell, 1962, p. 190)

(17) *l-yum huwa labəs s-səlham* ‘today he’s wearing a cloak’ (ibid.)

While the usage in (16) is typical of most Arabic varieties, where abstract nouns are always expressed with an article /(V)l-/, (17) is notable in that /l-/ does not appear here to refer to a referential entity. Harrell is not alone in having noted such occurrences. Brustad, for example, cites a number of nouns in her own data that likewise do not appear to refer to uniquely identifiable entities despite the presence of an element /l-/. Among these are:

(18) *hada waḥəd r-ražəl ma-ʕəndu-š l-wlad, ʕəndu yir l-mra* ‘there’s this man who doesn’t have children, he just has a wife’ (Brustad, 2000, p. 36)

(19) *dbəḥ t-tur, ʕrəḍ ʕla n-nas* ‘he slaughtered a bull, invited people’ (Brustad, 2000, p. 37)

---

<sup>7</sup> “L'article «əl» et, lui, associé à ce type de détermination où un objet est distingué.”

In such cases, Brustad seeks to explain the marking of nouns with /l-/ as a result of their relative animacy, relevancy to the text, and retrievability via shared cultural expectations, all of which could foreseeably push them “towards the definite end of the continuum;” this explanation appears to be an attempt to reconcile the /l-/prefixed nominal forms with prior expectations for what their semantics would entail. Maas, for his part, encounters a similar conceptual issue with such as examples as:

(20) f l-ləwwəl ma-ka-ntʃwwəd baʃ nʃuf d-dəmm ‘*in the beginning I wasn’t used to seeing blood*’ (Maas, 2011, p. 157)

(21) w dda mʃahūm kadalika l-kura baʃ ilʃəb l-wliyəd ‘*they also brought along a ball so the boy could play*’ (Maas, 2011, p. 158)

Maas’s solution for nouns of this type is to introduce a category he refers to as “discursive definite,” which includes both referential definite nouns as they are typically understood as well as “generic definites.” That at least some generic entities, particularly verbal subjects, would take the same morphological marking as definite-referential objects is not particular surprising if we consider Givón’s comments on the cross-linguistic prevalence of this grouping (2.2.2). It is unclear, on the other hand, what would semantically distinguish one of Maas’s “generic definite” entities from a zero-marked non-referential one in the case that it were a verbal object.

What emerges from this discussion is that at least some authors have hinted at challenges to the notion of /l-/ in MA as a definite article in the traditional sense, meaning one that refers to entities that are uniquely identifiable for both the speaker and the listener (perhaps as well as generic subjects). None have, that said, ever truly abandoned the claim that /l-/ is indeed a definite article, instead opting to explain apparently ‘type identifiable’ nouns marked with /l-/ as still somehow semantically definite via roundabout means. While I contend that these explanations have not been satisfactory, in taking them scholars have nonetheless managed to align their descriptions of /(V)l-/ in MA with that of it in other Arabic dialects, calling it a “definite article” and in turn suggesting it corresponds with a ‘definite’ status in the Givenness Hierarchy.

### 3.1.3. Proposed Indefinite Article *wahād* (/l-/)

In addition to the forms Ø and /l-/, both of which are familiar from other Arabic varieties, almost all descriptions of Moroccan Arabic recognize an “indefinite article” variously represented as *wahād* or *wahād* /l-/. Harrell (Harrell, 1962, p. 189) describes this article as “concretizing” and gives its translation equivalent in English as *a(n)*, but does little else to distinguish it from Ø besides stating that it “always refers to something clearly specific.” The syntax of *wahād*, according to Harrell, requires that the following noun have the definite article /l-/ unless it explicitly excludes it, as shown in his given examples *wahād r-ražal* ‘a man’ and *wahād l-ktab* ‘a book.’ Harrell also gives the example *wahād bəllarž* ‘a stork,’ where the noun *bəllarž* cannot, for reasons that appear tied to its lexical status, take what he refers to as the definite article, but no particular explanation is given for this behavior (see 3.2.2).

Other authors have given descriptions of *wahād* that parallel Harrell’s while offering refinements. Youssi (1992, p. 146) repeats the claim that the article *wahād* requires the following noun to be marked with the definite article /l-/ and argues that this use of *wahād* is syntactically identical to the Arabic construct state, in which the former of two nouns is annexed by the latter and in turn receives its definiteness status (Youssi gives *šahib l-məqha* ‘the café owner’ as an example, which would presumably compare with *wahād l-məqha* ‘a café’). It remains unclear, however, how or why this sort of annexation – which typically establishes the entire noun phrase as definite – would instead contribute to an indefinite meaning, as is typically the case when *wahād* occurs with nouns. In any case, Youssi does indicate that *wahād* can be used to indicate a specific entity<sup>8</sup>, a meaning that is not explicitly clear but can at least be inferred from context in his examples:

(22) *wahād l-məsɔala lli imkn ntkəllmu fliha* ‘an issue we can discuss’  
(Youssi, 1992, p. 146)

(23) *šəft wahād l-məžmuʕa ka-ydəhku* ‘I saw a group [of people] laughing’  
(ibid.)

---

<sup>8</sup> As witnessed in his translation for *wahād l-insan*: “une (certaine) personne.”

Brustad's treatment of the same construction, which she identifies as *waḥəd /l-/*, goes into more depth on its semantic connotations. Where Harrell and Youssi make only passing reference to the 'specific' quality of nouns marked with the article, Brustad describes specificity as a key feature of nouns marked with *waḥəd /l-/*. In particular, she notes its use as a "new topic marker" that is used to introduce textually prominent entities of high relevance to the upcoming discourse, equating it with similar uses of the etymologically identical *wāḥid* in other dialects. In keeping with this interpretation, she often translates the article as the indefinite 'this' of colloquial English rather than simply 'a(n).' Among the occurrences of *waḥəd /l-/* she cites are:

(24) *kayn waḥəd n-nuḥ axūr dyaḥ l-ḥut* 'there's this other kind of fish...' (Brustad, 2000, p. 33)

(25) *naḍ tẓəwwəž mra x<sup>w</sup>ra, tẓəwwəž waḥəd l-mra x<sup>w</sup>ra* 'he up and married another woman, married this other woman' (Brustad, 2000, p. 35)

As we saw in our earlier discussion of semantic statuses and givenness, both Brustad's identification of *waḥəd /l-/* as a new topic marker and her translation of it as English 'this,' which would occupy a similar semantic region, suggests that it does indeed correspond with the specific status of our modified Givenness Hierarchy. This is further established via the respective contexts of the above in-text usages of *waḥəd /l-/*, all of which can be read as instances of the speaker introducing a referential entity that he or she can uniquely identify but that is assumed to be new to the speaker. Appropriately, this view is explicitly confirmed in Maas (2011, p. 155), who not only shows that *waḥəd /l-/* refers to a referential entity known to the speaker but not the addressee, but also plainly calls it "specific."<sup>9</sup>

The semantic status indicated by *waḥəd (/l-/)*, then, seems to be clearly established in the preexisting literature as what we here call 'specific' as well, and there seems little reason to challenge this generalization on the basis of either other authors' or my own textual evidence. The primary question associated with this article, instead, lies

---

<sup>9</sup> "Bei [+ spezifisch] sind diskursive Differenzierungen grammatisiert: hier kann die referenzielle Bestimmtheit für den Sprecher gegeben sein, nicht aber für den Hörer."

in its syntactic behavior. In particular, none of the surveyed authors, despite their mostly accurate descriptions of *waḥad* (/l-/)'s semantic connotations, have ever adequately addressed a key conceptual issue related to it: namely, why a form that is exclusively used to indicate indefinite meanings would require what has otherwise been described as a definite article (/l-/) as part of the construction. This arrangement appears to be unique to Moroccan, because even the other Arabic dialects that do have a similar specific article (or presentative marker) *wāḥid* of the same etymological origin prohibit it from being followed by a definite entity (Brustad, 2000).<sup>10</sup> While the meaning of the article where it occurs is thus clear, its form can be considered problematic for current descriptions. This factor likely plays into authors' apparent uncertainty about whether /l-/ is simply part and parcel of a single article (as in *waḥad* /l-/, given by Brustad and Maas) or the second in a series of two articles where the former requires an otherwise definite construction to follow (as in *waḥad* + /l-/, in Harrell, Youssi, and Moscoso García).

#### 3.1.4. Proposed Indefinite Article *ši*

The final nominal form that has been widely recognized in MA is described as marked by yet another indefinite article, consistently identified as *ši*, which is prefixed to nouns that have otherwise been described as zero-marked. Harrell (1962, p. 189) contrasts this article with *waḥad*, calling it instead a “potential” article and giving the approximate English equivalent ‘some,’ as in *ši ɗar* ‘some house (or other),’ and *ši-ḥaža* ‘something.’ Other authors writing in English have echoed this choice of translation, with Brustad giving *ši* as “some (sort of).” Like *waḥad* /l-/, she sees the form as an “indefinite-specific” article indicating a greater degree of individuation (or perhaps what we would call referentiality) than would an unmarked noun. Among her examples are:

- (26) *ka-ybqa iqul ši kəlma qbiḥa, ka-yqul ši māsadil qbiḥa* ‘he keeps saying some nasty word, he says some nasty things’ (Brustad, 2000, p. 27)

---

<sup>10</sup> This in addition to the fact that other dialects have animacy restrictions for /wāḥid/, meaning it must make reference to a human actor, whereas Moroccan's equivalent can refer to inanimates, as in (24).

(27) *w ana ʕndi ʃi nas dīfan* ‘while I had some people as guests’ (ibid., my translation)

The French literature on MA generally reflects this interpretation, with both Caubet (1983) and Youssi (1992) typically translating *ʃi* in its articular sense as ‘une (certaine),’ or ‘quelques’ with plurals. Caubet, specifically, describes *ʃi* as a “quantifier” that indicates an entity’s unique existence (or what we might again interpret as ‘referentiality’) but does not necessarily specify an exact number or identify the referent uniquely. Various examples of hers include:

(28) *ʕtini ʃi zɫafa!* ‘give me a bowl!’

(29) *dəxlət ʃi bənt* ‘some girl came in’

(30) *ʕtawni ʃi ktub* ‘they gave me some books’

Maas, again, is more exacting, assigning nouns marked with *ʃi* the unique semantic status of ‘referential’ (but not ‘specific’), which overlaps quite ideally with our working Givenness Hierarchy. As previous discussion would lead us to expect of such nouns, Maas shows that their existence as individuated entities is important to the discourse, but that neither the addressee nor the speaker is able to uniquely identify them. Accordingly, in the example below, we can extract that the speaker assumes that a ‘straw-like’ entity is important to the discourse, but is not necessarily sure it must be ‘straw’ in a strict sense:

(31) *f ʃi tbən, wəlla...?* ‘in some [sort of] straw, or...?’ (Maas, 2011, p. 155)

From this brief review of previous work we can conclude that article *ʃi* has probably enjoyed the most accord in descriptions of MA’s definiteness system and does not present any immediate descriptive challenges in the way that other forms do. Adopting Maas’s interpretation, which gives a name to a semantic sense others have hinted at through their translation equivalents, we can thus consider *ʃi* to be an indefinite

marker corresponding with the ‘referential’ tier of the current study’s modified Givenness Hierarchy.

### 3.1.5. Overview of Previously Proposed System

Taking the points of agreement among previous studies into account, we are left with a fairly clear view of the system of definiteness marking in MA as it has been envisioned to date. If the above authors are correct, we would have a four-form system where each form corresponds with a unique semantic status. We can identify these proposed correspondences as:  $\emptyset$ : ‘type identifiable,’ *ši*: ‘referential,’ *wahəd* (/l-): ‘specific,’ and /l-/: ‘definite.’ This ‘definite’ status may include “generic” or “categorical” entities as described by Maas and Harrell as well. Maas gives a particularly useful diagram for understanding these semantic relationships (Figure 9), which likewise serves as an apt summary of these previous descriptions’ conclusions.

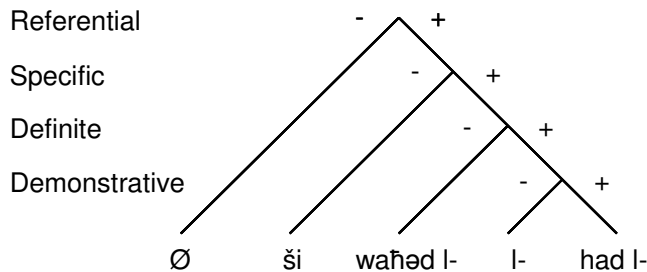


Figure 9: Proposed Semantic Statuses of MA Forms

(Maas, 2011, p. 156)

Disregarding the demonstrative tier that Maas includes (given that, in Arabic, a demonstrative implies that the following noun be marked for definiteness anyhow), we can easily map the four remaining forms into a simpler representation of our working hierarchy. This gives us a model by which we should theoretically be able to predict semantic relationships for different nominal forms with consistency and accuracy. To this effect, Figure 10 gives the semantic statuses of the Givenness Hierarchy from ‘definite’ through ‘type identifiable,’ their corresponding forms in MA as proposed in previous

descriptions, and examples of both an animate and inanimate entity expressed grammatically with these forms.

... > <b>definite</b> > <b>specific</b> > <b>referential</b> > <b>type identifiable</b>				
/l-/ N	<i>waḥəd</i> (/l-/) N	<i>ši</i> N	Ø N	
<i>l-wəld</i> 'the boy'	<i>waḥəd l-wəld</i> 'this boy'	<i>ši wəld</i> 'some boy'	<i>wəld</i> '(a) boy'	
<i>l-kas</i> 'the cup'	<i>waḥəd l-kas</i> 'this cup'	<i>ši kas</i> 'some cup'	<i>kas</i> '(a) cup'	

Figure 10: Proposed MA Forms for Givenness Hierarchy

I now turn to my own data. Indeed, for a number of occurrences of nominal entities, the above description works quite well. If chosen selectively, in fact, one can find myriad examples that would seem to confirm the efficacy of this prevailing model for definiteness marking in MA. We can cite, for example, the following instances of /l-/ used alongside contextually definite (either via previous mention in the discourse or inherent uniqueness) nouns:

(32) *ta-tdxəl l-ḡərusa təmma ta-ttysəl w ta-tx<sup>w</sup>ərž* ' the bride goes in there, washes herself and comes out' (9-18)

(33) *mnin ka-yži l-ḡimtiḥan, yəḡni ḡaruri xşş l-ḡinsan inuḡ mḡa s-sbaḥ* 'when the exam comes around, a person has to get up in the morning' (8-9)

(34) *yadi nnəqşu l-butə, w yadi nxlliu-ha tṭiyb mzyan* 'we'll turn down the stove, and let it boil really well' (3-4)

(35) *ka-ydiru l-fathə, şafi w ka-tmši l-ḡərus* 'they read the fatiha,<sup>11</sup> that's it and the bride goes off' (2-15)

<sup>11</sup> The first chapter of the Qur'an.



- (36) *l-ḡayalat ka-yštəḥu, ta-yḡanniu, ta-yduz n-nhar b xir w ḡala xir* ‘the women dance, sing, the day goes by wonderfully’ (9-11)
- (37) *l-wəzīr ka-yddi-h l l-mdina, ka-ybqa iṣṣayl-u tamma* ‘The wezīr [‘best man’] takes him to the city and keeps him occupied there’ (9-14)

We can also pinpoint a number of instances of /l-/ used with generic subjects. As established in 2.2.2, this is neither surprising nor problematic in the light of Arabic varieties’ habit of grouping generic subjects with referentially definite entities, as well as a cross-linguistic tendency toward the same:

- (38) *daba l-wəld ila bya itṣəwwəṣ l-bənt ta-ymši... l l-walidin dyalu ta-yxṭṭəbu-h* ‘now if a boy wants to marry some girl he goes... to his parents and they get him engaged’ (2-1)
- (39) *mnin ka-yṣi l-ḡimtiḥan, yəḡni ḡaruri xṣṣ l-ḡinsan inuḡ mḡa ṣ-ṣbaḥ* ‘when the exam comes around, a person has to get up in the morning’ (8-9)
- (40) *l-qəḥwa ta-tḡirha f l-briq, ta-tḡir ḡlah s-skḡar...* ‘[as for] coffee, you put it in the coffeepot, you add sugar to it...’ (4-6)

For the proposed article *wahəd* (/l-/), most uses seem to confirm with what we would expect from the above model, conveying the sense that a noun is uniquely identifiable to the speaker but not the addressee (i.e. ‘specific’ but not ‘definite’). Most likewise seem to confirm that /l-/ , as described, is a necessary part of the syntactic construction:

- (41) *ka-nqəbṭu... wahəd l-ḡina? dya... t-ṭin, w ka-ndiru fih... ka-ndiru fih l-lḥəm* ‘we grab this [certain] pot made out of clay and put meat in it’ (4-4)
- (42) *ta-yḡəlsu f wahəd l-kūrsi mūḡəyyən f wahəd ṣ-ṣala kbira* ‘they sit in this special chair in this big hall’ (9-8)

The same can be said of *ṣi*, which often appears to mark an expectedly referential meaning, where the individuation of the noun is of relevance to the discourse although the speaker cannot uniquely identify the referent:

- (43) *ka-txəlləş-u, ka-təʕti-h ʃi baraka, imma ʕəʃra drahəm wəlla xəmʃtaʃər dərham wəlla miyət dərham, ʃhal ma kan* ‘you pay him, give him some sort of tip, either ten dirhams or fifteen dirhams or a hundred dirhams, whatever you’ve got’ (4-16)
- (44) *w byina ʃi mihraʒan wəlla ʃi məʕrid, baʃ ttʕərrəf z-zərbiya l-wawazgitiya* ‘we want some sort of festival or some sort of exposition, so that carpets from the Wawazgit region will get exposure’ (1-22)

Finally, my data does include a few instances of hypothetically “zero-marked” or “bare” nouns, most of which can be interpreted as ‘type identifiable’ (or non-referential) and potentially translated as ‘a(n).’ A number occur as objects, as in:

- (45) *ka-taxʷəd hiya ʕəsa, ka-yaxʷəd huwa ʕəsa* ‘she takes a stick, he takes a stick’ (9-29)
- (46) *n-nhar t-tani... ka-ndəwwzu l-ʕulum w... mʷadda xʷəra nsit-ha* ‘[on] the second day we take sciences and another subject, which I’m forgetting’ (8-17)
- (47) *kaynin l-ʕayalat daba lli dayrin ʒəmʕiyat w tʕawniyat nisaʕiya* ‘there are these ladies nowadays who have put together women’s associations and cooperatives’ (1-18)

In addition, we can also find the proposed bare form Ø as a generic predicate:

- (48) *ana təlmid... b s-sana t-tanya ʔakaʕurya* ‘I’m a student in the second year of the baccalaureate’ (8-1)
- (49) *n-nhar l-axʷər yadi tnuḍ, ka-tʃqa f ʔarhūm ʕəla asas annəha ʃafi raha wəllat mərə ka-ttʕətamd təmmaya* ‘the next day she’ll get up and do housework, with the logic that she’s now become a woman and is being depended upon there’ (9-43)

I present the nominal entities in the above contexts precisely because they appear to substantiate the dominant view of how definiteness in MA is patterned. In doing so, that said, I do not necessarily seek to defend it. To the contrary, what I mean to show with these examples is how easily an author who already believes the MA system must be semantically similar to that of other Arabic varieties, and who has internalized the

generalizations made by his predecessors, can select for – or perhaps even invent, in the case of those whose grammatical descriptions do not draw on naturally occurring data – forms that correspond with whatever meaning is preemptively expected of them. I contend that this sort of selective engagement is exactly what has informed most previous scholars’ views on the MA definiteness system.

The issue with this approach, that said, is clear: the nouns from my own examples above have been explicitly selected to suit a preconceived model, and are not fully representative of the whole set of data. The fact of the matter is that if one attempts to look at the semantic connotations of various MA nominal forms objectively, a variety of cases arise that “break” the traditional descriptions of definiteness marking. In what follows I present these challenges, and how I believe the MA system needs to be re-envisioned in order to appropriately meet them.

### **3.2. CHALLENGES TO PREVIOUS DESCRIPTIONS**

The primary issue with previous descriptions, and one that some authors have hinted at but never explicitly stated, is that Moroccan Arabic’s “definite article” /l-/ does not in actuality seem to clearly correlate with definiteness. This is not to deny that semantically definite nouns of Arabic origin overwhelming do display the morphological form /l-/ (as in examples 32-40), but at the same time there is ample evidence that /l-/ can likewise be found throughout the indefinite range of the semantic continuum, including where previous descriptions would lead one to expect a “bare noun” that would resemble the prefixless indefinite form of other Arabic varieties. Inversely, there have been shown to be a group of MA nouns – which I argue below is an open class, and larger than typically assumed – that are not, even in semantically definite contexts, marked with the form /l-/. Both of these observations challenge the notion that /l-/ is truly a definite article, an assumption that is nearly universal in studies of Arabic varieties. Here I present in detail the evidence in favor of disassociating /l-/ with definiteness.

### 3.2.1. Presence of /l-/ in Indefinite Contexts

While none of the previous literature has taken the step of abandoning the term “definite article” when referring to /l-/, at least some works have already conceded that the semantic range of /l-/ is broader than one might expect of such an article. Harrell’s (1962, p. 190) comment that “the Moroccan definite article has a much wider range of use than the English definite article and is often used in situations where English uses no article at all” reflects such an observation, as perhaps does his insistence that /l-/ can be used to refer to categories as a whole (as in example 17). We find a similar conceptual issue in Brustad, who goes to great lengths to explain the presence of /l-/ in contexts where most evidence would otherwise suggest the nouns are fully indefinite (18 and 19). Finally, that Maas must invent a category for “discursive definites” to accommodate the somewhat problematic “generic definite” entities he encounters (as in 20 and 21), similarly, indicates that he does not feel that the regularly understood connotations of definiteness are appropriate for many nouns that incorporate /l-/.

My data presents the same complications, with /l-/ appearing in a number of semantic contexts that are difficult to identify as anything other than indefinite. In addition, such occurrences can be found not only in one subset of the indefinite semantic range, but throughout all of the indefinite tiers of the Givenness Hierarchy. It is reasonable, then, to treat these semantic statuses one by one as a means of elucidating over how wide a range /l-/ can truly occur.

#### 3.2.1.1. /l-/ with ‘Specific’ Statuses

We begin with the ‘specific’ status, or that which refers to a referential entity that can be uniquely identified by the speaker but not the addressee. As we have already seen, one of the primary forms for marking this status is what has previously been described as an article *waḥad* (/l-/), given in others’ examples (22-25) as well as my own (41-42). The parenthetical (/l-/) is an indication of the general agreement among scholars that the syntax of this construction requires that first component *waḥad* be followed by a “definite” noun phrase, in most cases marked by /l-/. What is surprising that no author has pointed out, in this light, is how the proposed semantics of the “article” are

themselves self-contradictory: if the case is truly as it has been described, it involves marking a nominal for indefiniteness (*waḥād*) and definiteness (/l-/) simultaneously, the equivalent of saying ‘*a the dog*’ in English. This apparent contradiction would apply to the majority of the uses of *waḥād* (/l-/).

Typically, the *waḥād* (/l-/) construction refers to a singular noun.<sup>12</sup> All of the previous examples of the proposed article elucidate this usage, and we can easily find more examples where it occurs alongside indefinite-specific meanings, such as:

(50) *ya-nqaddām waḥād l-wasfa dyal... wasfa myaribiya, ṣibara ḥla ḥarša*  
‘I’m going to present this recipe, a North African recipe, known as *harsha*’ (3-1)

(51) *ka-ykun bzzaf, xaṣṣetan waḥd n-nuṣ dyal l-ḥarira ka-tkun, küll waḥd*  
*ka-ydir l-ḥarira f l-ftur* ‘there’s a lot [of food], especially this  
[certain] type of *harira*; everyone makes *harira* for iftar’ (7-12)

Since in many MA dialects *waḥād* (/l-/) is restricted to single entities, a question that perhaps follows is what sort of marking may accompany plural nominals of the same ‘specific’ semantic status. My reading is that these entities, too, are often expressed via forms involving /l-/ , but involve no additional marking, which would make them formally identical to what has previously been described as ‘definite.’ A couple of examples from my oral text corpus are the following, in which we can assume a ‘specific’ meaning for the underlined nouns on the basis that they represent new topics of which the addressee would not be previously aware:

(52) *kaynin l-ṣayalat daba lli dayrin žəmṣiyat w tṣawniyat nisaḡiya* ‘there are these ladies nowadays who have put together women’s associations and cooperatives’ (1-18)

(53) *l...ntaḡiž ka-ybiynu fihūm lli nžəḥu, w kaynin n-ntaḡiž xwəra*  
*ka-ybiynu l... lli ma-žabu-š n-nūqta mzyan ṣandhūm l-ḡistidraki* ‘the results make it clear who passed, and there are these other results that show the... those who didn’t get a good score have the remedial exam’ (8-25)

---

<sup>12</sup> Though Harrell does note that it in the dialect of Fes, it can be used with plurals, which would indicate extension of its grammatical scope.

The generalization we can extract from these examples is that the proposed “definite article” /l-/ appears alongside what is probably a large majority of nominals of the status ‘specific,’ which by definition must be indefinite. This serves as preliminary evidence that /l-/ is violating the semantic restrictions on an exclusively definite article.

### 3.2.1.2. /l-/ with ‘Referential’ Statuses

As we move further to the right on our working implicational hierarchy, we arrive to the status of ‘referential,’ associated with individuated entities whose existence has been established as relevant to the discourse but which neither the addressee nor the speaker can uniquely identify. Previous work has shown that one of the primary means of marking this status is a proposed article *ši*, as in examples 28-31 and 43-44. There is nonetheless some tentative evidence that /l-/ can also represent solely ‘referential’ entities, although this is harder to confirm without direct insight into the speaker’s communicative intentions.

To make this distinction I therefore rely on Givón’s (1978, p. 296) definition of ‘non-definite,’ which we have already noted is in essence equivalent to ‘referential:’ “while the verbal expression indicates that the speaker is committed to the existence of some individual, the actual identity of that individual is left unspecified, presumably because it is of no import in that particular communication.” Following this description, I identify the following as occurrences of the form /l-/ in non-specific ‘referential’ contexts:

(54) *daba l-wəld ila bya itžəwwəž l-bənt ta-ymši... l l-walidin dyalu ta-yxṭṭabu-h* ‘now if a boy wants to marry some girl he goes... to his parents and they get him engaged’ (2-1)

(55) *w ka-ntmnnau zəfma y... ikun s-suq hnaya, hna f taznaxt, maši f mərrakš, ikun t-tabṣ dyal z-zərbiya hnaya f taznaxt* ‘we wish we had some sort of market [for them] here in Taznakht, not in Marrakesh, and that the impact of the carpet [industry] was here in Taznakht’ (1-21)

The appearance of /l-/ as an alternative to *ši* in these solely referential contexts, where not even the speaker can uniquely identify the individuated entity, casts even further doubt on the notion that /l-/ displays the traits we would expect of a true definite article.

### **3.2.1.3. /l-/ with ‘Type Identifiable’ Statuses**

Finally, and perhaps not unexpectedly now that we have already seen instances of the proposed “definite article” in co-occurrence with other indefinite meanings, we can also find numerous occurrences of /l-/ with apparently non-referential entities that are best described as exclusively ‘type identifiable’ only. It is worth noting that other authors have already identified the presence of such nouns (as in examples 17-21) and grappled with them, arriving at somewhat different solutions to the implied problem. Brustad entertains the idea that these seemingly indefinite nouns may actually be more specified in the mind of the speaker than is apparent from the context, a function of their relative individuation and animacy, and that this identification grants them more “definiteness.” Harrell and Maas, on the other hand, frame the presence of /l-/ as a syntactic rule, where “categories as a whole” or “generics” must be marked for definiteness. While this is plausible for the subject position – where the overlap between subject and topic, according to Givón (1978), often results in generic entities being treated as syntactically definite -- it is unclear why one would expect this to be the case for non-referential objects.

My opinion, on the other hand, is that the previous suggestions are overly complex solutions to a simple problem, born out of a desire to make apparently discrepant forms and meanings “work” in a preconceived model. The solution I propose is instead to take such nouns as exactly what they appear to be: exclusively ‘type identifiable’ entities, for which the speaker is not committed to the existence of any one individual (but rather that of the class), and which involve no element of definiteness whatsoever. In addition to the other authors’ examples, I provide the following:

- (56) *ɣadi nxəlltu l-ʕənaʕir, w ɣadi nax<sup>w</sup>du l-məqla, w ɣadi nxwiu-h f l-məqla, w ɣadi nṭiybu-ha fuq l-buṭa* ‘we’re going to mix the ingredients together, get a frying pan, pour it into the frying pan, and heat it over the stove’ (3-3)
- (57) *f l-luwl ka-ndiru-h... f l-kas, mn bʕdət tani ka-nʕəlləlu ḥbub atay...* ‘at first we do that in a cup, then after that we rinse the tea leaves again’ (5-5)
- (58) *kayn-ši... ɣir ka-yləʕbu ɣir binathūm, ma-kayn l-g<sup>w</sup>ərb wəlla ḥtta ši ḥaʕa* ‘there’s not... they just play [music] amongst themselves, there’s not a band or anything’ (2-16)
- (59) *baʕ ka-yʕhər ka-ykun... lli ka-yʕhər b t-taʕin, w l-xūbz, l-mūḥəm w atay...* ‘for shour, there are people who have a tagine, and bread, and tea...’ (7-4)

These structures correspond with other examples of /l-/ in apparently non-referential contexts discussed in Brustad, including:

- (60) *ma-ʕndha-ʕ l-wəld. naḍ gal lha ana xəʕʕni l-wəld* ‘she didn’t have a son. He up and told her, I need a son’ (Brustad, 2000, p. 36)
- (61) *...ka-ytbaʕu f l-ḥanut ʕəʕri* ‘they’re sold in a modern store’ (Brustad, 2000, p. 41)

The structure in (61) is particularly interesting, because it exhibits what Brustad considers an “asymmetrically definite construction” in which the noun *l-ḥanut* ‘store’ incorporates /l-/ but the modifying adjective *ʕəʕri* ‘modern’ does not, seeming to break an agreement rule common to all other major Arabic varieties. All of the non-Moroccan Arabic speakers with whom I was able to confer consider this structure ungrammatical. It nonetheless occurs in my data as well:

- (62) *ma-ka-tmši-ʕ l-ʕərusa l ʕənd n-nəggafa. ka-ydiru ɣir l-ʕərs ʕadi* ‘the bride doesn’t go to the hairdresser. They just have a normal wedding’ (2-13)
- (63) *ta-tdir l-ḥbub dyal atay f l-bərrad, w ka-txwa ʕlih l-ma tayb* ‘you put rolled tea leaves in the teapot, and you pour boiling water on it’ (6-1)



- (64) *ka-ybniū-h b l-gṣab w dakši, ka-ydāxxlu lih l-ma sxun bašh isxʷən*  
 ‘they build it with bamboo reeds and all that, and fill it with hot water  
 so it’ll warm up’ (9-17)

In addition to the above attestations, we can cite additional evidence that these structures are far from rare in MA. They are considered grammatical enough, for example, to appear in television advertisements (65) and even MA textbooks intended for foreigners (66):

- (65) *šhiwat ʕarāb kuḷa ka-tqāddām likūm: l-ḥut mṣammār* ‘Arab Cola  
 recipes presents to you: stuffed fish’ (Advertisement for Arab Cola<sup>13</sup>)
- (66) ... *w kayn l-bid māsluq wälla tayb* ‘... and there are cooked or boiled  
eggs’ (Chekayri, 2011, p. 477)

If the presence or absence of /l-/ is in any sense indicative of definiteness, then, it is the attributes expressing the quality rather than the head nouns, all of which above are indefinite but still have /l-/. This notion – that /l-/ is not exclusively associated with definite semantic statuses – is supported by the evidence we have now seen that the proposed “definite article” can actually occur alongside any and all of the in-focus tiers of our modified Givenness Hierarchy, including the indefinite ones ‘specific,’ ‘referential,’ and ‘type identifiable.’ Out of all the possible semantic contexts in which nominals may occur, the only one I can identify that appears to exclude forms with /l-/ is that of generic predicates, which we discussed earlier as a possible subcategory of ‘type identifiable’ (2.2.2). That such predicates are inherently attributional, like the above adjectives, might help explain the restriction on /l-/ (4.2.3; 5.2). For now, however, we turn to the definite range of the semantic continuum to examine the applicability of previous models there.

### 3.2.2. Absence of /l-/ in Definite Contexts

At this point we have established that, contrary to previous models of definiteness in MA, /l-/ can occur alongside any number of indefinite meanings, an observation that weakens the case for it as an exclusively “definite article.” I now raise the issue of /l-/

<sup>13</sup> <http://www.youtube.com/watch?v=kVp5Ce9LfCM>

as a marker for definite entities themselves, casting doubt on whether it truly correlates with definiteness even in the semantic range where it is expected to do so in Arabic varieties. In earlier examples (32-40) we saw that /l-/ does indeed appear to occur with a majority of nouns of the status ‘definite,’ but these examples were, again, selectively chosen. If we look at nominal form-meaning correspondences holistically, by contrast, what we find is that there are actually a number of nouns in MA that are never marked with /l-/ in any context, including those in which they are explicitly definite. These “prefixless” definites pose yet another challenge to the predominant descriptive model.

Scholars have long recognized that there are MA nouns that, to adopt Harrell’s wording, “never take the definite article [l-/]” (1962, p. 189). It has almost become a scholarly tradition among Moroccanists to include a list – typically 10-15 items – of nouns that display this morphology in their grammatical descriptions. Harrell, as a representative example, gives 11 well-known nouns that are apparently prohibited from being “modified” with /l-/ (Figure 11); no explanation is provided for why this may be the case. In addition to these common nouns, Harrell also notes that MA has a derived nominal pattern involving a prefix *ta-* and suffix *-t* (as in *tanəžžart* ‘carpentry, from *nəžžar* ‘carpenter’) that displays the same restriction. Harrell suggests that most nouns of the type that do not take /l-/ “must be memorized as individual lexical items,” contributing to the sense that they are a closed class that do not abide by the grammatical rules of other MA nominal entities.

**Harrell’s (1962, p. 190) List of Common MA Nouns Avoiding /l-/**

<i>atay</i>	‘tea’	<i>bibi</i>	‘turkey’	<i>bəllarž</i>	‘stork’
<i>bnadəm</i>	‘human being’	<i>maṭiša</i>	‘tomatoes’	<i>muka</i>	‘owl’
<i>səksu</i>	‘couscous’	<i>tamara</i>	‘trouble, pain’	<i>ṭaba</i>	‘tobacco’
<i>xizzu</i>	‘carrots’	<i>žahənnəm</i>	‘hell’		

Figure 11: Common Nouns that Avoid /l-/ (Harrell, 1962)

Moscoso García (2003) gives a similar list of items for which /l-/ cannot appear, dividing them into three types as a function of phonology and apparent origin. The first of these types includes “Berber or Berberized words that begin with *a-* (such as *atay* ‘tea’);” the second includes items of various origins – whether Berber, European languages, or even Arabic – that “do not begin with *a-*” (as in *ṭaba* ‘tobacco’); the third includes words that are prefixed with the etymologically Arabic *bən-* or *bu* (like *bnadm* ‘human being’). It is not necessarily clear what value there is to this grouping when nothing about the phonology of the second “type,” in particular, would distinguish its constituents from the majority of other Moroccan nouns; it in essence appears to be a sort of “catch-all” category for those /l/-less nouns that Moscoso García is unable to neatly categorize. Nonetheless, the simple act of creating such a list again gives the feeling that these “inherently definite” nouns are necessarily restricted in number and exceptions to the grammatical rule, corroborating Harrell’s impression.

The analysis of such nouns as “exceptional,” which is probably the predominate one, is convenient because it allows scholars to skirt the primary challenge they pose: that they do not fit the commonly described model of how definiteness in MA functions. If /l-/ were in fact a definite article, it would be expected that it could mark definiteness for all common nouns. These “prefix-resistant” nouns are thus problematic because their formal behavior cannot be accurately predicted using the standard set of morphosyntactic operations that (at least partially) work for most MA nouns. Figure 12 elucidates this

...	>	definite	>	specific	>	referential	>	type identifiable
		/l-/ N		<i>waḥəd</i> (/l-/) N		<i>ši</i> N		Ø N
		<i>l-wəld</i>		<i>waḥəd l-wəld</i>		<i>ši wəld</i>		<i>wəld</i>
		‘the boy’		‘this boy’		‘some boy’		‘(a) boy’
		<i>*l-bəllarž</i>		<i>*waḥəd l-bəllarž</i>		<i>ši bəllarž</i>		<i>bəllarž</i>
		‘the stork’		‘this stork’		‘some stork’		‘(a) stork’

Figure 12: Ungrammatical Results for Proposed Forms

challenge, showing how this previously described model produces grammatical results for a hypothetical unmarked form *wəld* ‘boy,’ but not *bəllarž* ‘stork.’

Moving forward, we can corroborate this theoretical problem with evidence from natural speech. Two very similar occurrences of *atay* ‘the tea’ are in my data, the first of which below can be identified as definite via the modifying adjective *l-luwl* ‘original,’ and the second of which displays the same adjective as well as a demonstrative *dak* ‘that.’

(67) *ka-nkəbb<sup>w</sup>u l-ħaža, zʕma ka-nħiydu-h, ka-nšəlləlu-h, ka-nəržž<sup>u</sup> atay l-luwl l l-bərrad* ‘we pour out a bit -- I mean we remove it -- we rinse it, then we put the original tea back in the teapot’ (5-6)

(68) *ka-tšəff<sup>i</sup> f l-kas dak atay l-ləwwəl, w tʕawd tšəlləl... ʕawd tšəlləl žuž xəṭrat* ‘you strain that original tea into a cup, and keep rinsing... you rinse it two times’ (6-3)

In neither case here, however, is the noun *atay* marked with /l-/ despite its contextually definite status, a reality that is in direct conflict with the model previous descriptions of MA definiteness marking have put forward. Furthermore, although some scholars (such as Moscoso García, above) have highlighted the etymological similarity of such nouns with those in Berber (where definiteness is unmarked), the fact that the completely monolingual Arabic speaker in (67) produced the same structure as the Berber-MA bilingual in (66) reiterates that such “inherently definite” nouns are fully a feature of MA and cannot, at least synchronically, be dismissed as an outcome of second-language interference.

If these items do represent a closed class, of course, they might be rightly dismissed as the frozen products of distant linguistic history, a small group of exceptions with no real implications for the syntax of the language today (similar to *-en* plurals in Modern English). This indeed seems to be the view that most previous studies have taken. My stance, however, is that this view disregards a good bit of evidence that “prefixless” definites are not only much more common than is typically allowed for, but also that they have productively entered the language in the recent past and may even still

be doing so. This serves as an argument against the notion that we are justified in excepting them from a model of definiteness marking in MA.

To begin, although most grammars tend to give a small list of such nouns, a closer look at the MA lexicon shows that one can rapidly expand their count. Harrell's own (1966) dictionary, for example, includes perhaps over a hundred more items that are described as having "no article." Among the easiest of these to identify are apparent loans that have maintained the Berber state markers *a-* and *ta-*, including nouns such as *azaglu* 'yoke,' *afrag* 'partition,' *abraz* '[type of] wedding ceremony,' and *takawt* 'gall.' Chafik's (1999) lexicon of Berber loans into MA registers yet hundreds more lexical items, many 'substrate-type' nouns that relate to flora, fauna, and local cultural traditions, that display the same morphology. In addition, we must keep in mind that MA has fully borrowed the Berber derivational pattern *ta-* + [nominal stem] + *-t*, which likewise precludes /l-/, meaning /l-/-less definite abstract nouns can be continually be added into the MA lexicon at will.

More /l-/-resistant common nouns can be found in Harrell's dictionary under the prefixes *bən-* and *bu-*, among them *bənnəfzuž* 'violet,' *bubriš* 'gecko,' *busəkka* 'rattlesnake,' *butəllis* 'nightmare,' *bušwida* 'pear,' and *bušwika* 'scarlet fever.' Diachronically, we can identify most of these compounds as derived from the Arabic elements *bən* 'son' or *bu* 'father' plus another noun, as in *bu* 'father' + 'twig' > *bušwida* 'pear.' These compounds may have originally served as generic epithets that applied to the entire class and for which referentiality was unimportant, but at some point a number of them must have begun to refer to nouns that can be individuated (as in *bušwida* 'the pear,' *waḥəd bušwida* 'this [particular] pear'). That this semantic transition happened without the morphological addition of /l-/ in definite contexts suggests that some sort of semantic framework that allows for such constructions may have already been in place.

One may argue that these etymologically Berber and restructured Arabic nouns could have entered into MA very early and simply be the remnants of some sort of distant semantic reshuffling that is no longer descriptive of today's paradigm, but even this reasoning would warrant an explanation as to why, after so much time, they have never

succumbed to the pressures of linguistic economy and began to be marked with /l-/ like other nouns (if it were indicative of definiteness, that is). We can likewise strengthen the argument against this view with the observation that /l-/less definite nouns can be shown to have entered via borrowing from European languages, with which MA was only in much later contact. Heath (1989, p. 157), for example, cites data from Brunot (1949) to claim that “there is evidence that at some chronological stages in the recent develop of [MA], there has been a general tendency to avoid adding definite /l-/ to borrowed nouns from [European] languages.” Although Heath clarifies that such nouns are rarer today, the fact that MA was productively adding nouns that disallowed a “definite” /l-/ as recently as the French colonial period is further evidence that we are speaking of an open class rather than a closed one.

In fact, a few such nominals even appear in my data. Although many MA dialects render the borrowed nouns underlined below as *l-bṣtila* ‘pastilla’ (< Sp.) and *d-disir* ‘fruit’ (< Fr. *dessert*), for example, the speaker in (68) gives them without /l-/ , which we might otherwise expect since it is present with all the other generic food items she lists:

- (69) *d-džəž, l-lhəm b l-bərquq, l-həlwa, bṣtila, disir, kǔlši dakši...*  
*ka-yak<sup>w</sup>əlu n-nas iṣərbu* ‘chicken, meat with prunes, sweets, pastilla,  
fruit, all of that... people eat and drink’ (2-5)

In yet another example from my data a borrowed noun *kayit* ‘[course] paper’ (< Turkish *kağıt* ‘paper’) shows the same pattern, resisting /l-/ even in contexts where it has been described as necessary. MA has never been in intense contact with Turkish (which has no definite article), so the term in question may well have been brought into the country via Ottoman Algeria, but in any case it shows the same morphological pattern of avoiding /l-/:

- (70) *mn bṣəd, ka-tqfəl ṣəla hadik l-ṛina? dyal t-ṭin, ka-tqfəl ṣəlih b ši... b ši*  
*kayit, kayit dyal qalb s-skk<sup>w</sup>ar, wahəd kayit zṛəq...* ‘Afterward you  
close off that clay pot, you close it off with some sort of course paper,  
the course paper from a block of sugar, this blue course paper...’  
(4-12)

For both (69) and (70), the speakers grew up as MA monolinguals and there is no reason to doubt their competency in the dialect. I also have in my field notes a similar construction *dak amkawsu* ‘that scythe’ – rather than the *dak l-amkawsu* previous models might predict – produced by a native MA speaker from Zagora with no knowledge of Berber. That all of these speakers readily produce forms without /l-/ in contexts that previous descriptions would lead us to believe require them, then, is a serious challenge to the adequacy of those grammatical models. Furthermore, we have seen that the primary defense one could make in those models’ favor – that “inherently definite” nouns represent a closed class of exceptional nominals with their own grammar – is hard to maintain in the light of evidence that this “class” can and has historically been expanded via processes of borrowing and morphological derivation.

At this point the final defense one might make for the traditional model of definiteness in MA is to claim that it is the dominant one of two separate systems for marking definiteness, and that the “inherently definite” nouns we have discussed here are subject to a different but parallel set of morphosyntactic processes. I do not find this notion appealing, and contend that that burden of proof is upon whomever does to explain what sort of cognitive processes would prompt native MA speakers to maintain two different formal paradigms for expressing the same set of semantic notions. Phonological cues might seem like a tempting explanation, with nouns that begin with *a-* almost universally excluding /l-/, but this suggestion does nothing to explain how a speaker would know to derive definite *t-ṭaqa* ‘the window’ from *ṭaqa* but not *\*t-ṭaba* ‘the tobacco’ from *ṭaba*, where the initial segments are identical. As I take a different view entirely, I leave such speculation to others.

We have now seen that previous descriptions of /l-/ as a “definite article” fail to accurately describe actual MA speech when viewed from two primary angles. On one hand, /l-/ has been found to occur throughout the entire range of the Givenness Hierarchy, meaning it can occur alongside semantically indefinite nouns of all statuses. On the other, there is a great degree of evidence that definite semantic statuses do not actually require /l-/ in order to be expressed. I take all of the above as the basis for my central theoretical

claim: that /l-/ in Moroccan Arabic, for whatever it may have been in the past, is no longer a definite article. In keeping with this claim, I hereafter refer to it as \*/l-/, a reminder that we are not looking at the traditional Arabic article but rather a reanalyzed form with quite different semantic implications.



## 4. A New Model for Definiteness in MA

If \*/l-/ is not in fact a definite article, the question that logically follows is what it instead represents. My answer to this is relatively simple: \*/l-/ is no longer an article at all, but rather a lexicalized component of what is now the unmarked form of most etymologically Arabic nouns, as well as of those borrowings that have (optionally) been given it by analogy. Discussion of how such a reanalysis may have occurred diachronically follows in 4.2.1, but I begin with a focus on how this description suits the present-day behavior of MA nouns. My claim is that by re-envisioning \*/l-/ not as a definite article but rather as a lexical component, we can simultaneously explain both why it appears alongside nouns in indefinite contexts as well as why some nouns can exhibit definiteness in lieu of it.

### 4.1. SYNCHRONIC DESCRIPTION

In the new model I hereby propose for definiteness marking in MA, the semantic status ‘definite’ is exclusively zero-marked, as can sometimes be the statuses ‘specific,’ ‘referential,’ and ‘type identifiable’ as well. For most MA nouns of Arabic origin, the ‘bare’ or ‘unmarked’ form is that which expresses the historical article \*/l-/ (as in *l-wəld* ‘boy’), which explains why \*/l-/ gives a superficial sense of representing definiteness. Other nouns have unmarked states that do not express \*/l-/ (such as *bəllarž* ‘stork’), which is likewise a function of their etymology rather than any exceptional semantic status. For each noun, this unmarked form – which we can imagine as corresponding with the speaker’s most basic mental representation of a given nominal entity – is solely sufficient for predicting all of the possible nominal forms that correspond with various tiers of the Givenness Hierarchy.

To clarify, I mean by this that the form *l-wəld* expresses as its most basic sense the concept ‘boy,’ without necessarily implying anything about the noun’s givenness status. In the same way, *bəllarž* simply expresses ‘stork’ without additional semantic information. Once the concept expressed via the bare noun is present mentally, a speaker may proceed to make additional morphosyntactic modifications to the nominal form in

order to bring it into line with a particular semantic status. For truly definite entities, the semantic status is indicated by zero-marking, so that the definite form remains phonologically identical to the bare form (i.e. *l-wəld* ‘the boy;’ *bəllarž* ‘the stork’). For specific entities, a speaker can prefix the bare form with the article *waħəd* to indicate that particular semantic notion (as in *waħəd l-wəld* ‘[indef.] this boy,’ *waħəd bəllarž*, ‘[indef.] this stork’). Since \*/l-/ here is considered a lexical component rather than an article and has no actual semantic implications for definiteness, this model simultaneously answers the question of why the article *waħəd* would seem to require it for some nouns but not others, as well as how \*/l-/ can be found alongside an indefinite an indefinite status.

To claim that the unmarked state for most MA nouns of Arabic origin is that which contains \*/l-/ (as in *l-wəld* ‘boy’), of course, runs directly counter to previous literature’s insistence that the “prefixless” MA form (*wəld*) represents the bare or unmarked nominal. I give more detailed justification for this view in the following discussion, but for now it is worthwhile to answer the question of how one can explain the continued attestation of nominals for which \*/l-/ is possible but not present. My view is that this ‘historical indefinite’ – which I will hereafter give as \*Ø – is, like \*/l-/, no longer associated with any particular givenness status, but has rather undergone reanalysis into what is primarily a marker of quantification, with relatively strict syntactic restrictions. I also hold that it is no longer best envisioned as ‘bare’ form, but rather one that is productively derived from the synchronic unmarked form by *disfixing* the etymological element /l-/ where it occurs, meaning we are looking at a subtractive morphological process rather than an additive one (*l-wəld* ‘boy’ - /l-/ > *wəld* ‘a [single] boy’). I represent this ‘quantifying disfix’ as {- /l-/}. My data supports the notion that where the disfix {- /l-/} occurs independently, it indicates absolute quantity; where it occurs alongside the article *ši*, it indicates unspecified quantity, or simply discursive ‘existence’ (which, in turn, accounts for *ši*’s association with referentiality).

This model has a number of advantages, not least among them the fact that we can for the first time accurately predict formal expressions of givenness in all MA nouns, including even those that have previously been treated as grammatical exceptions, by

applying a single unified set of morphosyntactic operations to the unmarked form for any nominal entity. Figure 13 shows how the model produces fully grammatical forms for unmarked MA nouns regardless of their etymology and whether or not the bare form contains \*/l-/. Since we can assume that disfixation of \*/l-/ is phonologically conditioned, the process is productive for those nouns for which it is part of the bare form (*l-wəld*, *l-kas*) but is not necessary or possible for those for which it is not (*bəllarž*, *maṭiša*). Each of these resulting forms is associated – though in the case of the zero-marked form, not exclusively – with a certain semantic sense.

Ø N		<i>wahəd</i> (Ø) N		<i>ši</i> {- /l-/} N		{- /l-/} N
unmarked (definite)	>	specific	>	referential	...	(quantified) type identifiable
<i>l-wəld</i> '(the) boy'		<i>wahəd l-wəld</i> 'this boy'		<i>ši wəld</i> 'some boy'		<i>wəld</i> 'a (single) boy'
<i>bəllarž</i> '(the) stork'		<i>wahəd bəllarž</i> 'this stork'		<i>ši bəllarž</i> 'some stork'		<i>bəllarž</i> 'a (single) stork'
<i>l-kas</i> '(the) cup'		<i>wahəd l-kas</i> 'this cup'		<i>ši kas</i> 'some boy'		<i>kas</i> 'a (single) cup'
<i>maṭiša</i> '(the) tomato'		<i>wahəd maṭiša</i> 'this tomato'		<i>ši maṭiša</i> 'some tomato'		<i>maṭiša</i> 'a (single) tomato'

Figure 13: Reanalyzed MA Forms & Givenness Implications

Figure 14 gives the currently proposed system of definiteness in MA using Givón's wheel diagram (described in 2.2.2). The four nominals forms given in Figure 13, as we can see, are distributed with some overlap over Givón's six semantic sectors. Using this diagram, we can elucidate some of the generalizations I herein make about givenness in MA. First, it is the unmarked form Ø (typically corresponding with the presence of \*/l-/) that has the widest semantic range, able to express the notions represented by five of the six sectors. It likewise has exclusive domain over the 'generic-subject' and

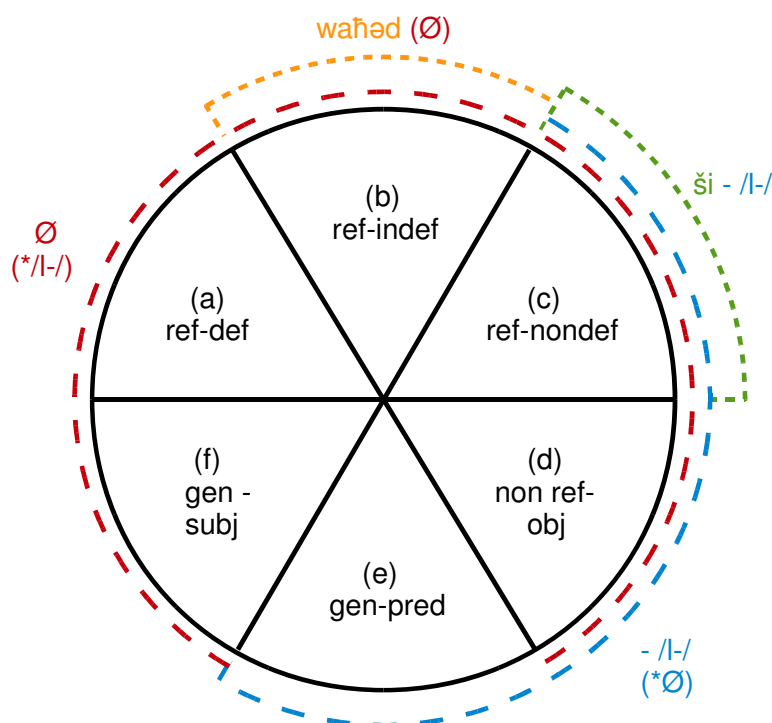


Figure 14: Wheel Diagram for Moroccan Arabic

Equivalencies for  
(Modified)  
Givenness Hierarchy:

- (a) Definite
- (b) Specific
- (c) Referential
- (d) Type Identifiable
- (e) Type Identifiable\*
- (f) Definite\*

*In-Context Examples of Extant Forms:*

- (a) 1. *l-ṣarusa ka-thətt gdamha wəšt l-gəsriya lli fiha l-ma w fuq d-dəbliyž* ‘the bride puts her foot in the middle of the bowl of water and on top of the bangle’ (9-39)
- (b) 1. *ta-ygəlsu f waḥəd l-kūrsi mūṣəyyən f waḥəd s-sala kbira* ‘they sit in this special chair in this big hall’ (9-8)  
2. *kaynin l-ṣayalat daba lli dayrin žəmfiyat w tṣawniyat nisaḍiya* ‘there are these ladies nowadays who have put together women’s associations and cooperatives’ (1-18)
- (c) 1. *ka-txəlləš-u, ka-təṣṣi-h ši baraka, imma ṣəšra drahəm wəlla xəmṣəšəṣ dərḥəm wəlla miyət dərḥəm, šḥal ma kan* ‘you pay him, give him some sort of tip, either ten dirhams or fifteen dirhams or a hundred dirhams, whatever you’ve got’ (4-16)  
2. *daba l-wəld ila bya itžəwwəž l-bənt ta-ymši... l l-walidin dyalu ta-yxṭṭəbu-h* ‘now if a boy wants to marry some girl he goes... to his parents and they get him engaged’ (2-1)
- (d) 1. *ka-tṣəffi f l-kas dak atay l-ləwwəl, w tṣawd tšəlləl... ṣawd tšəlləl žuž xəṭṭrat* ‘you strain that original tea into a cup, and keep rinsing... you rinse it two times’ (6-3)  
2. *w ka-ndiru kas dyal l-ma, ṣafi* ‘and we put in a [single] cup of water, that’s all’ (4-8)
- (e) 1. *n-nhar l-axwər yadi tnuḍ, ka-tšqa f ḍarḥūm ṣəla asas annəha ṣafi raha wəllat mərə ka-tṣətamd təmmaya* ‘the next day she’ll get up and do housework, with the logic that she’s now become a woman and is being depended upon there’ (9-43)
- (f) 1. *l-mərḥəla l-ḍəwla hiya l-mərḥəla dyal d-dəzzan yəṣni ka-ydəzz... ka-ndəzzu l-ynəm, w had l-mərḥəla ka-yqum biha r-ražəl* ‘the first phase is that of gathering the wool: we shear the sheep, and this phase is done by a man’ (1-2)

‘referential-definite’ sectors, which we can effectively group together as simply ‘definite.’ In addition, another form *waḥad* is “built on” the form Ø. The disfixed form {- /l-/}, on the other hand, not only is distributed throughout a smaller semantic range but is the exclusive form for only one sector (‘generic-predicate’); I see this co-distribution as the outcome of syntactic constructions that require {- /l-/} where Ø would otherwise suffice semantically. Among these constructions is *ši*, which can likewise be said to be “built on” an underlying disfixive process {- /l-/} and is associated with the status ‘referential.’

In what follows I give a brief synchronic description for each of these four nominal forms, respectively marked Ø, *waḥad* (Ø), {- /l-/}, and *ši* {- /l-/}. In particular, I seek to show how they support a model that solves most of the theoretical challenges one can raise for past descriptions, as laid out in 3.2. We begin with the most widely-distributed and basic of these forms, the unmarked or bare form Ø, which for most nouns of Arabic origin corresponds with forms containing the historical article \*/l-/.

#### **4.1.1. Unmarked Form Ø (etymological \*/l-/)**

The unmarked nominal form in MA can be best described as that which a noun takes when explicitly referential and definite, i.e. when the speaker is using it to refer to an individuated entity that is uniquely identifiable for both the speaker and addressee. This is not to say that the unmarked form cannot occur alongside other semantic statuses, because – as we saw in 3.6.1 – it certainly can; rather it is the fact that definite entities are exclusively expressed via the unmodified, zero-marked form that allows us to identify it in this way. There are no particular phonological requirements for the unmarked form, although as a function of their Old Arabic etymology the lexicalized prefix \*/l-/ appears as part of it for a majority of nouns (as in *l-kas* ‘cup,’ *l-wald* ‘boy,’ *š-šəžra* ‘tree,’ *t-tbən* ‘straw’). The phonological shape for nouns of non-Arabic origin varies and can be seen as an outcome of borrowing routines that are susceptible to historical change; for this reason, some have entered as their bare forms in the source languages (atay ‘tea,’ *bəllarž* ‘stork,’ *ṭaba* ‘tobacco’) while others have been given the lexical component \*/l-/ out of

analogy with the majority of MA forms (*t-ṭumubil* ‘automobile,’ *l-buḷa* ‘lightbulb,’ *r-rwiḍa* ‘wheel’).

In claiming the MA form that is used as the exclusive means for expressing definite entities to be the unmarked one (i.e. \*/l-/ rather than the historical indefinite \*Ø for most etymologically Arabic nouns), I base my judgement on Croft’s (1990) definition of markedness as given in Winford (2003, p. 230), who states that “in general, the unmarked value of an opposition [1] has less complex structure, [2] is found in more environments, [3] and has a wider-cross-linguistic distribution.” One can make an argument that the semantically definite form satisfies each of these diagnostic criteria, which I present here.

Perhaps the easiest of these criteria to comment on is [2], because we have already seen a great deal of evidence that the forms through which definiteness is expressed are found in more semantic environments than others. This is not necessarily a new discovery, but rather a refinement of the impressions that other authors have already given. Corriente, (2008, p. lxvi), for example, notes that the use of what he considers to be the “definite article” \*/l-/ does not have “a distribution strictly governed by the category of determination,” and that it is on the whole used much more frequently than forms not preceded by \*/l-/. In the current study we have looked at this distribution in more depth, and seen that the \*/l-/ forms, as shown in Figure 14, can occur for meanings of the statuses ‘definite’ (including referential entities and generic subjects), ‘specific,’ ‘referential,’ and ‘type identifiable’ (3.2.1). This wide semantic distribution serves as a preliminary argument in favor of \*/l-/ forms being unmarked.

Answering criterion [1] is more involved, because ‘complexity’ can be defined in different ways. If we were to take phonological bulk as the primary measure of complexity, for example, it would indeed appear that what am I claiming is the unmarked form of an etymologically Arabic noun (as in *l-wəld*) is in fact more complex than the disfixed form (*wəld*), simply because it has comparatively more phonological material. Croft’s (1990) notion of structural complexity, however, is not primarily interested in phonology but rather morphology, meaning that a less complex structure should be the

result of fewer morphological operations. In this light what I see as an ideal metric for a more ‘simplex’ nominal form is whether it alone is sufficient for deriving all other possible forms that are needed to express the full range of semantic possibilities. To put my claim in another way, if we are given only the unmarked (or simplest) nominal form and have no knowledge of the other possible forms, we should be able to predict the latter with full accuracy using the former. The opposite, however, may not be true: more complex forms should not necessarily be independently sufficient for back-forming more basic forms.

As a proof of how this complexity metric qualifies the ‘definite’ series of forms for an unmarked status but disqualifies those that have traditionally been considered ‘bare’ (and which I consider to be ‘disfixed’ instead), I give the following sets of nouns (Figure 15). Set (1) gives three groups of two nouns with the forms they will necessarily take when a speaker means to refer to a semantically definite entity. Set (2) gives what previous descriptions have taken as the unmarked forms of the same nouns (3.1.1), which are essentially “prefixless” in that they have no lexical component \*/l-/.

	(a)		(b)		(c)	
(1)	<i>l-bulisi</i>	<i>buʃwida</i>	<i>t-ṭaqa</i>	<i>ṭaba</i>	<i>ž-žbəl</i>	<i>l-žasad</i>
(2)	<i>bulisi</i>	<i>buʃwida</i>	<i>ṭaqa</i>	<i>ṭaba</i>	<i>žbəl</i>	<i>žasad</i>
	‘policeman’	‘pear’	‘window’	‘tobacco’	‘mountain’	‘body’

Figure 15: Contested Unmarked Forms

The notion advanced in previous descriptions is that the nouns in set (2) are the bare or unmarked forms from which the “definite” nouns in set (1) can be derived by adding the morphological element \*/l-/. As we saw in section 3.2.2, this fails to accurately predict the morphological behavior of a number of nouns which are never prefixed with \*/l-/ even when definite, as seen with *buʃwida* and *ṭaba* above in set (1) above. In groups (a) and (b) I have purposefully juxtaposed these nouns with others containing phonologically identical initial segments to show that there is no phonological

cue that would allow speakers to make this derivational distinction if they were indeed “building” the forms on set (1) from the forms in set (2). The same is true of group (c), which presents a phenomenon noted in Heath (1989, p. 53) where \*/l-/ in the vicinity of /ž/ is realized as gemination for native MA nouns (ž-ž**ba**l) but as a phoneme /l/ for CA borrowings (l-ž**asad**; also see 4.1.3). Heath suggests for such nouns an underlying formal distinction between /ž<sub>1</sub>/ and /ž<sub>2</sub>/ that determines the realization of \*/l-/, but this provides no explanation as to how native speakers – who do not necessarily have knowledge of the nouns’ provenance – would cognitively differentiate two identical phonemes, so it too is problematic for the notion that definite forms can be adequately predicted from the traditionally given bare forms.

If we take the forms in set (1) as the most basic representations of the noun, however, the opposite is true: we can, via a single set of morphosyntactic rules, accurately predict all of the forms in set (2). This productive process is the ‘disfixation’ I refer to elsewhere, and is in essence the inverse of affixing nouns in other Arabic dialects with an article /(V)l-/. The rule governing this disfixation is that *if a nominal form either begins with a phoneme /l/ or a geminate consonant, either the /l/ will be dropped or the consonant will be degeminated, respectively; if neither is present then the noun undergoes no morphological change*. Additional evidence for the productivity of this process and its theoretical plausibility is given in section 4.1.3, but for the current argument this description is sufficient to show that it is possible to predict all other possible nominal forms from only that one which overlaps with definiteness, whereas for other scenarios this is not the case. This observation strengthens the argument that those forms which express the ‘definite’ semantic status are in fact formally co-identifiable with the most basic or simple nominal forms, which in turn satisfies criterion [1] for considering them unmarked.

Finally, we can consider the notion of cross-linguistic distribution (metric [3] in Winford’s quote, above) as relevant to the case for nominals with \*/l-/ in MA being unmarked. If \*/l-/ is to be interpreted as an etymological component of the nominal stem rather than an article, and if it can, as we have seen, appear alongside the majority of



givenness statuses, the implication is that most of these semantic categories in MA can be expressed without any article at all. This would actually make MA more typologically similar to the majority of the world's languages, only about a third of which have been shown to have articles of any sort (Dryer, 1989 in De Mulder & Carlier, 2011). In addition, if we look at the immediate areal distributions of such features, the notion that there is no true “definite article” in MA would align it typologically with nearby Berber languages, in which definiteness is also zero-marked. Later I make an argument that this areal arrangement has likely played a role in the development of the current system, but for now it suffices to reiterate that zero-marked definiteness is in fact typologically common both worldwide and locally, and that this again works in favor of my interpretation of the definite MA noun as structurally identical to the bare form.

Following this logic, I take the nominal form that can be used to explicitly express definiteness as the unmarked form, even if it occurs in other semantic contexts (the important trait being that it could be used to refer a definite noun without formal modification). In the below example, for instance, all of the underlined nouns are contextually definite, having been introduced in previous discourse. Since we can consider the definite form identical to the unmarked form, we can thus assume the unmarked forms *l-ḡarusa* ‘bride,’ *l-ḡaṣriya* ‘bowl,’ and *d-dābliyiž* ‘bangle:’

- (71) *l-ḡarusa ka-thətt ḡdāmha wəšt l-ḡaṣriya lli fiha l-ma w fuq d-dābliyiž*  
‘the bride puts her foot in the middle of the bowl of water and on top  
of the bangle’ (9-39)

That all of these unmarked nouns contain \*/l-/ is, again, simply a function of their etymology. Other MA nouns will not display \*/l-/, but they are no less definite than the nouns in (70). In the following example, for instance, we can cite the unmarked forms *l-kəbš* ‘ram’ and *bəllarž* ‘stork,’ where one bare form has etymological \*/l-/ but the other does not; both, however, refer to nominal entities that have already been established in the discourse:

- (72) *l-qəṣṣa dyał waḥəd l-kəbš, huwa w waḥəd bəllarž. huma šhab. waḥəd*  
*n-nhar ḡərəd l-kəbš ḡla bəllarž...* ‘[here is] the story of this ram, him

and this stork. They're friends. One day the ram invited the stork...' (Destaing, 1937, p. 89)

Although the form Ø itself does not necessarily imply anything about definiteness, it often expresses such meanings simply because the 'definite' status is exclusively zero-marked. The unmarked form itself, however, is not limited to only the 'definite' status and can appear alongside indefinite meanings as well, a possibility that is allowed for specifically because it lacks an association it with any particular semantic status. Section 3.2.1 gives a number of occurrences of nouns containing \*/l-/ – which we can now identify not as an article but rather as a lexicalized component of the bare noun – in indefinite contexts. These, then, are evidence of how the unmarked form can be used in formal representations of nearly any givenness status. There is no need to repeat all of the previous examples here, but we can nonetheless briefly review the contexts that they exemplify.

First, as we saw, the unmarked form is the primary means for expressing *unquantified*, non-referential 'type identifiable' objects, as seen with the nouns *š-šərbil* 'pair of slippers,' *s-səlham* 'cloak,' and *t-tur* 'bull' below:

(73) *yəʕni ida žab liha š-šərbil, ta-yžib l mm<sup>w</sup>u š-šərbil ta-yžib l mm<sup>w</sup>ha š-šərbil* 'meaning if he brings her a pair of slippers, he brings his mother a pair of slippers and brings her mother a pair of slippers'

(74) *l-yum huwa labəs s-səlham* 'today he's wearing a cloak' (Harrell, 1962, p. 190)

(75) *dbəh t-tur, ʕrəð ʕla n-nas* 'he slaughtered a bull, invited people' (Brustad, 2000, p. 37)

Second, the unmarked form alternates with the more dominant form *ši* {- /l-/} (see 4.1.4) as an expressive means for 'referential' entities, as in *s-suq* 'some [sort of] market' in example (76):

(76) *w ka-ntmnau zəʕma y... ikun s-suq hnaya, hna f taznaxt, maši f mərrakš, ikun t-tabɿ dyal z-zərbiya hnaya f taznaxt* 'we wish we had some sort of market [for them] here in Taznakht, not in Marrakesh,

and that the impact of the carpet [industry] was here in Taznakht'  
(1-21)

Finally, the unmarked form can appear with entities of the status 'specific.' For plurals, it can occur independently, as it does with *n-ntaʔiž* 'these results' in the following sentence:

- (77) *l-...ntaʔiž ka-ybiynu fihūm lli nžəhu, w kaynin n-ntaʔiž xwəra ka-ybiynu l-... lli ma-žabu-š n-nūqta mzyan ʕəndhūm l-ʔistidraki* 'the results make it clear who passed, and there are these other results that show the... those who didn't get a good score have the remedial exam' (8-25)

Alternatively, the unmarked form can be prefixed with the article *waḥəd* for what is typically (but not necessarily) a specific meaning. We now look at synchronic uses of this article specifically.

#### 4.1.2. Article *waḥəd* Ø

The MA article *waḥəd* is a modification of the unmarked nominal form, a prefix added to the bare form Ø that implies an indefinite meaning, generally of the status 'specific' (though this may not be true for all usages; see below). In terms of the article's form and use, I differentiate my account from that of previous descriptions (which gave an article *waḥəd* /l-/) in that I do not see the presence of \*/l-/ to be related or relevant to the article at all. The tendency of \*/l-/ to appear in the vicinity of *waḥəd*, I again claim, has nothing to do with the nature of the article but is simply a reflection of most MA's nouns' etymology, where historical \*/l-/ is retained in the unmarked form. The following opening line from one of Destaing's (1937) MA folk stories, for example, shows how *waḥəd* is identically prefixed for the unmarked nouns *l-kəbš* 'ram' and *bəllarž* 'stork' to give the indefinite-specific meanings 'this ram' and 'this stork,' respectively:

- (78) *l-qəšša dyal waḥəd l-kəbš, huwa w waḥəd bəllarž* '[here's] the story of this ram, him and this stork' (Destaing, 1937, p. 89)

In the same way, this explanation accounts for the indefinite constructions *wahəd rəddəfət l-bqər* ‘this lizard,’ *wahəd tata* ‘this chameleon,’ and *wahəd kayit* ‘this course paper’ in examples (78) and (79). In each of these cases *wahəd* is simply prefixed to the unmarked forms *rəddəfət l-bqər* ‘lizard,’ *tata* ‘chameleon,’ and *kayit* ‘course paper.’

(79) *l-qəssa dyal wahəd rəddəfət l-bqər hiya w wahəd tata...* ‘[here’s] the story of this lizard, her and this chameleon...’ (Destaing, 1937, p. 49)

(80) *mn bʕəd, ka-tqfəl ʕəla hadik l-ʔina? dyal t-ʔin, ka-tqfəl ʕəlih b ši... b ši kayit, kayit dyal qalb s-skk<sup>war</sup>, wahəd kayit zraq...* ‘afterward you close off that clay pot, you close it off with some sort of course paper, the course paper from a block of sugar, this blue course paper...’ (4-12)

The noun *rəddəfət l-bqər* ‘lizard’ is etymologically a compound of two Arabic nouns (*rəddəfa* ‘sucker’ + *l-bqər* ‘cow’), but it behaves the same as other nominal entities in that it can be made indefinite simply by prefixing it with *wahəd*. Other grammatical constructions that display this compound morphology include *wahəd bit l-ma* ‘this [particular] bathroom’ and *wahəd ɖar š-šabab* ‘this [particular] youth center,’ all of which again confirm that the presence of noun-initial \*/l-/ is not required by *wahəd*.

As we saw earlier in our discussion, previous descriptions have emphasized, whether explicitly or implicitly, the role of *wahəd* as a marker for ‘specific’ givenness statuses. In this sense, it is used to refer to an entity that the speaker can uniquely identify, but that the addressee cannot; this description would also account for its attested role as a ‘new topic’ marker. For the most part, I am in accord with previous authors on the semantic implications of *wahəd*, which indeed appears to indicate a ‘specific’ entity in the majority of cases where it occurs.

In (78) and (79), for example, it is clear that Destaing’s storyteller is introducing the main characters in a tale of which he is already aware, meaning they are known to him but not to the listener. Similarly, the speaker in (80) uses *wahəd* to clarify that he is not (after the first mention, at least) referring to just any sort of course paper, but very specifically to a blue-colored type that sugar comes wrapped in, which would not be

immediately clear to the listener. These are typical usages of *waḥād*, where we can expect that some sort of elaboration will be given for the modified noun, either by means attribution or adopting it as a subject. The same expectation is likewise met in all of the following occurrences of *waḥād*:

- (81) *ka-nqəbṭu... waḥād l-ṭina? dyal... t-ṭin, w ka-ndiru fih... ka-ndiru fih l-lḥəm* ‘we grab this [certain] pot made out of clay and put meat in it’ (4-4)
- (82) *ta-ygəlsu f waḥād l-kūrsi mūṣəyyən f waḥād s-sala kbira* ‘they sit in this special chair in this big hall’ (9-8)
- (83) *ya-nqəddəm waḥād l-wəsfə dyal... wəsfə myaribiya, ṣibara ṣla ḥərša* ‘I’m going to present this recipe, a North African recipe, known as *harsha*’ (3-1)
- (84) *ka-ykun bzzaf, xaşşətan waḥd n-nuṣ dyal l-ḥərira ka-tkun, küll waḥd ka-ydir l-ḥərira f l-ftur* ‘there’s a lot [of food], especially this [certain] type of *harira*; everyone makes *harira* for iftar’ (7-12)
- (85) *ṣəḥqas kayna waḥād l-m<sup>w</sup>adda x<sup>w</sup>əra ka-nduwz-ha f l-ṣəṣiya, hiya l-anglay* ‘because there’s this other subject we’ll take in the afternoon, which is English’ (8-14)
- (86) *ila šədditi l-ḥakaḷurya rah ka-tkun fərḥan w l-ṣaṭila dyalk ka-tfərḥ mṣak, w ka-ykun... ka-ykun waḥād s-sṣada kbira wəṣt l-ṣaṭila* ‘if you get the baccalaureate you’re really happy and your family’s happy with you, and there’s this huge happiness in the family’ (8-37)
- (87) *b n-nəṣba l l-ṣərusa ka-tdir waḥād t-təqlid smitu taḥəmmamt* ‘as for the bride, she does this tradition called *tahəmmamt*’ (9-15)

These examples confirm that, at least in most cases, *waḥād* denotes a ‘specific’ usage. At the same time, there are examples in my data where *waḥād* is used to refer to an entity that does not appear to be uniquely identifiable in context, and thus cannot be specific. These are given below, where I translate *waḥād l-biḍa* as ‘an egg’ and *waḥād l-qənba* as ‘a piece of burlap:’

(88) *ʕanna nəʃs kilu dyal l-ħərša, ʕanna šwiya dyal l-məlħa, w waħəd l-bida, w kas dyal l-ħlib* ‘we have a half kilogram of harsha [mix], a little bit of salt, an egg, and one cup of milk’ (3-2)

(89) *ka-tšədd-u b waħəd l-qənba... w ka-tddih l l-fərnaṭši* ‘you grab it with a piece of burlap... and you take it to the furnace operator’ (9-15)

I do not, of course, have a direct window into the speakers’ intentions and cannot say with absolute certainty that were not envisioning unique entities, but the discursive contexts here – in which neither entity is elaborated upon in any way, and for which specificity would seem to have no effect on the discourse – make this possibility unlikely. Instead I believe it is fair to give these instances as tentative examples of *waħəd* expanding its scope to non-‘specific’ indefinite statuses, a notion that is actually unsurprising in the light of the diachronic discussion of the form we will engage in 4.2.2.

#### 4.1.3. Disfixed Form {- /l-/} (etymological \*Ø)

Traditional outlooks on Moroccan Arabic have, as we have seen, emphasized a basic formal dichotomy between nouns marked with a “definite article” \*/l-/ (as in *l-kas* ‘the cup’) and what they have considered to be a “prefixless” unmarked form (*kas* ‘a cup;’ see 3.1.1). In the preceding discussion I have complicated these notions, arguing not only that the description of “definite article” is inaccurate for \*/l-/ (3.2) but also that the “prefixless” form given as unmarked in other works does not meet the criteria for what we would expect of an unmarked form (4.1.1). Instead I have put forth a model where all definite nouns in MA are zero-marked, meaning that the unmarked form is that which can be used to express definiteness without further modification, and that where \*/l-/ occurs it is not a determiner or marker of any particular givenness status but rather an etymologically-determined component of the unmarked noun stem.

At the same time, I have never denied that there is indeed a formal distinction between nouns which express \*/l-/ and those that do not, nor that \*/l-/ remains morphologically significant. Indeed, a majority of MA nouns can and do take both forms (*l-wəld* ‘boy,’ *wəld*; *l-kas* ‘cup,’ *kas*; *l-ktab* ‘book,’ *ktab*), and if one is to claim that \*/l-/ is not a determiner then he must inevitably provide an explanation for what its absence

indicates. Similarly, if \*/l-/, where it appears, is to be taken as a component of the unmarked nominal form from which other forms are derived, one must provide an explanation for what sort of morphological process would allow derivation of forms in which it is not present (e.g. *l-kas* > *kas*). Here I give my answer to both of these questions, stating my view that nouns for which \*/l-/ is present in the unmarked form can productively lose the element via a process of ‘disfixation’ {- /l-/}, itself representing a marked grammatical structure that is not associated with any particular givenness status but instead conveys quantification or status as a generic predicate.

I begin with an explanation of the morphological process that governs derivation of the form {- /l-}. First of all, ‘disfixation’ – or subtractive morphology, where a form loses phonological bulk while accumulating meaning – is cross-linguistically rare but not unattested. It has been shown to occur, for example, as a means of expressing repeated action with Alabama verbs (Hardy & Montler, 1988); subtractive morphology is also used to derive plural nouns in some varieties of German (Golston & Wiese, 1995) and Gaelic (Dorian, 1978). These cases serve as evidence that languages can and sometimes do equate less basic meanings with the loss of phonological material. My claim is that the disfixation of \*/l-/ in MA, in similar fashion, is used to add an additional semantic sense to the unmarked form of the noun as required by certain syntactic constructions.

In section 4.1.1 I have argued that it is the nominal form which coincides with definiteness that serves as the unmarked form of the MA noun, partly on the grounds that it represents the form from which all others can be predictably derived. Disfixation, then, represents the means through which a speaker can produce *žbəl* ‘a (single) mountain’ from the unmarked form *ž-žbəl* ‘mountain’ or *wəld* ‘a (single) boy’ from unmarked *l-wəld* ‘boy.’ This process is governed by a relatively simple phonologically conditioned rule, which I give again as: *if the unmarked noun begins with either a phoneme /l/ or a geminate consonant, either the /l/ will be lost or the first consonant will be degeminated as appropriate*. Figure 16 again shows how that this rule can be universally applied to unmarked MA nouns to predict grammatical forms of the disfixed type; no such rule, however, can be devised for the inverse operation.

<i>l-wald</i>	‘boy’		<i>wald</i>
<i>maṭiṣa</i>	‘tomato’		<i>maṭiṣa</i>
<i>amkawsu</i>	‘scythe’		<i>amkawsu</i>
<i>l-wad</i>	‘river’	{ - /l-/ } →	<i>wad</i>
<i>d-dar</i>	‘house’		<i>dar</i>
<i>l-žanb</i>	‘side’		<i>žanb</i>
<i>ž-žutiya</i>	‘pawn shop’		<i>žutiya</i>

Figure 16: Deriving the Disfixed Form from the Unmarked Noun

Because this rule is theoretically active for all nouns, speakers do not need to be aware of the etymology of a given noun to derive the disfixed form, but can instead do so on the basis of phonology alone, meaning we have an inherent answer for how *all* speakers can cognitively associate one set of forms with another. This contrasts with previously proposed models, where no such explanation is available.

In fact, the notion of disfixation even helps explain otherwise odd morphological behavior that has been documented for some borrowings into MA. Heath (1989, p. 130), for example, gives the definite forms *l-iṭru* ‘liter’ (< Sp. *litro*) and *l-iṭra* ‘monthly payment’ (< Sp. *letra*), both of in which the initial /l/ can be lost to produce the forms *iṭru* and *iṭra* where syntactically required. Since the nouns were necessarily borrowed with the initial /l/, which exists in the source stems, we can explain the forms in which it is lost only by positing a subtractive morphological rule.<sup>14</sup> Presumably speakers who were not aware of the source morphology treated the original borrowings *liṭru* and *liṭra* like other unmarked MA nouns, after which they proceeded to apply normal disfixive morphology since the phonological condition involving the presence of initial /l-/ was met.

Another point in favor of a phonologically conditioned disfixation rule is the fact that *\*l-/* is sometimes maintained for the second noun component of genitive constructions even in cases where the syntax of most Arabic varieties would typically

<sup>14</sup> Otherwise, and were additive morphology to be at play as other descriptions have claimed, we would expect *liṭru*, *l-liṭru* and *liṭra*, *l-liṭra*, which do not reflect the linguistic reality. Other Arabic dialects, by contrast, do typically preserve the source stems of such borrowings, as in Eastern Arabic *liṭr* ‘a liter,’ *il-liṭr* ‘the liter.’



require its absence. While the particle *ši* is typically accompanied by loss of \*/l-/ (*l-wəld* ‘boy’ > *ši wəld* ‘some boy’), for example, compound nouns like *bit l-ma* ‘bathroom’ (< *bit* ‘room’ + *l-ma* ‘water’) and *ħabb l-mluk* ‘cherries’ (< *ħabb* ‘seed’ + *l-mluk* ‘kings’) appear unchanged in its vicinity, yielding *ši bit l-ma* ‘some bathroom’ and *ši ħabb l-mluk* ‘some cherries.’ This behavior can be explained by the notion that disfixation is only active for or triggered by the first phoneme of a noun phrase. In addition, that such constructions convey indefinite semantic statuses is yet another argument against the presence or absence of \*/l-/ having any relevance for definiteness.<sup>15</sup>

Now that we have surveyed its morphological properties, we arrive to the question of the disfixed form’s semantic implications. In addition to calling the equivalent form (\*Ø) “unmarked,” a claim we have at this point dismissed, previous descriptions have also considered it to be the primary means for expressing non-referential objects. We have already falsified this view by showing that unmarked forms (often containing \*/l-/) very frequently occur in this semantic position (see 3.6.1.3); in fact, they occur much more frequently in this role than does the disfixed nominal form. My data, contrary to previous descriptions, show disfixed nouns that exclude the stem component \*/l-/ to be on the whole relatively rare, and likewise show that when they do occur it is with two primary grammatical functions. The first is for indicating quantity; the second is as a predicate complement. Here I give in-context examples of both types of occurrence.

Perhaps the most common context in which we can find MA nouns of the form {- /l-/} is that of absolute quantification, i.e. where the noun is part of a syntactic construction that establishes the number of the entity being referred to with complete certainty. Commonly this is in the vicinity of numbers and fractions, as in the following:

- (90) *had l-mərħəla ka-tstəyərq ərbəʕ... rbaʕa w ššrin safa, w had l-mərħəla dyal t-txmir ka-nxəmməru-ha... f š-šəbba – š-šəbba w l-ma*

<sup>15</sup> If \*/l-/ were truly analyzable as a definite article we would expect an indefinite \**ši bit ma*, not attested; instead the \*/l-/ appears to be frozen into the noun stem. A parallel case would be MA *küll bit l-ma* ‘every bathroom;’ in other Arabic varieties the /l-/ would instead indicate a definite noun, giving the translation ‘all of the bathroom.’

‘[for] this phase, it gets soaked for twenty-four hours, and during this phase we condition it in alum – alum and water’ (1-5)

- (91) *ʕanna nəss kilu dyal l-ħərša, ʕnna šwiya dyal l-məlħa, w waħəd l-biḏa, w kas dyal l-ħlib* ‘we have a half kilogram of *harsha* [mix], a little bit of salt, an egg, and one cup of milk’ (3-2)
- (92) *daba kaynin f l-mayrib kaynin žuž tanžiyat – kayn t-šanžiya l-mərakšiya w t-šanžiya š-sfrawiya* ‘now there are two tanjias in Morocco – there’s Marrakech style tanjia and Sefrou style tanjia’ (4-15)
- (93) *ka-txəlləs-u, ka-təʕti-h ši baraka, imma ʕəšra drahəm wəlla xəmstašər dərħəm wəlla miyəṭ dərħəm, šħal ma kan* ‘you pay him, give him some sort of tip, either ten dirhams or fifteen dirhams or a hundred dirhams, whatever you’ve got’ (4-16)
- (94) *ka-tšəffi f l-kas dak atay l-ləwwəl, w tʕawd tšəlləl... ʕawd tšəlləl žuž xətrat* ‘you strain that original tea into a cup, and keep rinsing... you rinse it two times’ (6-3)
- (95) *mnin ka-tʕawd nəfs l-ʔimtiħan, ka-tsənnə waħəd təlt yyam, bima ytšħħu l-wraq* ‘once you repeat the same exam, you wait three days or so, so the papers can be graded’ (8-29)
- (96) *ta-yžibu žuž qsaʕi, ka-yšəṭħu, ka-yakʷəlu-ha, ka-yšərbu, ka-ytfərtəṭ l-ʕərs* ‘they bring her two platters, they dance, they eat it, they drink, and the wedding comes to a close’ (9-46)

Relatedly, the disfixed form appears to be used to denote quantify for entities that are explicitly single. In both of the sentences below, for example, the speakers are giving instructions for how to make a certain dish. In turn, they each use the disfixed form *kas* ‘a (single) cup’ to refer to mark the nominal *l-kas* ‘cup’ as specifically quantified:

- (97) *ʕanna nəss kilu dyal l-ħərša, ʕnna šwiya dyal l-məlħa, w waħəd l-biḏa, w kas dyal l-ħlib* ‘we have a half kilogram of *harsha* [mix], a little bit of salt, an egg, and one cup of milk’ (3-2)
- (98) *w ka-ndiru kas dyal l-ma, ʕaʕi* ‘and we put in a single cup of water, that’s all’ (4-8)

These quantified usages contrast with occurrences of the same noun in other non-referential contexts where its quantification is unimportant. In lieu of expressing this additional semantic notion, the speakers simply use the unmarked form:

(99) *f l-luwl ka-ndiru-h... f l-kas, mn bʕdət tani ka-nšəlləlu ħbub atay...* ‘at first we do that in a cup, then after that we rinse the tea leaves again’ (5-5)

(100) *ka-tšəffi f l-kas dak atay l-ləwwəl, w tʕawd tšəlləl... ʕawd tšəlləl žuž xəṭrat* ‘you strain that original tea into a cup, and keep rinsing... you rinse it two times’ (6-3)

This quantitative contrast is witnessed yet again with the noun *l-ʕəša* ‘stick’ in the examples below, both of which are from the same speaker. In (101) the noun is merely a disposable instrument for an action and does not need quantification; therefore it takes the unmarked form. In (102), however, the important notion is that each the bride and the groom take a *single* stick (for the contest they are about to engage in), signifying that they are supposedly on equal ground. I give ‘she takes one stick, he takes one stick’ here for the sake of transparency, but an alternate translation is ‘she takes one stick, he takes another.’<sup>16</sup>

(101) *mnin ka-tšərrəb-u xəšša thrəb l-yūrfa dyalha ʕlaḥqaš ila bqat hašla təmma yadi iqtəlu-ha b l-ʕəša* ‘once she’s given him a drink she has to flee from her room because if she stays there surrounded [by them] they’re going to beat her up with a stick’ (9-29)

(102) *ka-taxʷəd hiya ʕəša, ka-yaxʷəd huwa ʕəša* ‘she takes one stick, he takes one stick’ (9-29)

The observation that disfixation is associated with quantity expression also helps unravel a problem encountered in Brustad (2000, p. 38), in her analysis of nominal forms in a Moroccan folk tale. The tale revolves around a woman whose husband is angry because she can only have girls – of which she has already has seven – and he wants a son. Brustad notes that the speaker does not use \*/l-/ when referring to ‘a girl’ (*bənt*) but

<sup>16</sup> Also refer to (45) for an example of one might ignore the quantitative aspect entirely, though I feel this is an oversight.

does use \*/l-/ when referring to ‘a boy,’ (*l-wəld*) a fact that she tentatively attributes “to the social importance of the male child, giving him a higher degree of individuation.”

- (103) *gat-lu wlədt bənt. gal lha guli li šnu wlədti rah ila wlədti l-bənt*  
*ya-ndəbh-ək w ndbəh-ha. ta šaft-u zayd lha b l-mus, gat lu had, wlədt*  
*l-wəld* ‘she told him, “I had a girl.” He told her, “tell me what you had  
– if you had a girl, I will slay you and slay her. Until she saw him  
coming at her with the knife. She told him, “calm down, I had a  
son.”’ (9-29)

What I see as more relevant to the formal difference, on the other hand, is that the speaker’s choice of the disfixed form *bənt* reflects an additive element, particularly since we are specifically aware that the woman in the story has *seven* daughters, and *one* more girl will give her *eight* (the number of children she has had without having any boys being important to the narration). In this light an apt translation may be not ‘I had a girl’ but rather ‘I had yet another girl.’ The non-referential ‘boy’, on the other hand, does not need to be quantified (or have \*/l-/ disfixed) because what is important to the characters is not that there be *one* boy specifically, but simply a boy.

My impression is that most previous literature has failed to recognize this quantitative dimension for singular nouns of this type, perhaps because authors have automatically assumed what they considered the “bare” form (and I call \*/l-/disfixed) to indicate indefiniteness rather than another semantic notion. As a result, their glosses may not have fully captured what speakers intended with their choice of morphology. For the following example from Maas we in 3.1.1, for example, I would pose an alternative gloss:

- (104) *küll nhar ka-ntiyyəb hawli* ‘every day I cook a sheep’ (Maas, 2011,  
p. 154) > ‘every day I cook one sheep’

In addition to strictly numerical uses, the disfix {- /l-/} can also indicate absolute quantities involving ‘all’ or ‘none.’ In keeping with this function, we find it used consistently alongside the particles *küll* ‘every,’ *həttə* ‘(not) even,’ and *bla* ‘without.’ We also find it alongside *ayy/ašmən* ‘any/which,’ which again implies a specific numerical

quantity, as well as with superlative constructions (such as *axər* N ‘the last N’) that necessarily specify a single entity. A number of examples from my data are:

- (105) *küll waḥəd fin ka-ymši ka-tbqa l-ḡarusa f d-ḡar dyalha* ‘everyone goes his way and the bride remains with the groom at her [new] house’ (2-11)
- (106) *ka-nzidu ḡlih s-skk<sup>w</sup>ar, b š-šiba wəlla b ši aḡšab, küll waḥəd aš ka-yšərb, küll waḥəd w ašmən ḥaža ḡziza ḡndu* ‘we add some sugar to it, with wormwood or some herbs – everyone drinks whatever [he likes], everyone and anything he likes’ (5-7)
- (107) *küll mdina, baš ḡawd tani ka-ttmiyz b l-mudun* ‘every city is distinct from the others’ (7-33)
- (108) *küll šhur, küll waḥd ka-ybyi ḡawd tani aš ka-yak<sup>w</sup>əl fih w hakk<sup>wa</sup> t-taqalid dyalna f l-mayrib* ‘for each shour<sup>17</sup>, everyone eats whatever he wants, and that’s how our customs go in Morocco’ (7-29)
- (109) *ka-tsaḡdək tḡawzk... l-muškil, w lli bya itḡawd l-ḡam yəḡni ši ḥaža tabiḡiya ma-kayn ḡtta muškil* ‘they help you get through the problem, and if anyone wants to repeat the year, I mean it’s something natural, there’s no problem at all’ (8-39)
- (110) *mnin ka-thəzz ž-žlal dyalha, axər ḥaža ka-yžbəd hiya dik l-ḡuta* ‘as she’s taking her gifts, the last thing that he pulls out is that fish’ (9-28)

All of the above usages again support the notion that {- /l-} is very closely associated with absolute quantification. This usage typically overlaps with indefinite meanings, though not exclusively (as in 109). At the same time, as we have seen elsewhere, it is not appropriate to think of {- /l-} as an exclusive indefinite marker when the unmarked form can also indicate indefinite statuses for the same entities as well. The most we can say, then, is that in one of its major usages the form {- /l-} behaves as a sort of morphologically-derived quantitative classifier, and is in turn required by particles that necessarily indicate quantity. All of the above particles demonstrate this requirement, as

---

<sup>17</sup> The pre-fast meal during Ramadan.

does the categorical negative particle *la*, as in *la wəld* ‘(there’s) no boy.’<sup>18</sup> In addition, the particle *ši* has implications for quantity and requires the nominal form {- /l-}, but we treat it independently below.

There is one other major role that {- /l-} plays productively, which is to indicate the predicate complement. This usage hints at a larger discussion that I do not fully engage in this study, but a brief overview is nonetheless relevant. Nouns of the semantic type that Givón calls ‘generic predicate’ are actually relatively infrequent in my data. In fact, there are only a couple of overt examples in the entire set of texts, both of which we have already seen and I give again here:

(111) *ana təlmid... b s-sana t-tanya bakaḷuṛya* ‘I’m a student in the second year of the baccalaureate’ (8-1)

(112) *n-nhar l-ax<sup>w</sup>ər yadi tnuḍ, ka-tšqa f darhūm ʕəla asas annəha šafi raha wəllat mərə ka-ttʕətamd təmmaya* ‘the next day she’ll get up and do housework, with the logic that she’s now become a woman and is being depended upon there’ (9-43)

Although such clear examples are infrequent, we can nonetheless also identify a few more instances of ‘implied’ predicates, typically parenthetical asides where the speaker interjects a disfixed noun to clarify the identity of an entity that has already been introduced. We can see these usages in the following:

(113) *bħal matalan l-maʕriḍ lli ka-ykun bħal qəlʕat mguna, bħal haduk l-məʕrufin... iyih, mwasim* ‘like for example the exposition that they have in Kalaat M’Gouna, like those well known ones... yeah, [they’re] seasonal festivals’ (1-23)

(114) *ya-nqəddəm wahəd l-wəšfa dyal... wəšfa myaribiya, ʕibara ʕla hərša* ‘I’m going to present this recipe, [it’s] a North African recipe, known as harsha’ (3-1)

---

<sup>18</sup> Categorical negation can also be expressed for MA verbs (Brustad, 2000, p. 306). Verbal negation is typically expressed with a circumfix *ma-* + *-š* as (*ma-byā-š* ‘he doesn’t want’), but for categorical negation the suffix *-š* is dropped (*ma-byā* ‘he doesn’t at all want’), showing that lack of phonological bulk can indicate the quantity “none” in other contexts as well.

- (115) *w ma-nžəhti-š, kayn lli ka-ymši ka-yxdəm, imši idxəm ši blaša... ši məʕməl wəlla ka-ymši idir l-mikanik wəlla... küll waḥəd aš ka-ydir, yəʕni mxəddmat ʕadiyin* ‘and if you didn’t pass, there are some people who go work, go work somewhere, some sort of factory or they become mechanics or... everyone does his own thing, I mean [they’re] normal vocations’ (8-31/32)

Although such attestations are limited in my data, these morphological patterns are typical of MA speech and I have not heard non-disfixed nouns in the predicate position for any competent speakers, with the possible exception of a few Berber-dominant late bilinguals. For the entire span of semantic notions expressed in Givón’s wheel diagram, then, the generic-predicate sector represents the only semantic sense for which the unmarked form described in 4.1.1 does not generally occur.

Unlike with ‘type identifiable’ objects, I do not see quantification as relevant to this grammatical requirement. Instead my best explanation of why generic predicates would disallow \*/l-/ where other indefinite meanings allow it is that, while the historical association of \*/l-/ with definiteness for the head noun has dissolved, it has been maintained for the nominal attribute, of which the predicate is expressive. This possibility would also account for constructions such as *l-ma tayb* ‘boiling water’ and *l-ʕərs ʕadi* ‘a normal wedding, where the adjective but not the noun shows a correspondence between the absence of \*/l-/ and indefiniteness. Since this study’s primary concern has been definiteness marking as it is expressed in the form of the primary noun, I do not develop this argument further here, but it is certainly in need of attention in future research.

Besides for quantification and with predicates, other attestations of the disfixed noun can be identified, but all of these are of limited productivity or, alternatively, appear to be under the direct semantic influence of MSA (where the comparable form does represent indefiniteness). As is the case in other Arabic varieties, nouns with attached possessive pronouns necessarily exclude the element \*/l-/, as does the first element of the Semitic “construct state” indicating a genitive relationship between two nouns:

- (116) *ka-tnuḍ ṣ-ṣbaḥ ṣadi, ka-tlbəs, ka-ttgədd, ka-tsnna ražəlha ha yži f l-ṣəṣiya* ‘she gets up in the morning like normal, gets ready, and awaits her husband, who will be there in the afternoon’ (9-21)
- (117) *l ya-tži, bb<sup>w</sup>aha, mm<sup>w</sup>ha, ya-yžibu lha l-fṭur* ‘once she’s there her father and mother are going to bring her breakfast’ (2-7)
- (118) *ṣawd tani faš ka-tak<sup>w</sup>əl l-fṭur, ka-y... ka-tx<sup>w</sup>ərž l s-saḥa tx<sup>w</sup>ərž mṣa l-ṣaṭila dyalk, wəlla šhabk* ‘once you’ve eaten iftar you go out to the [public] square – perhaps with your family, or your friends’ (7-19)
- (119) *xu l-ṣəris... ka-yži, ka-yṣiyṭ ṣəla l-ṣərusa* ‘the groom’s brother comes, calls for the bride’ (9-38)
- (120) *daba l-ṣayalat ila ma-naḍu-š baš ihəḍru ṣla l-ḥəqq dyalhūm w ḥəqq z-zərbiya, ana ka-nḍenn bnni... blli had z-zərbiya yaḍi... yaḍi tmūt nihaṭiyan* ‘now, unless women get up and start speaking about their rights and the rights of the carpet, I think that the carpet’s going to... it’s going to die once and for all’ (1-29)
- (121) *had šhadət l-bakalurya ka-təṣtabər mūhimma, ḍaruriya* ‘this baccalaureate degree is considered important, indispensable’ (8-35)

I consider neither of these constructions to be particularly productive because, excepting their use with a small class of words that includes kinship terms (116-119) and body parts, they only very rarely with native MA nouns, genitive relationships for which are typically constructed analytically with a particle *dyal* or *d*. Where they otherwise occur they are almost always in MSA-like constructions typical of educated speech (as in 120 and 121), which are much more like code-switching and likely do not reflect the native syntax of MA.

For the same reason, while I will not argue that \*/l/-less nouns in MA speech cannot indicate non-referential objects that do not involve a quantitative element, I do hold that where they appear they are almost always sourced from MSA, the syntax of which influences the nominal forms. In (122), for example, *žamṣiyat* ‘associations’ and *tṣawniyat* ‘cooperatives’ represent direct borrowings from MSA, which may account for why they are not given as *l-žamṣiyat* and *t-tṣawniyat* like we might expect for native MA nouns (at least based on the patterns we have seen in this study); further evidence of



MSA influence in the immediate phrase can be witnessed in the speaker's production of the phoneme /ʔ/ in *nisaʔiya* 'womens'.<sup>19</sup> The same would go for the example (123) we earlier saw from Youssi, in which *rtakəb ʒarima* is a direct borrowing of the literary phrase *irtakaba ʒarīmatan* with only phonological modification:

(122) *kaynin l-ʕayalat daba lli dayrin ʒəmʕiyat w tʕawniyat nisaʔiya* 'there are these ladies nowadays who have put together women's associations and cooperatives' (1-18)

(123) ...*rtakəb ʒarima* 'he committed a crime' (Youssi, 1992, p. 144)

Whether increasing levels of literacy in Morocco will lead to more MSA-like uses of the disfixed form – where it does not have a quantitative connotation and is the primary means for representing an object of the status 'type identifiable' – remains to be seen. For the present, however, this effect only appears marginally with items that already have formal semantic connotations, and I have seen no evidence of it in play with core MA vocabulary (which is why *kas* 'a single cup' continues to contrast with *l-kas* 'a cup,' although the former would convey both meanings in the MSA model).

To review this section's key arguments, the disfixed form {- /l-/} is a marked structure that is productively derived from the unmarked form via a rule where the etymological element \*/l-/ is removed from the stem. It plays two main functions. The first of these is to indicate objects that are necessarily quantified, whether numerically or via a particle that indicates a certain number. The second is to indicate the generic predicate, for which it is the only form that can be used (to the exclusion of the unmarked form). Where it otherwise occurs, what appears to be the disfixed form may simply be insertional MSA code-switches with slight phonological adaptation, displaying the semantic connotations of the formal variety rather than the speaker's native MA. We now look at a special usage case for {- /l-/} alongside a particle that has typically been given as an indefinite article in previous descriptions.

---

<sup>19</sup> The glottal stop /ʔ/ is typically pronounced only in recent borrowings from MSA, and does not occur in MA items inherited from Old Arabic.

#### 4.1.4. Particle *ši* {- /l-/}

In my descriptive model, *ši* is a particle that functions similarly to *küll* ‘every,’ *həttə* ‘(not) even,’ and *bla* ‘without’ and *ayy/ašmən* ‘any/which’ in that it indicates quantity and syntactically requires that the following noun be marked with the related disfix {- /l-/}. As opposed to the other particles which indicate a known absolute quantity, however, *ši* merely indicates potential quantity, or that the entity is individuated and could be more specifically qualified were more information available to the speaker. This description, in turn, simultaneously resembles Harrell’s claim that *ši* is a “potential article” as well as Caubet’s insistence that it is a quantifier. As we have seen in 3.1.4, *ši* has typically been associated with a ‘referential’ status, an observation that can be explained by the fact that a speaker, in using *ši*, does intend to refer to a unique (and quantifiable) entity or set of them as opposed to any of their type. It cannot, however, indicate the status ‘specific’ because the speaker does not have the prerequisite knowledge to uniquely identify the entity, despite the importance of its individuation in the discourse.

Besides the fact that I see *ši* as requiring a morphologically disfixed form {- /l-/} rather than what previous descriptions have called a “bare” form, I generally agree with earlier conclusions that *ši* is used to mark ‘referential’ statuses. A majority of its occurrences in my data support this notion. In (124), for example, the speaker uses *ši* to indicate that the *l-baraka* ‘tip’ to which he refers is an important part of the chain of events in the discourse, but that he cannot explicitly identify (or quantify) it, as seen in the different possible monetary values he suggests:

- (124) *ka-txəlləş-u, ka-təʕti-h ši baraka, imma ʕəšra drahəm wəlla xəmşəşər dərham wəlla miyət dərham, šhal ma kan* ‘you pay him, give him some sort of tip, either ten dirhams or fifteen dirhams or a hundred dirhams, whatever you’ve got’ (4-16)

Even where the ‘potential’ aspect of *ši* is not as clearly delineated as above, the contexts in which it occur still strongly suggest that the speaker is referring to an entity that can be individuated (i.e. is referential) but cannot be uniquely identified from his or

her perspective as the speaker, often because he or she is describing events from another person's perspective. In the following, for example, the speakers are referring to the underlined objects from the respective perspectives of the student who didn't pass his exam (125), the newlyweds' family (126), the carpet weavers in Taznakht (127), and Moroccans in general (128):

(125) *w ma-nžəhti-š, kayn lli ka-ymši ka-yxdəm, imši idxəm ši blasa... ši məsməl wəlla ka-ymši idir l-mikanik wəlla...* 'and if you didn't pass, there are some people who go work, go work somewhere, some sort of factory or they become mechanics or... (8-31)

(126) *ya l-ṣaʔila l-qriba hiya lli hədəra, ma-ta-ykun ši hədd bərrani bʕid* 'it's just close family who's in attendance, there's not anyone distant or from outside [the family]' (9-4)

(127) *w byina ši mihražan wəlla ši məṣriḍ, baš ttfərrəf z-zərbiya l-wawazgitiya* 'we want some sort of festival or some sort of exposition, so that carpets from the Wawazgit region will get exposure' (1-22)

(128) *ka-nzidu ʕlih s-skkʷar, b š-šiba wəlla b ši aṣṣab, küll waḥəd aš ka-yšərb, küll waḥəd w ašmən haža ʕziza ʕndu* 'we add some sugar to it, with wormwood or some herbs – everyone drinks whatever [he likes], everyone and anything he likes' (5-7)

All of these usages would then qualify as 'referential' in our working model of givenness. This semantic connotation seems to account for the majority of the occurrences of *ši* in my data set. There is only one possible exception to this tendency, which is that in which *ši* occurs alongside the particle *hətta* '(not) even' in negative contexts:

(129) *kayn-ši... yir ka-yləṣbu yir binathūm, ma-kayn l-gʷərb wəlla hətta ši haža* 'there's not... they just play [music] amongst themselves, there's not a band or anything' (2-16)

(130) *hətta ši ktab ma tbaʕ* 'not a single book was sold' (Caubet, 1983, p. 241)

Since the entity referred to does not even exist, it is unlikely that the speaker means to assign it a referential status. In turn, what we might identify here is the semantic extension of *šī* into the status ‘type identifiable.’ This is likely only one of many ongoing semantic changes that Moroccan has seen in the past and is still undergoing, as we discuss below. In general, however, *šī* still appears to be the dominant form for marking referentiality and is largely restricted to this sense. This is in keeping with previous descriptions, meaning the traditional analysis of this form is in lesser need of qualification and complication than that of the others above.

## 4.2 DIACHRONIC DIMENSIONS

In the preceding discussion I have made the novel argument that Moroccan Arabic, contrary to all previous descriptions’ suggestions, no longer has what we can truly call a definite article in the synchronic sense (see 3.2). Relatedly, I have shown evidence that it is typically the nominal form that incorporates the etymological article *\*l-* that now serves as the unmarked form of MA noun, appearing throughout the full range of the Givenness Hierarchy and serving as the base form from which other forms with more specific semantic connotations can be derived (4.1.1). Among these are the article *waḥad* (Ø) that usually indicates ‘specific’ statuses (4.1.2), the disfixed form {- /l-/} that is used for quantification and generic predicates (4.1.3), and the particle *šī* {- /l-/} that typically expresses referentiality (4.1.4). Because this description diverges quite radically from others, which have considered MA forms to express semantic notions similar to those of their equivalents in other Arabic varieties, a question that it necessarily introduces is how MA might have undergone the sort of semantic reshuffling and formal reanalysis that led to its divergence from related Arabic varieties.

In this section I provide a brief overview of not only how I believe MA has undergone the semantic shift in question, but also why we might even have been able to expect some of the outcomes proposed in the synchronic description above. Some of this discussion is inevitably speculative for the simple reason that a great deal of debate still surrounds the origins of the modern Arabic dialects (see Watson, 2011), and we cannot

know with certainty what MA looked like in the distant past. At the same time, where MA diverges from all other major dialects and CA it would seem that there are grounds for claiming that there has indeed been diachronic change. I take those form/meaning associations that are common to all other Arabic varieties as the likely starting point for MA definiteness system, and the current study's synchronic description as its end point. What is at issue, then, is how the dialect group moved the first point to the latter in the time that has passed since Arabs arrived to North Africa fourteen hundred years ago.

There is more than one candidate for what may have sparked this shift. One is a natural tendency for change in a certain direction, which may have already been present in the language even before Moroccan dialects were geographically separated from other Arabic varieties. This, however, cannot single-handedly account for whatever change has occurred in MA because it has apparently not effected the same results in these sister dialects. What we might then posit as a catalyst for the semantic changes in MA is a history of language contact. As we saw in 2.1.1, MA has been in relatively intense long-term contact with Berber, original speakers of which probably represented a large contingent of second-language learners who swelled the ranks of Arabic speakers soon after the language's arrival to North Africa. In doing so, they may have reconfigured the semantic connotations of MA forms to more closely approximate those of their native languages, thus establishing the system of definiteness marking we see today. I am of the opinion that neither internal tendencies nor contact pressures are mutually exclusively as agents of semantic change, and that they likely worked together to co-facilitate the diachronic developments MA has apparently witnessed. For this reason I mention them both, as appears applicable, in the following discussion of how each of the current MA forms and its semantic implications may have evolved.

#### **4.2.1. Unmarked Form Ø (etymological \*/l-/)**

We begin with a treatment of \*/l-/, or how a form that was once probably marked and had very specific semantic restrictions came to display the opposite traits, with a former article \*/l-/ being incorporated into the bare noun stem (see 4.1.1). In essence, this

is the question of how *l-wəld*, once exclusively ‘the boy,’ was semantically reanalyzed to simply represent ‘boy.’ There is little doubt that \*/l-/ was at one point in MA’s history a fully definite article, semantically restricted along the same lines as its cognate /(V)l-/ in all other major Arabic dialects. This was likely the case when Arabs arrived to Morocco, an assumption we can deduce from the fact that no major historical or modern varieties outside of western North Africa seem to use /(V)l-/ in any other capacity. Since we today have strong evidence that \*/l-/ can appear with MA nouns falling nearly anywhere along a definite-indefinite continuum, however, we must assume that the historical article \*/l-/ at some point lost its primary semantic association with the ‘definite’ status, even if the form itself was maintained in one way or another.

As a means of understanding the historical semantic trajectory that \*/l-/ appears to be have been following, it is helpful to consider its origins far back in the history of the Arabic language, even long before the post-Islamic expansions that brought Old Arabic to North Africa. Here I turn to Semiticists’ work on proto-Arabic, which is largely in agreement on the processes that led to the development of /(V)l-/ in Arabic varieties. Although opinions differ on precisely what its original form may have been, the dominant consensus among such scholars is that the Arabic definite article /(V)l-/ developed out of a demonstrative or deictic particle (Rubin, 2005). This demonstrative form, the reasoning goes, underwent a process of semantic extension in which speakers ultimately generalized it to an article that could represent definiteness for any entity, without regard to deixis.

This theory finds support in the fact that similar developments have been observed for many other languages. De Mulder & Carlier (2011), for example claim that “even though the grammatical category of [the definite article] is far from being universal, the grammaticalization process that leads to its development exhibits cross-linguistic regularities: in the majority of cases, the definite article originates from a weakened demonstrative.” Adopting a proposal by Greenberg (1978), the authors describe this semantic change as involving the first two of a series of stages through which the original demonstrative form evolves: it begins as a demonstrative, then

<b>1. Demonstrative</b>	An item serves as a nominal modifier for both spatial deictic and for anaphoric reference.
<b>2. Definite Anaphoric Marker</b>	The item is no longer associated with spatial reference; its main function is now to refer to entities mentioned earlier in discourse.
<b>3. Context-definite marker</b>	In addition to previous mentions, the item also refers to definite entities that are recoverable via contextually available knowledge.
<b>4. Marker of “Semantically Definite”</b>	The item is no longer restricted to contextual knowledge; it may refer to any entity that is identifiable via world knowledge, including both individual and generic entities.
<b>5. Indefinite-Specific Marker</b>	The item is no longer restricted to definitely identifiable entities; it may in addition refer to specific indefinite entities, that is, entities that are not necessarily identifiable to the hearer. It can simply assert existence.
<b>6. Article Loss</b>	The item loses its association with referentiality; it no longer has a pragmatic or semantic function, it can occur in any context and with any noun, and it may be exapted for other functions such as noun classification.

Figure 17: Heine’s (2012) Grammaticalization Stages for the Definite Article

becomes a definite article, then a ‘specific article,’ then a ‘noun marker.’ Heine (2012) gives a more nuanced six-stage version of this model (Figure 17), detailing the semantic changes that each successive stage entails. All of the authors seem to agree on the same major points here, namely that the grammaticalization chain typically begins with a demonstrative and ends with loss of the article entirely, with the form’s semantic role as a definite article representing an intermediate phase.

Following this model, what we know about the history of old Arabic suggests that by the time of the Islamic expansions and the dispersion of the Old Arabic varieties that gave rise to the modern dialect groups, the form /(V)l-/ had already passed through the first few stages of the above grammaticalization process and arrived to stage 4, where most Arabic varieties remain today. In some senses it is actually rather surprising, after more than a millennium, that there have not been shown to be multiple Arabic varieties that have advanced to the latter stages of this common cross-linguistic cline. At the very least, in any case, it should not be unexpected that at least one of the many extant Arabic dialect groups would have done so, which – when we consider how closely MA seems to match Heine’s description of ‘article loss’ – certainly seems to be the case.

If we map the latter three stages (4-6) of Heine's grammaticalization cline onto our modified Givenness Hierarchy, we can envision what it entails as a rightward extension of the forms that originally represented definiteness into the indefinite tiers of the hierarchy. Similar, if we map it onto our wheel diagram of Moroccan Arabic (as in Figure 14, in 4.1), we can imagine this diachronic process of semantic change as a clockwise expansion of the same forms into the indefinite sectors of the wheel. After the evidence we have seen in this study, it is clear that \*/l-/ has indeed followed this path. The truly challenging question that remains is why, especially when other Arabic dialects have remained at stage 4 on Heine's cline, Moroccan Arabic seems to have advanced into the latter stages more quickly.

The most convincing answer to this question is probably that language contact has played a role, perhaps acting as a catalyst to push MA further along a cline the language had already been following historically. As we established in section 2.1, the historical evidence suggests that one of the major factors associated with the spread of Arabic in Morocco was the linguistic Arabization of already extant Berber-speaking tribes. What we can infer from this is that the original Arabic dialects to arrive to the area must have, at points in their history, had fairly significant numbers of second-language (L2) learners, meaning Berber would have played the role of a substrate language (Thomason & Kaufman, 1988). In turn, it is not unreasonable to expect that the process of imperfect group learning may have led to semantic rearrangement of the sort that seems to have occurred in MA, particularly as adult learners made "interlingual identifications" (Weinreich, 1953) between forms that they saw as structurally congruent.

Winford's (2003, p. 251) overview of language contact in scenarios of shift from one language to another zeros in on this notion of perceived congruence, showing that it is one of the major factors that leads speakers to restructure an L2. He gives the following "congruence-based constraints," which he frames as the outcomes of the psycholinguistic processes governing the acquisition of a new language:

- "L2 structures that are highly congruent with those in the L1 will be acquired more easily (and successfully) than those that are not."



- L2 forms that are partially congruent with or partly similar in semantics or function to L1 forms will tend to be reanalyzed on the model of the latter.
- Certain L2 structures or elements that have no counterparts in the L1 may be difficult to learn. Learners may simply ignore such structures or employ L1 strategies by way of compensation.”

How these constraints may have applied to Berber L1 speakers acquiring MA, especially as it relates to the acquisition of definiteness marking, is fairly easy to envision. Although knowing what Berber looked like fourteen hundred years ago poses many of the same challenges we have seen for Arabic, if not more, that no Northern Berber varieties of which I am aware have anything resembling a definite article (Abdel-Massih, 1971; Dell & Elmedlaoui, 2002; Kossmann, 2007) suggests that this has been the case since at least before Arabs arrived in North Africa. These Berber speakers, then, would have found the then definite article \*/l-/ of MA’s predecessors without a readily identifiable counterpart in their native language. They would have, on the other hand, been able to recognize that forms incorporating \*/l-/ overlapped in meaning with the “inherently definite” unmarked nouns of their own L1 in the definite range of the semantic spectrum, and perhaps seen them as congruent in this sense. Once conceptual equivalency had been established between the Arabic definite form containing \*/l-/ and the unmarked Berber noun, which has nearly full range of all possible givenness statuses, the stage was set for rapid extension of \*/l-/ into the semantic range of the ‘congruent’ Berber form, including the indefinite statuses.<sup>20</sup> This may have sparked the erosion of \*/l-/’s association with definiteness.

I am not the first to posit Berber speakers’ L2 acquisition of Arabic as responsible for the wide semantic range \*/l-/ has in Moroccan and nearby varieties (though I do believe I am the first to specifically call this change ‘article loss’ in contemporary MA). This notion has, in fact, been discussed as responsible for “article agglutination”

---

<sup>20</sup> As a corollary to this process we can cite the fact that all but the earliest Arabic loans into Berber tend to display \*/l-/ (Kossmann, 2012; Marouane, 2009), supporting the notion that L2 learners did seem to take the \*/l-/ -prefixed forms as the most congruent with their unmarked nouns.

(Corriente, 2008, p. lxvi) in Arabic borrowings into Spanish, where both the Romance and historical Arabic article are expressed (as in *el alcalde* ‘the judge’ < Sp. art. + Ar. *al-qāḍī*). Corriente summarizes this view as it applies to Iberia, saying that “the Muslim invaders were superficially Arabized Berbers who, lacking an article in their native language and being therefore scarcely able to master the rules of its usage, attached it permanently to the Ar. loanwords acquired by Br., as well as to every substantive in the Ar. they learned, spreading this usage in the areas invade by their troops, the Iberian Peninsula and wide expanses of Western Africa.” While Corriente’s view is primarily interested in Spain, I see no reason why the same dynamics would not have been at play on the North African mainland.

Substrate influence, that said, is not necessarily the only contact scenario that can possibly describe the relationship between MA and Berber. The fact that, even after centuries of language shift, there still remains a large minority of Berber speakers in Morocco shows that there has been also long-term maintenance of Berber in many communities even in the face of intense contact. On these grounds I ultimately side with Tilmatine (2011, p. 1012), who claims that contact between the two families is not as simple as a substrate/superstrate relationship and “there are indications to suggest a likely hypothesis that there is a convergent evolution of the two North-African native languages.” This interpretation too, though, would support the notion that Berber contact played a role in the semantic reinterpretation of definite \*/l-/ and its diffusion into indefinite contexts as speakers of both languages bowed to the pressures of “structural convergence” (Heath, 1984) and rearranged inherited material to reflect similar semantic categories.

At this point we have looked both at cross-linguistic trends in the development and decline of definite articles and at contact explanations for how such an article might be reinterpreted by L2 learners, and seen that both could contribute to an explanation for how MA’s former definite article \*/l-/ lost its strict association with that semantic sense and came to both represent other statuses and eventually serve as an unmarked form. My view is that the most plausible scenario involves a combination of these two outlooks,

where contact with the article-less Berber sped MA along a grammatical cline on which it had lay dormant since the original Semitic demonstrative was grammaticalized into a definite article in proto-Arabic. I likewise see this development, whatever specific mechanism facilitated it, as the primary driving force in the semantic reshuffling that has taken place in MA, in that while other morphosyntactic arrangements can express various degrees of givenness, their forms and semantics have ultimately been delineated by the diachronic shift \*/l-/ > Ø. We now briefly turn to these other forms.

#### 4.2.2. Article *wahād*

The etymological origins of *wahād* are thoroughly transparent, so that even today it is phonologically identical to the MA numeral *wahād* ‘one,’ itself cognate with the *wāhid* of CA and most other Arabic varieties. What is not immediately clear, on the other hand, is how or when what was originally a numeral was semantically extended into the role of an indefinite article that largely marks specific statuses. As we have seen in 2.3.3, MA is not the only modern Arabic dialect that has a reflex of Old Arabic *wāhid* as a sort of indefinite article, though the similar forms found in eastern Arabic do have syntactic restrictions that MA’s *wahād* does not display. In Classical Arabic, by contrast, *wāhid* expresses an exclusively numerical meaning and does not show any signs of behaving as an indefinite article. The primary questions we can pose for the the article *wahād*, then, relate to how long it has been present in MA and what its relationships with similar forms in other dialects may be.

It is not at all out of the ordinary to find what was originally a numeral playing the role of indefinite-specific article. This is, in fact, one of the most common cross-linguistic means through which an indefinite article arises (Heine, 2012). In the same way that that the semantically definite article is grammaticalized out of an earlier demonstrative, the typical grammaticalization cline for an indefinite article begins with a numeral ‘one.’ According to Heine, who gives this process as yet another series of stages of semantic reanalyses (Figure 18), the original numeral is first semantically extended to the role of a presentative marker that marks entities important to the discourse, then to ‘specific’

<b>1. Numeral</b>	An item serves as a nominal modifier denoting the numerical value ‘one.’
<b>2. Presentative Marker</b>	The item introduces a new participant presumed to be unknown to the hearer, and this participant is then taken up as definite in subsequent discourse.
<b>3. Specific-Indefinite Marker</b>	The item presents a participant known to the speaker but presumed to be unknown to the hearer, irrespective of whether or not the participant is expected to come up as a major discourse participant.
<b>4. Non-Specific Indefinite Marker</b>	The item presents a participant whose referential identity neither the hearer nor the speaker knows.
<b>5. Generalized Indefinite Article</b>	The item can be expected to occur in all contexts and on all types of nouns except for a few contexts involving, for instances, definiteness marking, proper nouns, predicative clauses, and so on.

Figure 18: Heine’s (2012) Grammaticalization Stages for the Indefinite Article

entities on the whole, then to solely referential entities, and then to indefinite nouns of the status we have previously referred to as ‘type identifiable.’ Here we can again envision this semantic shift, at least beginning at stage 3, as an extension of the form from the ‘specific’ (Givón’s ‘referential-indefinite’) status either rightward or clockwise on the Givenness Hierarchy or wheel diagram, respectively.

It is clear from our discussion of *waḥad* in 4.1.2 that the article, at least as it is primarily used, would correspond with stage 3 of Heine’s grammaticalization cline, from which we can conclude that it has already passed through the former two stages. In addition, as we saw in the same section, there is some preliminary evidence that for at least some speakers, it has begun to be generalized to non-specific entities, as witnessed in *waḥad l-biḍa* ‘an egg’ and *waḥad l-qanba* ‘a piece of burlap.’ This is in keeping with what we would expect were the article indeed following a typical cross-linguistic grammaticalization cline, slowly losing its association with specificity and being generalized to all indefinite entities. If this is the case, then, progress into stages 4 and 5 appears to be in an early phase.

Brustad (2000) has shown that other Arabic dialects – specifically, Syrian, Egyptian, and Kuwaiti – also have an “indefinite-specific” marker *wāḥid*, of obvious

etymological relationship to MA's *waḥad*. Brustad does not speculate on the relationship between these forms explicitly, but the question of whether *waḥad* in MA may have a common origin is an intriguing one. In favor of a common origin is the etymology of the markers in question; against it, however, is their semantics and syntax. First, when we consider how common a cross-linguistic development the process numeral > indefinite article is, there is no particular reason to doubt that it may have been developed independently more than once. Furthermore, my reading of the use of *wāḥid* in Eastern dialects is that it can only be used to introduce especially prominent discourse participants and would probably best be described as a 'presentative marker.' This contrasts with Moroccan's stage 3 (or nascent stage 4) article *waḥad* and means the Eastern Arabic equivalent has really only advanced one stage, perhaps suggesting an origin much younger than the early Islamic expansions. The final and most convincing argument piece of evidence for MA's *waḥad* having been developing independently, that said, is the fact that it is built on the nominal form that would have historically represented definiteness (typically those with \*/l-/). What this suggests is that the former definite article \*/l-/ would have already had to have been semantically extended into the indefinite semantic range at the time MA's *waḥad* began to be grammaticalized as an indefinite article. Since we have seen above that this development likely occurred after the arrival of Arabs to North Africa, it works in favor of a local post-migration origin for *waḥad* as well.

If *waḥad* did arise in Morocco, it is likely that contact played at least some role in its development. There has been no shortage of scholars who have pointed out the etymological and functional similarity of the MA structure *waḥad* N and the Berber arrangement *yan/yat* N, where *yan* and *yat* can be traced back to the masculine and feminine forms for numeral 'one,' respectively (Chafik, 1999; Corriente, 2008; Marçais, 1977; Tilmatine, 2011). Not only is the congruence between the Berber and Arabic structures difficult to deny, but my impression is that they also tend to express the same explicit semantic notion of specificity. In the following two examples from a Tashlhiyt Berber text, for example, we can contrast the indefinite-specific *yan l-kas* 'this

[particular] cup,’ which refers to a unique entity that has exceptional qualities, with the non-referential *l-kas* ‘a cup’ of no import.<sup>21</sup>

(131) *iga l-mjdub, iga l-ḥakim. ifka yas ṛbbi yan l-kas. ad ukan gis imikk n watay, ig as yat talwizt n d-dhab* ‘once there was a Sufi who was a magician. God gave him this cup. As soon as he poured some tea in it, it made him a golden coin’ (Stroomer, 2001, p. 117)

(132) *iṣmmr talbriqt nns, imdi tt, iffī y l-kas n ž-žaz* ‘he filled his teapot, put it [on the fire] and poured [tea] in a glass cup’ (ibid.)

What these examples reveal is that *yan/yat*, in at least some Berber varieties, does indeed serve as a specific article but has not yet been generalized to all indefinite nouns, just as is the case with *wahād* in MA. The question these obvious similarities prompt, then, is that of exactly how contact might have played into such isomorphism; a comprehensive answer, however, is difficult to delineate at the moment. Previous discussions have given an almost exclusive sense that MA modeled its article *wahād* on a pre-existing Berber construction, but without historical evidence to show with certainty that Berber did have a specific article *yan/yat* at a time when MA did not have an equivalent there is nothing to rule out the possibility that *wahād* and *yan/yat* could have been co-grammaticalized in both languages at the same time, or that the feature could have even arisen independently in MA and then served as a model for Berber. If anything, the fact that grammaticalization of the numeral ‘one’ into a specific article is so common cross-linguistically only makes the latter possibilities more plausible. Since neither of the two colloquial varieties in question has traditionally been written, that said, deciding between the various accounts is a difficult prospect.

Some work that could perhaps cast light on this problem is Heine (2012) and Heine & Kuteva’s (2003) writing on ‘polysemy copying,’ which has shown that although one language can copy grammatical structures directly from another, these structures must still pass through the typical stages of a grammaticalization cline and tend to lag a

---

<sup>21</sup> These examples also highlight the structural/semantic isomorphism of MA and Berber, in both of which *l-kas* ‘cup’ can be used for non-referential objects as well as definite entities.

stage behind that of the model language (as is the case with Basque's indefinite article, borrowed from Spanish). On one hand, this might imply that if MA had calqued the structure *waḥad* from Berber, we would expect the Berber article to be a "step ahead" in the grammaticalization process and thus have a wider semantic range, which at least provisionally does not seem to be the case. In turn, one could dispute the notion that the development of *waḥad* represents a unidirectional process of MA copying Berber. On the other hand, one might argue that pervasive bilingualism of that sort we find in Morocco essentially neutralizes any such lag, thus rendering its presence or absence an ineffective metric. Future work may be able to more effectively ascertain whether previous work has been justified in claiming MA < Berber grammatical copying, or whether it has prematurely rushed to conclusions.

Regardless of the questions surrounding *waḥad*'s relationship with its Berber equivalent, we can nonetheless still sketch a reasonably viable account of the article's development. I thus summarize what I see as the most likely account for the history of *waḥad* as follows: it first began to develop out of the numeral 'one' only sometime after Arabs arrived to North Africa and their dialects became the target of L2 learning for native Berber speakers, meaning the semantic erosion of the definite article \*/l-/ had already begun to occur. From this point the specific-indefinite article *waḥad* was then simultaneously grammaticalized via a typical cross-linguistic cline, either on the model of or alongside the Berber articles *yan/yat*. Today this process continues as the form shows initial signs of spread to non-specific indefinite contexts, and if the trend remains in place one might predict that *waḥad* will continue to follow the same cline, ultimately coming to represent a fully indefinite article that is able to express any non-referential entity.

#### **4.2.3. Disfixed Form {- /l-/} (etymological \*Ø)**

If the major diachronic change that has led to the rearrangement of MA's definiteness system is \*/l-/ > Ø, where the former definite article was disassociated with definiteness and instead became lexicalized, the development of the former unmarked form into a phonologically conditioned disfix that is syntactically constrained can be

envisioned as a complementary process  $*\emptyset > \{- /l/\}$ . Since we have already looked at the linguistic developments that likely played into the semantic extension of  $*/l-/$  into indefinite tiers of the semantic hierarchy, it is not difficult to envision the pressures that would have led to the loss of  $*\emptyset$  in the same environments. Once forms that had been diachronically associated with definiteness lost this connotation and were no longer semantically restricted from representing indefinite meanings (*l-kas* ‘the cup’ > ‘cup’), it is only natural that the form  $*\emptyset$  would have lost its primary role as a marker of indefiniteness because it no longer had exclusive domain over this semantic range. In this light, what is perhaps a more intriguing question is why forms analogous to the Old Arabic  $*\emptyset$  would have been maintained at all, leaving disfixation via  $\{- /l/\}$  a productive process.

My response to this question is that while the historical group learning that seems to have been at least partially responsible for the observed semantic changes in MA was imperfect, it was still remarkably precise on the whole. Even if L2 speakers of MA – modeling the semantic implications of MA nominal forms on their Berber L1 – unwarily extended definite nouns into indefinite contexts, they never would have had to actually produce ungrammatical forms, but rather only semantically inappropriate ones. To give an explicit example, that speaker might say *bya l-wəld* ‘he wants a son’ at a time when  $*/l-/$  still implied definiteness might have struck a native Arabic speaker as contextually inappropriate, but there could be no denial that the form itself was acceptable (albeit with the meaning ‘he wants the son’). I relatedly see the maintenance of a disfixed form  $\{- /l/\}$  as the by-product of pre-existing Arabic constructions that grammatically required a nominal form  $*\emptyset$ . Because forms like *\*ašmən l-bənt* ‘which the girl’ would have been ungrammatical and never produced by the native speakers upon whom L2 learners would have sought to model their speech, L2 speakers with a competent command of the L1’s grammar (but not semantics) would have found a way to relate them to what they saw as the unmarked nominal form  $*/l-/$ , thus establishing the productivity of disfixation. That  $\{- /l-/$  subsequently developed an association with quantification, then, probably reflects the generalization that the dominant grammatical constructions that explicitly required



historical \*Ø to the exclusion of all other possibilities (i.e. numbers and certain particles) already reflected this element. Once this relationship was established, it could then be employed with more nuance to indicate quantification contrasts (*l-kas* ‘a cup,’ *kas* ‘a single cup’) rather than indefiniteness as it formerly had.

The maintenance of {- /l/} in the generic predicate position is more difficult to explain, particularly because the act of extending \*/l-/ into this semantic domain (*huwa wəld* ‘he is a boy’ > \**huwa l-wəld*) would not have been subject to the same grammatical restrictions we see with quantification. It may even be possible that at some point in the past, some MA L2 speakers and their descendants regularly did use the unmarked form Ø for generic predicates, although if this were the case it would have had to have since been reversed in most MA dialects. What I instead see as a possibility is that the predicate’s semantic overlap with adjectives (c.f. *huwa mʕəlləm* ‘he is a knowledgeable [person]’ and ‘he is knowledgeable’) somehow allowed it a special status where it continued to follow the Old Arabic pattern for definiteness marking, with \*/l-/ indicating definiteness and its absence indefiniteness. This, as we have seen with indefinite forms such as *l-ma sxun* ‘hot water’ and *l-ʕərs ʕadi* ‘a normal wedding,’ does indeed seem to be the case for adjectives. Why definiteness would still be marked for the MA attribute but not the noun is a question that poses its own historical problems, and one that will be critical for future research on MA to address (see 5.2).

#### 4.2.4. Particle *ši*

The particle *ši*, an existential quantifier and marker of ‘referentiality,’ poses many of the same questions that does *waḥəd*, and the generalizations we can posit about its history follow the same logic but suggest different conclusions. Like for *waḥəd*, the etymology of MA’s *ši* is quite transparent, with the particle derived from Old Arabic *šay?* ‘thing’ and etymologically related to a modern *ši*, of the same meaning, that is found in a number of currently extant dialects. In addition, *ši* plays an apparently congruent article function that marks referential states in at least one other major variety Arabic. A Syrian example from Brustad is:

- (133) *lāzim naʔmil lu šī muqaddime la-ḥatta mā yinšidim* ‘we must arrange some kind of preparation for him so that he won’t be shocked’ (Brustad, 2000, p. 27)

The semantic function of *šī* in Syrian clearly parallels that of its Moroccan counterpart, where the speaker is committed to the existence of an individual entity but is not able to uniquely identify it. In addition, unlike with *waḥad* and its cognates in other Arabic dialects, the grammatical requirements of *šī* are similar in both dialects in that they necessarily exclude the phonological material /(V)l-/, whether the latter is interpreted as a definite article (as in Syrian) or an internal component of the lexical stem (as in Moroccan). In turn, the similarity of these colloquial forms again raises the question of whether they more likely stem from a common origin or arose independently in their respective dialect groups. Brustad (2000, p. 43) is of the former opinion, stating that “the fact that Moroccan and Syrian dialects share the article [*šī*] ... suggests that these articles have fulfilled this [indefinite] function for a very long time.”

My own analysis supports this notion, particularly if we assume that the shift \*/l-/ > Ø was an early post-migration development in North Africa sped along by second-language learners adopting Arabic. Since *šī* necessarily requires what is now the disfixed form {- /l-/} in MA, corresponding with the historical indefinite form \*Ø, it is only logical that the particle would have already been present in the dialect before the former definite article \*/l-/ began to be extended to indefinite statuses; otherwise we would expect forms like \**šī l-wəld* ‘\*some boy’ along the lines of *waḥad l-wəld* ‘[indef.] this boy.’ That such hypothetical forms are not generally attested in MA, then, suggests that *šī* would most likely have been grammaticalized before the predecessors of MA ever even reached the Maghreb. This, subsequently, would give *šī* a plausible origin in the Arabian Peninsula. The semantic and formal similarity of the same particle in Syrian Arabic, despite its geographic distance from Morocco, lends credence to this view on one count. In addition, a Saudi informant of mine indicates that that *šī* is used as an article in the Jazan area of southern Saudi Arabia (as in *šī wāḥid* ‘someone’), which would give us

direct evidence of a form *ši* in its proposed area of origin and even stronger cause to posit an original development there.

If the origins of *ši* do indeed date back to the early or pre-Islamic dialects of the Arabian Peninsula and it is truly co-identifiable with etymologically similar particles in other dialects, the form and its semantics would seem to have shown a remarkable amount of stability over the centuries. I am unable to speculate as to why this may be the case, especially when other historical forms have seen such drastic semantic reshuffling over the course of the same time period. Even in contemporary terms, however, there seems to be little indication that *ši* is either losing productivity or yielding much of its semantic scope to the unmarked form  $\emptyset$ , which can (as in examples 54 and 55) but only infrequently does represent ‘referential’ meanings. If anything, the expressiveness of *ši* may actually be expanding (as in examples 129-130, 4.1.4) where it can be seen to represent solely ‘type identifiable’ entities. Partly responsible for this shift may be the routine use of *ši* in politeness routines, where speakers underspecify an object they desire out of deference to the speaker:

(134)  $\text{ʕndək } \text{ši } \text{xūbz?}$  ‘Do you have [some sort of] bread?’ (Caubet, 1983, p. 141) > ‘Do you have bread?’

Formally the construction might thus represent its typical ‘referential’ connotation, but functionally it is equivalent to any other marker of ‘type identifiable.’ I see it as a possibility that such routines, upon becoming formularized rather than explicitly selected for by the speaker, could act as a vehicle for semantic extension and replacement of other forms. Even if such change is under way, however, it has not happened in full, so that the expressions  $\text{ʕndək } \text{ši } \text{xūbz?}$  and  $\text{ʕndək } \text{l-xūbz?}$  ‘do you have bread?’ essentially remain synonymous.

## 5. Conclusions

This study has aimed to recast the current discussion of definiteness marking in Moroccan Arabic by challenging traditional views on how definiteness and givenness are expressed in the dialect, providing a new synchronic description to account for what previous models cannot, and supporting this description with a provisional diachronic look at how it may have come about. Throughout it I have used unelicited, contextualized linguistic data from MA native speakers and early bilinguals to both evaluate the theoretical claims of previous authors and support my own alternative views. Among the most essential of these views is that MA has undergone article loss, a process via which the form \*/l-/ – which once represented a definite article – lost its association with definiteness, was extended into indefinite semantic contexts, and became a lexically-determined component of the new unmarked form of the noun (4.2.1). This diachronic shift has, in turn, determined the semantics of other MA nominal forms, so that the historically indefinite \*Ø has been reanalyzed as a disfixed form {- /l-} that is derived via subtractive morphology (4.2.2) while an indefinite article *waḥad* has been developed (4.2.3) and an older particle *ši* maintained (4.2.4). I have likewise argued, among other things, that this semantic shuffling is not only explained by MA's linguistic history, which shows both internal and external impetuses for such change, but could perhaps even be expected on the basis of it.

The ideas presented in this study are certainly nowhere near conclusive, and deserve critique and refinement just as do those of previous work. If even a few of the precepts I have outlined in this study are to be taken as valid, however, they should have immediate implications for a number of different fields to which the study relates. In addition, the study opens the door for a variety of other research questions in the future, the answers to which could be useful not only for the study of Moroccan Arabic but for the discipline of linguistics as a whole. In what follows I briefly outline these current implications and future directions, and then register my final remarks.

### 5.1. IMPLICATIONS FOR PREVIOUS WORK

Perhaps the most obvious impact the conclusions of this study will have, should they be accepted, is on dialectological descriptions of Moroccan Arabic. As I have shown, up until this point nearly all descriptions of the dialect have treated the forms \*/l-/ (my Ø) and \*Ø (my {- /l-/}) as semantically analogous to similar forms in Classical Arabic and other dialects, where they represent definiteness and indefiniteness respectively. In this study I have challenged this description, providing evidence that \*/l-/ is not exclusively associated with definiteness but can instead occur for all semantic statuses on the Givenness Hierarchy (3.2.1); I have likewise argued that there is an open and historically productive class of nouns for which \*/l-/ is never present even when semantically definite (3.2.2). In this light, future descriptions should no longer refer to \*/l-/ as a “definite article” unless they can produce convincing counter-evidence to these conclusions, as to do otherwise would be a disservice to an accurate description of the dialect. Similarly, where past studies have referred to \*Ø (my disfixed form {- /l-/}) as “bare” or “unmarked” we should take a critical view of these claims, asking what synchronic evidence would support this assumption and, in particular, how speakers would derive other forms from it when there seems to be no consistent rule for doing so.

I have made these claims and put forth an alternative model, that said, not merely to critique previous descriptions, but with the primary goal in mind of answering a number of fundamental questions that they have already raised. My newly proposed synchronic description of MA’s definiteness system settles the debate over why an indefinite article *waḥad* would often but not always require a “definite article” \*/l-/, namely because the latter is not actually an article but rather a lexicalized component of a majority of MA nouns. It also explains the semantic patterns behind what would otherwise seem to be an apparently erratic distribution of forms \*/l-/ and \*Ø for non-referential entities by giving the latter as a marker of quantification rather than indefiniteness. Finally, the description given here allows us, for the first time, to conceptualize the cognitive processes that would allow speakers to construct semantically appropriate forms for *all* MA nouns using the same morphosyntactic derivation

processes, regardless of whether or not they occur with the element \*/l-/, because definiteness in MA is given as zero-marked rather than the product of additive morphology, as it has been viewed in the past.

The currently proposed model, in turn, has residual benefits for studies that have taken the accuracy of previous descriptive models for granted. One field in which these benefits could be particularly resonant is that of bilingualism, where studies of code-switching between Moroccan Arabic and other (primarily European) languages have made significant contributions to the field's general theoretical outlook. Bentahila and Davies's (1983) pivotal study of Moroccan Arabic-French code-switching, as an example, makes the characteristic claim that "the indefinite [*waḥad*] and the demonstratives [*had*], [*dak*] must... be followed by a definite article [/l-/], so that NPs containing these have a sequence of two determiners, as in [*waḥad l-bənt*], literally 'one the girl' and [*had l-ma*], literally 'this the water.'" Although my model would thoroughly dispute the view that \*/l-/ in such contexts is either a definite article or required, in Bentahila and Davies's case the authors use it as a basis for grammaticality judgments by claiming that code-switched nouns will also require determiners in the same constructions, as in:

(135) *haduk les gens* 'those people' (Bentahila & Davies, 1983)

(136) *dak la chemise* 'that shirt' (ibid.)

(137) *waḥad le liquide* 'a liquid' (ibid.)

This view apparently suffices to explain Bentahila and Davies's data, but does not necessarily do so for the findings of others who seem to have internalized the same claims about how definiteness is expressed in MA. Nortier (1990, 1995) and Boumans (1998), for example, have both found examples of code-switched Dutch insertions into MA speech that include neither a Dutch determiner nor MA \*/l-/ where they believe one or the other would be "obligatory." Some examples from Nortier include:

(138) *waḥad gesprek* 'a conversation' (Nortier, 1995)

(139) *dik cultuur* ‘that culture’ (ibid.)

(140) *wahed bejaardencentrum* ‘an old people’s home’ (Nortier, 1990)

Because the authors have been under the impression that a “definite article” \*/l-/ is necessary in such contexts, they have in turn labeled such constructions as ungrammatical or “in violation” of MA syntax. In response, the same authors have labored to explain why they nonetheless occur, with Nortier (1995, p. 88) suggesting that the appearance of “null articles” is mostly likely a “suspension of grammar” unique to the phenomenon of code-switching. This explanation has, in turn, been enthusiastically picked up by scholars (such as Myers-Scotton, 2002; Owens, 2005) who see it as evidence of how processing difficulties in code-switching impede full implementation of either code-switched languages’ grammar, and who have incorporated into their own models.

The description I have presented in this study, however, would suggest the opposite. Since I do not see \*/l-/ as an article at all but rather a lexicalized component of the unmarked noun, the fact that a borrowed or code-switched noun would not display it is not at all surprising, and structures in which it does not appear would certainly need not be considered ungrammatical. That some code-switched nouns do show the source language determiner and some do not, then, can be viewed as a question of borrowing routines rather than one of grammatical requirements, the primary variable being which source form a given bilingual community has habitually taken as congruent to the unmarked MA noun.<sup>22</sup> Subsequently, these routines could be susceptible to variation from speaker to speaker or community to community. This notion is supported by the fact that even for French, where code-switched nouns typically do include the source

---

<sup>22</sup> In-depth speculation as to why different borrowing routines may be at play for different language pairings is beyond the scope of this study, but future research might explore the role of phonological similarity between \*/l-/ and the French determiners *le, la, les* as opposed to Dutch *de, het*, as well as the geographic and ethnic origins of the migrant communities: an estimated 60-80% (Gazzah, 2008, p. 12) of the Dutch Moroccan population is of Rif Berber origin and speakers may have modeled Dutch > MA routines on pre-established Berber > MA ones, where there is no determiner to be borrowed.

language article (as in 135-137), Post (2013) has found a number of French nouns with “null articles” in online written code-switching, such as:

(141) <*kayen chi had ygoli wach had taman ghadi yahbet f mois de septembre*> ‘is there anyone who can tell me whether or not this price will go down in the month of September?’ (Post, 2013)

(142) <*had video machi sehih rah machi police dyal almaghreb rah police dyal algerie*> ‘this video isn’t correct, that’s not the Moroccan police, that’s the Algerian police’ (ibid.)

More research will be needed to substantiate which factors may be at play in the selection of borrowing routines from community to community. Nonetheless, if the synchronic description of MA I have given in this study is in fact correct, various instances of code-switching that have previously been considered aberrant can actually be shown to be fully grammatical under the rules of monolingual MA syntax. In turn, they would no longer serve as useful for the theoretical argument that the cognitive demands of code-switching can result in a departure from the preexisting grammars of the code-switched languages. While there are certainly examples from other language pairings that scholars might use to maintain this claim, that Moroccan has played such a vital role in previous iterations of the theory could be a setback for its champions, at least temporarily.

Another type of previous scholarship that this study is set to impact is that on language contact, particularly as it relates to Moroccan Arabic and Berber. There has been a good deal of writing on the apparent substratal influence of Berber on MA (Chafik, 1999; Chtatou, 1997; El Aissati, 2006) and evidence of the two languages’ ongoing convergence (Maas & Procházka, 2012; Maas, 2000; Tilmatine, 2011), but the notion of definiteness has rarely played a particularly prominent role in these arguments for contact. Where it has, the focus has generally been on the syntactic similarities of the indefinite MA and Berber articles *wahād* and *yan/yat*, without a truly in-depth discussion of how their semantics overlap and what this may imply for the directionality of contact effects, which is typically assumed to be Berber > Arabic (though I see a process of co-



grammaticalization as equally likely). What has never been explored, however, is the notion that the entire MA definiteness system may have converged toward the Berber one, which is in essence what I argue for by saying that MA has lost the definite article \*/l-/ and begun to treat semantically definite forms as identical to the unmarked nominal form. This would indicate semantic and structural convergence on a truly wide scale and be a much more compelling argument for the role of contact than is the existence of a single indefinite article following a common cross-linguistic grammaticalization cline.

While the potential value of the findings to those who deal with Moroccan Arabic directly are thus fairly clear, the study could likewise be useful for linguists in general. In particular, it functions as a case study of how one dialect group can diverge from related varieties as it undergoes semantic restructuring and a related reanalysis of inherited morphophonological material. For those working on language contact, it tentatively corroborates many of the assumptions voiced in prominent works such as Thomason & Kaufman (1988), particularly regarding how cases of language shift and group learning can lead to semantic and morphosyntactic changes in the adopted language. The study also has valuable implications for grammaticalization models insofar as Moroccan Arabic, as described here, would represent a language that has passed through all of the stages of the full grammaticalization cline from demonstrative to article loss. Whereas there are certainly other languages that can be presumed to have done the same, MA's clear advantage for such research lies in the fact that we have fairly substantial historical records for both Arabic and Semitic that might help us map these stages to approximate historical time periods, and perhaps even social developments.

## **5.2. FUTURE DIRECTIONS**

While we can cite a number of immediate implications that the current study may have, it likewise raises its own unique set of questions and demands, some of which I have already briefly mentioned or hinted at. One of the biggest that I have not explicitly discussed, however, is that of whether Moroccan Arabic is truly unique in having undergone loss of the definite article and the semantic rearrangements it has entailed. As

we saw in 2.3.3, all of the other major Arabic dialects appear to have maintained /(V)l-/ as a true definite article, but this should not lead us to the immediate conclusion that there are not at least some other dialects that have also seen article loss, particularly when there are a great number of peripheral Arabic varieties that have not had the benefit of the attention given to Moroccan (for which the phenomenon still had not been explicitly recognized until now). In fact, there is already agreement that some dialect groups, such as Central Asian Arabic (Jastrow, 2005; Ratcliffe, 2005), have lost the definite article /(V)l-/, a conclusion that has perhaps been more easily drawn because the morphophonological material historically associated with the article simply is not present in these varieties. I again contend that what has made recognition of article loss in MA so difficult for scholars is that this original phonological material has been maintained, contributing to a surface illusion that it is representative of definiteness. That the same issue may be at play in the case of other, lesser-studied Arabic dialects is certainly a possibility.

Looking beyond Arabic, that said, there may already be established parallels to Moroccan Arabic among other related languages of the Semitic family. The most fascinating and relatively well-documented of these is Eastern Late Aramaic and its descendants, in which a historical definite article suffix *\*-ā* became lexicalized as part of the noun and lost its association with definiteness (Kaufman, 1974; Pat-El, 2010; Rubin, 2005).<sup>23</sup> This development closely resembles what I have here proposed for Moroccan Arabic, where a definite prefix *\*/l-/* has ceased to mark a ‘definite’ status and has instead come to represent the unmarked nominal form. The details of how this process has been realized in Late Aramaic reveal even more striking similarities to the MA case: like in MA, the only historically indefinite forms that are preserved in these varieties are syntactically conditioned, appearing in positions “such as following numbers and the words *kul* ‘every’ and *dlā* ‘without,’” as well as for predicates (Rubin, 2005, p. 86). That we could cite two geographically separated Semitic languages that have independently undergone a process of article loss but nonetheless seen remarkably analogous outcomes

---

<sup>23</sup> An example from Rubin (2005, p. 68) is *baytā* ‘the house’ > ‘house, a house, the house.’

would be an enormous boon to the argument that these outcomes may be typologically predictable. In addition, further examination of the social conditions that may have accelerated this process for either of the language groups in question could shine light on those of the other.

Even without a great deal of additional investigation, the Eastern Late Aramaic case sparks some immediate questions for future research on MA. It has been relatively well established, for example, that following the loss of the historical definite article *-ā*, a number of Neo-Aramaic dialects actually re-developed a definite article from their demonstratives (Pat-El, 2010; Rubin, 2005). Whether the same might be a possibility for MA is intriguing, and it would doubtlessly be an exciting development to be able to watch the grammaticalization of a new definite article in progress were it indeed under way. Although in this particular study I have paid little attention to use of demonstratives, my initial observation is that at least in some cases MA demonstratives are used for semantic distinctions that would typically be made with a definite article. In the following discourse, for example, the speaker uses a demonstrative *dak* to differentiate the first mention of a non-referential object (143) from a second reference to it later (144):

(143) *ka-yžib l-luz, ka-yžib l-gərgaʕ, ka-yžib t-tmər, w ka-yžib ħuta* ‘He brings almonds, brings walnuts, brings dates, and brings a single fish’ (9-25)

(144) *dak l-luz, w dak l-gərgaʕ, w t-tmər lli žab, ka-nfərrqu-h ʕla q-ḏyaf, ta-yakʷəlu, ka-yḏəħku* ‘those almonds, and those walnuts and the dates that he brought – they distribute them among the guests, and they eat and they laugh’ (9-35)

The above example can be seen as anaphoric, but the demonstrative in it introduces a critical semantic distinction for a form that could otherwise indicate either a definite or indefinite entity. In another text, a different speaker uses a demonstrative *duk* to refer to an entity that has not even yet been introduced, but that could theoretically be deduced from the context:

(145) *hna kayn s-suq l-ʔusbuʕi... kayn s-suq l-ʔusbuʕi, ka-nddiu z-zərbiya l s-suq l-ʔusbuʕi, ka-yšriu-ha šndna duk l-mūhtəkirin, w šafi* ‘There’s a weekly market here – we take the carpet to the weekly market, and those monopolizing [middlemen] buy it from us, and that’s it’ (1-17)

What we might identify in such an example is the initial incursion of MA demonstratives into Heine’s “contextually definite” grammaticalization stage. Further analysis of MA speech may give us more such examples and clues as to the typological direction in which the dialect may be heading. The same analysis would be useful for Moroccan Berber varieties, in which I expect we would find semantic patterning for demonstratives that resembles their use in MA, much in the same way that we seem to with the indefinite *yan/yat* (4.2.2). The semantic contexts for demonstrative use thus remain a promising target for future research.

In addition to demonstratives, the semantic connotations of a number of other MA forms seem to be susceptible to future change and researchers would do well to monitor them. We have already seen that *wahəd*, for example, displays some preliminary signs of being grammaticalized into a general article that can signify not only ‘specific’ statuses but indefinite ones in general (4.2.2); future studies with a larger set of data may be able to better quantify how far this process has actually progressed. The particle *ši*, similarly, has been shown to have some overlap with other forms for representing ‘type identifiable’ statuses (4.2.4). Whether the usage of *ši* in this semantic sense is increasing or otherwise is a question that future work might ask, perhaps confirming or refuting my speculation that such a change might be driven by politeness routines. Finally, there is the question of how increasingly literacy in and familiarity with Modern Standard Arabic may impact MA speakers’ cognitive associations between certain forms and semantic notions. My impression is that this factor actually has little effect on the realization of core lexical items, but more detailed studies may be able to show otherwise.

Another unanswered question that has appeared at least a few times in this study is the notion that, while there is strong evidence that \*/l-/ has lost its association with definiteness for nouns, this association may have been maintained (or reestablished) as a feature of the nominal attribute (4.2.3). This possibility is reflected in constructions such

as *l-ma tayb* ‘boiling water,’ which would theoretically contrast with *l-ma t-tayb* ‘the boiling water,’ as well as in the fact that generic predicates do not typically show \*/l-/. I do not currently, however, have sufficient data to draw conclusions. A future study that focuses exclusively on this problem may be able to sort out exactly how consistent this form-meaning correlation is and look into questions of how it may have arisen. If it is in fact accurate to say that definiteness is marked for the attribute in MA, it could pose yet another interesting tie-in to the history of Semitic languages as a whole. Specifically, Patel (2009) has argued that original Semitic definite article arose via a similar pattern, marking the attribute but not the head noun; showing that this pattern is found in Moroccan synchronically would be strengthen the case that it could also have been viable in Semitic historically.

A final avenue for future work relates to the very theoretical frameworks that scholars use to conceptualize definiteness and givenness. Throughout this study I have made ad hoc use of two different models that embody many of the same concepts, namely Givón’s (1976) wheel diagram and Gundel et al.’s (1993) Givenness Hierarchy. Both of these models have advantages and weaknesses; it is precisely for this reason that I have used them both as seemed appropriate, and tried to map them onto each other where possible. Gundel et al.’s hierarchy, for example, encodes implicational relationships between definiteness, referentiality and type identifiability in a way that Givón’s wheel cannot, and better contextualizes the entire range of givenness from pronouns to fully indistinct entities. Givón’s wheel, on the other hand, makes a crucial distinction between generic and non-generic entities that is not found in the Gundel et al. model, and likewise differentiates what I have here called ‘specific’ entities from those that are ‘referential’ but non-specific. Ideally, a future study will be able to combine the advantages of both of these previous models into a unified framework for representing givenness, perhaps even one that interfaces it with recent work on grammaticalization by outlining “pathways” along which new forms would tend to develop and expand their semantic scope.

### 5.3. CLOSING REMARKS

What has been perhaps the most difficult part of making this study's central claims has been doing so with the full knowledge that they run contrary to nearly every description ever published of how definiteness is realized in MA, including those of authors whose work on the dialect is outstanding and for whom I have immense respect. In many ways it seems counterintuitive that definiteness marking – seemingly a fundamental and transparent component of a language's grammar – would not have been adequately described in one of the more widely spoken Arabic dialects, especially one that has enjoyed the attention of so much scholarly work over the course of the last century. Despite the cause for hesitation, I have tried in this study to lay out not only a new description of how definiteness is marked (or, perhaps more appropriately, unmarked) in Moroccan Arabic and an explanation of how this description is diachronically plausible, but also a sense of the reasons why previous scholars may never have even considered the description I present here as a possibility (see 2.3.1). What I ultimately see in authors' continued insistence on calling \*/l-/ a definite article, even when they themselves have documented many of the ways in which it does not actually correlate with definiteness, is the residual effect of standard language ideology (Milroy, 2001), the full chains of which the field of Arabic dialectology has yet to cast off.

In this light, it seems to be little coincidence that the MA forms \*/l-/ and \*Ø, both of which have apparent equivalents in Classical and Modern Standard Arabic, are the ones that have been most drastically misrepresented in past descriptions. Because /(a)l-/ is a true definite article in the standard language, scholars seem to have approached MA with the same assumptions, writing off or explaining around the many instances in which MA's etymologically equivalent \*/l-/ co-occurs with indefinite meanings or fails to mark definite entities. Similarly, because the prefixless nominal form in the standard language is the only one that can indicate an indefinite status, the same researchers seemed to have missed that it is not actually the primary means for representing indefinite entities, and instead typically only occurs for them under certain syntactic conditions. Inversely, however, the semantic descriptions for the two “colloquial” articles *wahād* and *ši* appear

to have been remarkably accurate, perhaps because their very absence in the standard language has allowed scholars to approach them unbridled by the same sort of *a priori* assumptions.

The lesson that we should perhaps take from this discussion is that it is not only linguistic form that should inform our judgments, but linguistic function as well. As mentioned in the introduction to this study, a safe assumption to make is that language is always in the process of change, and these changes can appear any number of domains, often even those in which we may not be actively seeking them. In the case of a language group with as many speakers as Arabic – many of whom are spread across thousands of different communities that have had only limited contact with each other for centuries – it should not be surprising to find grammatical systems that diverge from the normative model, and it should be even less surprising to find that these systems came about via cross-linguistically common processes. In essence, variation should be one of the first things we look for in Arabic dialects, and the expression of definiteness in Moroccan is only one of many places in which we can likely expect to find it.

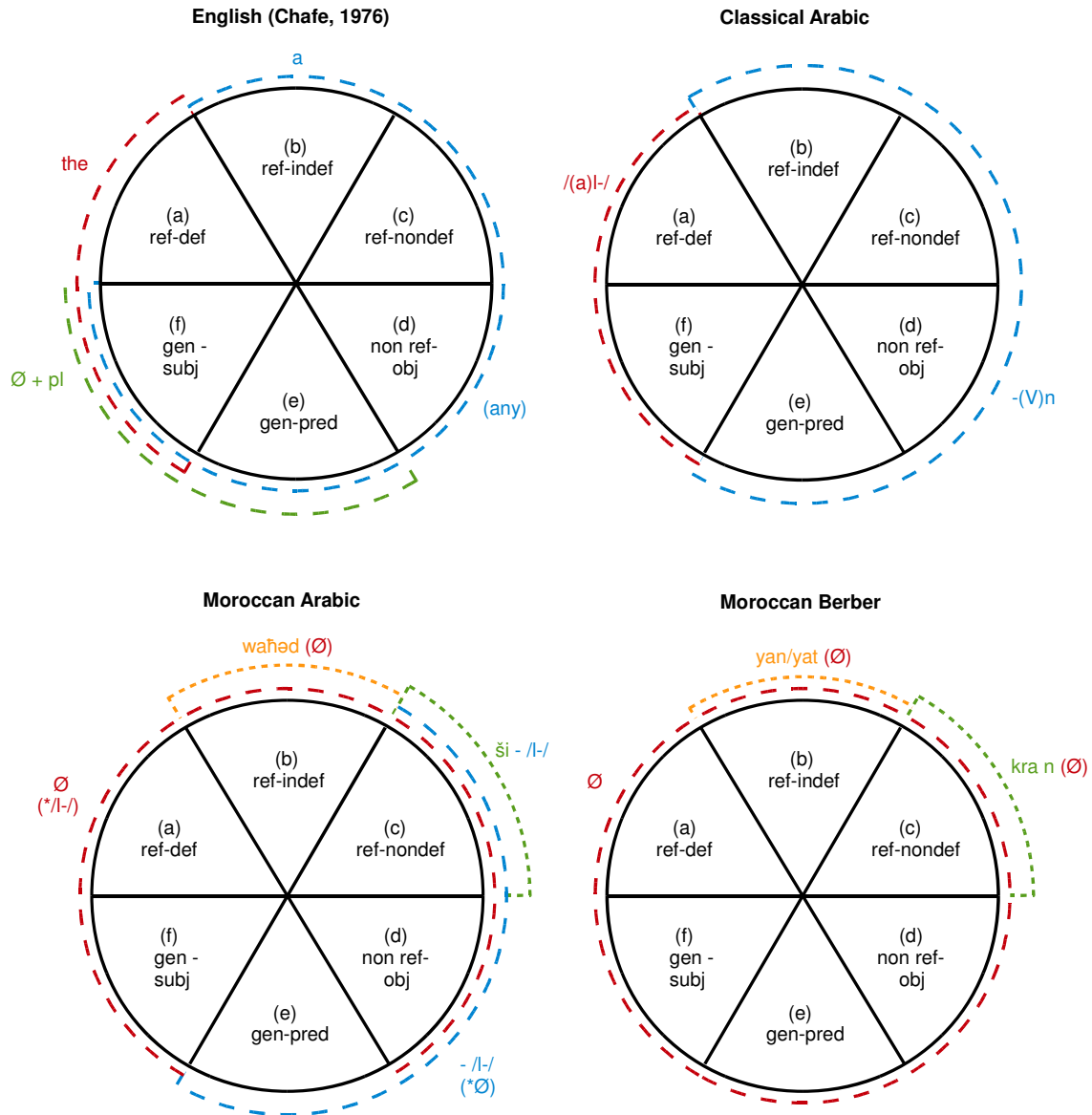
## Appendices

### A. GIVÓN'S (1976) WHEEL DIAGRAM FOR RELEVANT LANGUAGES

The diagrams below are provided to facilitate the process of visualizing how semantic notions of givenness are formally represented in the various languages discussed in this study.

*Equivalencies for (Modified) Givenness Hierarchy:*

(a) Definite (b) Specific (c) Referential (d) Type Identifiable (e) Type Identifiable\* (f) Definite\*





## B. IN-TEXT OCCURRENCES OF VARIOUS MA NOMINAL FORMS

The following lists are provided for easy reference to some of the more infrequently-attested nominal forms in the included Moroccan Arabic texts. Unmarked nouns that display \*/l-/ in their etymology, which represent a large majority of nouns in the texts (500+ items), do not appear here.

### Occurrences of *wahəd*

3-1	wahəd l-wəşfa
3-2	wahəd l-biḍa
4-12	wahəd kayiṭ
4-13	wahəd l-qənba
4-4	wahəd l-ʔinaʔ
7-12	wahd n-nuʃ
7-17	wahd l-ʃəya
7-18	wahd l-ʃəya
8-14	wahəd l-m <sup>w</sup> adda
8-18	wahəd l-m <sup>w</sup> adda
8-37	wahəd s-sʃada
8-9	wahd t-təydiya
9-15	wahəd t-təqlid
9-8	wahəd l-kürsi
9-8	wahəd ʃ-ʃala

### Occurrences of *ši* {- /l-/}

1-22	ši mihražan
1-22	ši maʃriḍ
1-24	ši haža
2-16	ši haža
4-12	ši kayiṭ
4-16	ši baraka
5-7	ši aʃšab
7-21	ši wahdin
7-21	ši haža
7-25	ši wahdin
7-26	ši wahdin
7-27	ši haža
7-6	ši ʔažin
8-31	ši blaša
8-31	ši məʃməl
8-39	ši haža
9-4	ši hədd

### Occurrences of {- /l-/}

1-11	ṭariqa	7-36	taqalid
1-12	mwadd	7-36	wahd
1-18	žamʃiyat	7-4	wahd
1-18	tʃawniyat	8-1	təlmid
1-23	mwasim	8-1	ṭalib
1-5	saʃa	8-12	swayʃ
2-11	wahəd	8-17	m <sup>w</sup> adda
3-1	ʃibara	8-19	yyam
3-1	hərša	8-20	yyam
3-2	kiļu	8-20	wahəd
3-2	kas	8-22	simana
4-16	drahəm	8-28	yum
4-16	dərħəm	8-29	yyam
4-16	dərħəm	8-32	wahəd
4-2	ʔanžiyat	8-32	m <sup>x</sup> əddmat
4-8	kas	8-34	wahəd
4-9	ṛəbṭa	8-34	wahəd
4-9	ṛəbṭa	8-39	muʃkil
5-4	həbub	8-4	sana
5-7	wahəd	8-6	sana
5-7	haža	8-6	mwadd
6-2	dqayq	8-7	šhər
6-3	xəṭrat	9-1	kiłumiṭər
6-4	xəṭrat	9-2	yyam
7-12	wahd	9-25	huta
7-16	wahd	9-28	haža
7-17	wahd	9-29	ʃəša
7-27	wahd	9-29	ʃəša
7-27	wahd	9-31	haža
7-29	shur	9-38	qəʃʃa
7-29	wahd	9-43	məra
7-3	təwqit	9-46	qʃaʃi
7-3	mdina	9-47	yyam
7-30	wahd		
7-33	mdina		
7-36	ixtilaf		

### Occurrences of nouns for which \*/l-/ is not present in etymology

1-15	dak aṣṭa	5-1	atay	7-14	arxṣiṣ
1-15	aṣṭa	5-10	atay	7-16	atay
1-8	aṣṭud	5-3	atay	7-4	atay
1-8	tarubiya	5-4	atay	9-15	taḥəmmamt
1-8	assay	5-5	atay	9-16	taḥmmamt
2-5	bṣṭila	5-6	atay	9-45	kʷəṣksu
2-5	disir	6-6	hadak atay	9-8	bnadm
4-12	kayit	7-13	wahd		

### Occurrences of nouns with attached possessive pronouns

2-18	raḏəlha	9-21	raḏəlha	9-31	mnṭəqha
2-6	ḏarhūm	9-22	raḏəlha	9-31	ḏlalha
2-7	bbʷaha	9-23	mmʷha	9-39	gdəmhə
2-7	mmʷha	9-23	mmʷu	9-43	ḏarhūm
7-19	ṣḥabk	9-24	mmʷu	9-44	ḏarhūm
8-21	ṣḥabu	9-24	mmʷha		

### Occurrences of nouns as first element of Semitic construct state

1-20	tswiq	6-1	atay	8-40	had ṣḥadət
1-29	ḥəqq	7-14	mmʷalin	8-41	ṣḥadət
1-8	wəraq	8-3	ṣḥadət	9-38	xu
1-8	wəraq	8-30	ṣḥadət	9-44	qəṣṣət
2-6	ḏar	8-35	ṣḥadət	9-45	qəṣṣət
4-10	ras	8-36	ṣḥadət	9-7	ṣaʔilət
5-5	ḥbub	8-40	qimət	9-7	qrab

### C. MOROCCAN ARABIC TEXTS

The oral texts below were collected by the author and represent the primary source for the linguistic data in this survey. They are transcribed directly from the digital recordings and no intentional editing has been performed; as such, pauses and false starts are frequently represented. The texts are included to allow readers to independently verify the discursive context of examples given in the study.

#### Text 1: Carpet-Making in Taznakht

*female, Berber, 28, high-school education, Taznakht area*

##### Text

(1) bəsmilla r-rəḥman r-rəḥim... ana daba yaḍi nhdər ʕla l-mraḥil lli ka-tmərr mnha z-zərbiya – z-zərbiya l-wawazgitiya, aw z-zərbiya t-taznaxtiya. (2) l-mərḥəla l-ʔəwla hiya l-mərḥəla dyal d-dəzzan yəʕni ka-ydəzz... ka-ndəzzu l-ynəm, w had l-mərḥəla ka-yqum biha r-ražəl. (3) mn bʕd ka-n... ka-nhəzzu ʕ-ʕufa, ka-nddiu-ha l l-wad, ka-nšəbbənu-ha, mn bʕd ka-nnšru-ha htta... htta tnšəf. (4) w mn bʕd ka-nnqqiu-ha mn ʕ-ʕuk, w mn bʕd ka-nqəršəlu-ha, w mn bʕd l-qəršəl ka... ka-nyəzlu-ha, w mn bʕd l-ʔzil ka-tntaqəl l-mərḥəla dyal t-txmir. (5) w had l-mərḥəla ka-tstəyərɔ ərɔəf w ʕšrin... rɔəfa w ʕšrin safa, w had l-mərḥəla dyal t-txmir ka-nxəmməru-ha f... f ʕ-ʕəbba – ʕ-ʕəbba w l-ma. (6) w mn bʕd ka-tntaqəl l-mərḥəla dyal ʕ-ʕibaya. (7) ʕ-ʕibaya, kayna ʕ-ʕibaya t-ṭabiʕiya w ʕ-ʕibaya l-kimawiya. (8) ʕ-ʕibaya t-ṭabiʕiya matalan ka-nšəbyu b l-aʕšab t-ṭabiʕiya bḥal aʕfud, tarubiya, wəraq t-tin, wəraq l-gərgaʕ, l-ḥənna, assay... kaynin bzzaf dyal... z-zəʕfran, (9) kaynin bzzaf dyal l-aʕšab walakin daba hit kayn ʕ-ʕafaf sʕib bzzaf baʕ tləqa duk l-aʕšab hit ka-ykunu f ʕ-ʕbəl, (10) ma... ma-kaynin-ʕ bzzaf daba ka-ndəʔru ka-nšəbyu b... b l-kimawi. (11) iyih, b ʕariqa kimawiya. (12) waxxa hiya ma-mzyana-ʕ bzzaf hit n-naḍira ʕwiya f dakši mwadd kimawiya.

(13) mn bʕd ka-n... ka-n... ma ygan ngg<sup>wr</sup>? (14) kan... mn bʕd ka-nsəddi... ka-nsəddiu... ka-nžibu s-səda ka-nsəddiu. (15) ʕad ta-ndiru... ka-nwəqqfu dak aʕta wəlla... z-zərbiya, aʕta... l-mnwal, iyih, l-mnwal, ʕafi w ka-nxədmu fiha ta tkəmməl, ta tkəmməl z-zərbiya. (16) mn bʕd ka-nddiu-ha l s-suq. (17) hna kayn s-suq l-ʔusbuʕi... kayn s-suq l-ʔusbuʕi, ka-nddiu z-zərbiya l s-suq l-ʔusbuʕi, ka-yʕriu-ha ʕndna duk l-mūhtəkirin, w ʕafi. (18) w kaynin... kaynin... kaynin l-ʕayalat daba lli dayrin ʕəmfiyat w tʕawniyat nisaʕiya, daba lhəmduʕla bʕda xəddamin ʕwiya, maʕi bḥal... maʕi bḥal ʕḥal hadi. (19) waxxa... z-zərbiya... ʕwiya ma... t-taman habt bzzaf, w hada huwa l-muʕkil dyal z-zərbiya. (20) lli ka-tʕani mnu z-zərbiya bzzaf huwa l-muʕkil dyal t-tswiq, tswiq z-zərbiya, (21) w ka-ntmnau zəʕma y... ikun s-suq hnaya, hna f taznaxt, maʕi f mərrakʕ, ikun t-ṭabʕ dyal z-zərbiya hnaya f taznaxt, maʕi... maʕi f mərrakʕ, (22) w byina ʕi mihražan wəlla ʕi məʕriq, baʕ ttʕərrəf z-zərbiya l-wawazgitiya, baʕ i... baʕ iʕriu mn ʕndna, (23) bḥal matalan l-maʕriq lli ka-ykun bḥal qəlʕat mguna, bḥal haduk l-məʕrufin... iyih, mwasim, (24) iyih baʕ ḥta hna tkun ʕndna ʕi ḥaža baʕ ttʕərrəf had z-zərbiya (25) w byina nrəddu liha l-ʔiʕtibar l z-zərbiya w l-qima dyalha...

[kifaš zəfma ka-tšufi l-mustəqbəl dyal z-zərbiya f taznaxt?]

(26) şarahatan l-mustəqbəl dyal z-zərbiya hnaya fi taznaxt ma... la yubašširu bi xayr. (27) la yubaššir... yəfni z-zərbiya bhal hiya yađiya u... bhal hiya yađiya w ka-tmut. (28) yađiya w ka-tmut. (29) daba... daba l-şayalat ila ma-nađu-ş baš ihəđru ſla l-ħəqq dyalhüm w ħəqq z-zərbiya, ana ka-ndənn bnni... blli had z-zərbiya yađi... yađi tmut nihađiyan u... (30) b had t-ťariqa yəfni ma- yađišt tštəmərri, had z-zərbiya.

### Gloss

(1) In the name of God... now I'm going to talk about the steps involved in making a carpet – a carpet from the Wawazgit area, or from Taznakht. (2) The first phase is that of gathering the wool: we shear the sheep, and this phase is done by a man. (3) Afterwards we take the wool and carry it down to the river, where we wash it with soap, and then we spread it out until it dries. (4) Next we pick out all the thorns and debris, and then we roll it up into balls, and after we've done that we spin it into yarn. Once the wool's been made into yarn, it's time for the conditioning phase. (5) [For] this phase, it gets soaked for twenty-four hours, and during this phase we condition it in alum – alum and water. (6) After that comes the dyeing phase. (7) When it comes to dyes, there are both natural dyes and chemical dyes. (8) With natural dyes, for example, we do the dyeing with natural herbs like *ashfud*, *tarubiya*, fig leaves, walnut leaves, henna, *assay*, and saffron. (9) There are a lot of herbs, but right now because of the drought it's really hard to find those herbs because they grow in the mountains. (10) There aren't a lot right now, so we're forced to do the dyeing with chemicals. (11) Yeah, the chemical way. (12) Although it's not really good because people don't look positively at that – at chemical ingredients.

(13) After that we... how do you say “to set up the warp?”<sup>24</sup> (14) So then we prepare the warp<sup>25</sup> – we bring the warp fabric and string it up. (15) Then we do... we stand up the loom, or? Carpet, loom... loom, yeah, the loom, that's it – and we work on the carpet until it's finished. (16) Afterwards we take it to the market. (17) There's a weekly market here – we take the carpet to the weekly market, and those monopolizing [middlemen] buy it from us, and that's it. (18) But there are these ladies nowadays who have put together women's associations and cooperatives; now, thank God, they're actually working a bit, unlike in the past. (19) Although the prices for carpets are really low, and that's the problem with carpets. (20) What carpets are suffering from is a marketing problem – carpet marketing. (21) We wish we had some sort of market [for them] here in Taznakht, not in Marrakesh, and that the impact of the carpet [industry] was here in Taznakht, not in Marrakesh. (22) We want some sort of festival or some sort of exposition, so that carpets from the Wawazgit region will get exposure, so people will buy from us, (23) like for example the exposition that they have in Kalaat M'Gouna, like those well-known ones...

<sup>24</sup> This segment is a code-switch to Tashlhiyt Berber.

<sup>25</sup> Set of parallel strings running lengthwise that form the core of the carpet.

yeah, seasonal festivals. (24) So we too can have something, so these carpets will get exposure. (25) We'd like to give the carpet back its esteem and value.

[How do you see the future of the Taznakhti carpet?]

(26) Honestly, the future the carpet here in Taznakht doesn't... it doesn't look good. (27) It doesn't look... I mean it's like the carpet is going... like it's going and dying. (28) Going and dying. (29) Now, unless women get up and start speaking about their rights and the rights of the carpet, I think that the carpet's going to... it's going to die once and for all. (30) The way things are going, it's not going to stick around, this carpet.

## Text 2: Getting Married

*female, Arab, 44, uneducated, Asfi area (same as in Text 6)*

### Text

(1) daba l-wəld ila bya itžəwwəž l-bənt ta-ymši... ka-ymši l l-walidin dyalu ta-yxtəbu-h. (2) ta-yŋttu-h s-skk<sup>w</sup>ar, l-hənnə, t-tmə, l-xatəm – baš ka-yŋlləmu dik l-bənt, w ta-yhəddən t-tŋəyyəd d l-ŋərs. (3) məlli ka-yhəddən t-tŋəyyəd d l-ŋərs, ka-ydədd b ihəyyəd l-kswa... (4) ka-ymšu n-nas, ka-ydiru n-nas l-ŋərs, ta-tmši l-ŋərusa ŋənd l-kūfuṛa, ka-tži, ka-yak<sup>w</sup>əlu n-nas iŋrəbu... (5) d-džəž, l-lhəm b l-bərquq, l-həlwa, bštīla, disir, kūlši dakši... ka-yak<sup>w</sup>əlu n-nas iŋrəbu, ka-tmši l-ŋərusa – ɖar l-ŋəris. (6) məlli ka-tmši l ɖar l-ŋəris, bhal hakk<sup>w</sup>a nit dakši lli daru f ɖarhūm ya-ydiru f ɖ-ɖar dyal l-ŋəris. kif kif. (7) l ya-tži, bb<sup>w</sup>aha, mm<sup>w</sup>ha, ya-yžibu lha l-ftur. l-ftur ašnu fih? (8) fih d-džəž, fih l-lhəm b l-bərquq, fih l-kswa dyalha, fih duk t-tbiqat dyal l-fəddə ŋamrin b l-məska w l-fanid w dakši – ka-yžibu lha l-ftur dyalha. (9) ɖ-ɖar dyal l-ŋəris. (10) l-mūhim ta-yak<sup>w</sup>əlu n-nas iŋrəbu... l-ŋəšiya, ka-ytfərrqu. (11) kūll wahəd fin ka-ymši ka-tbqa l-ŋərusa f ɖ-ɖar dyalha. hada huwa l-ŋərs dyalna.

(12) l-ŋərubīya la. ta-y... (13) ma-ka-tmši-š l-ŋərusa l ŋənd n-nəggafa. ka-ydiru yir l-ŋərs ŋadi. (14) l-ŋərs ŋadi – l-muhim ila l-wəld ila bya l-bənt ta-ydiru dik d-dbiha w idiru dakši... ta-ydiru t-təlba – l-fatḥa. (15) hadši f l-blad. ka-ydiru l-fatḥa, ŋafi w ka-tmši l-ŋərus. (16) kayn-ši... yir ka-yləŋbu yir binathūm, ma-kayn l-g<sup>w</sup>ərb wəlla htta ši haža. (17) maši bhal l-mədina. ka-ydiru-š l-g<sup>w</sup>ərb. (18) yir t-təlba w l-fatḥa, ka-yak<sup>w</sup>əlu n-nas iŋrəbu, ŋafi ha l-ŋərusa mšat [f] ŋənd ražəlha l ɖ-ɖar.

### Gloss

(1) Now if a boy wants to marry some girl he goes... he goes to his parents and they get him engaged. (2) They give him sugar, henna, dates, and the ring – to let the girl know, and so he can figure out the date of the wedding. (3) Once he's figured out the date of the wedding, they take out the dress... people go and have the wedding, the bride goes to the hairdresser, she comes back, people eat and drink... (5) chicken, meat with prunes, sweets, pastilla, fruit, all of that... people eat and drink, and the bride goes off – [to] the

groom's house. (6) Once she's gone to the groom's house, they do exactly what they did in their house in the groom's house. The same thing. (7) Once she's there her father and mother are going to bring her breakfast. What does breakfast include? (8) There's chicken, there's meat with plums, there's her dress, there are those silver platters full of gum and candy and all that – they bring her her breakfast. (9) [At] the groom's house. (10) Anyway people eat and drink... in the afternoon, they all split up. (11) Everyone goes his way and the bride remains with the groom at her [new] house. That's our type of wedding.

(12) Not in the countryside. (13) The bride doesn't go to the hairdresser. They just have a normal wedding. (14) A normal wedding – if a boy wants some girl they slaughter an animal and do all that... they have a religious ceremony – reading the *fatiha*. (15) That's in the countryside. They read the *fatiha*, that's it and the bride goes off. (16) There's not... they just play [music] amongst themselves, there's not a band or anything. (17) Not like in the city. They don't have a band. (18) Just the religious ceremony and the *fatiha*, people eat and drink, that it's and there goes the bride along with her husband to the house.

### **Text 3: Harsha Recipe**

*female, Arab, 28, high-school education, Taznakht*

#### **Text**

(1) bəsmillā r-rəḥman r-rəḥim, ɣa-nqəddəm wahəd l-wəʃfa dyal... wəʃfa myaribiya, ʃibara ʃla *ħərša*. (2) yəʃni ɣadi nax<sup>w</sup>du l-maqadir – ʃənnə nəʃʃ kilu dyal l-ħərša, ʃnna ʃwiya dyal l-məlħa, w wahəd l-biḍa, w kas dyal l-ħlib. (3) ɣadi nxəlltu l-ʃənaʃir, w ɣadi nax<sup>w</sup>du l-məqla, w ɣadi nxwiu-h f l-məqla, w ɣadi ntiybu-ha fuq l-buṭa. (4) yəʃni ɣadi nnəqʃu l-buṭa, w ɣadi nxlliu-ha ttiyb mzyan, w... w ʃafi.

#### **Gloss**

(1) In the name of God... I'm going to present this recipe, a North African recipe, known as *harsha*. (2) We'll start with the measurements – we have a half kilogram of *harsha* [mix], a little bit of salt, an egg, and one cup of milk. (3) We're going to mix the ingredients together, get a frying pan, pour it into the frying pan, and heat it over the stove. (4) We'll turn down the stove, and let it cook boil well and... that's it.

### **Text 4: How to Make Tanjia**

*male, Arab, 24, college education, Sefrou*

#### **Text**

(1) waxxa, daba ɣadi nqul lkūm kifaʃ ka-ntiyb t-ṭanžiya – t-ṭanžiya ʃ-ʃfrawiya. (2) daba kaynin f l-mayrib kaynin žuʃ ṭanžiyat – kayn t-ṭanžiya l-mərakšiya w t-ṭanžiya

š-šfrawiya. (3) t-šanžiya š-šfrawiya ka-tdar b l-... b l-lhəm dyal l-bəgri, aw dyal l-šžəl, w ka-t... w b l-btata, btata šyira gaš, šy<sup>wi</sup>wra bzzaf. (4) w mn bšəd... mühim ka-n... ka-nqəbtu wahəd l-... wahəd l-šina? dyal... dyal t-šin, w ka-ndiru fih... ka-ndiru fih l-lhəm, ka-ndiru fih bzzaf dyal l-bšla, w z-zit l-bəldiya. (5) bzzaf dyal z-zit l-bəldiya. (6) w ka-ndiru, ka-nzidu šliha... ka-ndiru lih s-smən, s-smən. (7) w ka-ndiru l-kammun bzzaf, l-fəlfla l-həmra, z-zəšfran l-bəldi w z-zəšfran r-rumi, w kan-diru t-tuma – ka-ndiru ši rbəš d r-ryus d t-tuma – w ka-ndiru l-krūmb w l-btata. (8) w ka-ndiru kas dyal l-ma, šafi. (9) w ka-ndiru l-... rəbta dyal l-qšbur w l-məšdnus, rəbta šyira... (10) w ka-ndiru, lli bya izid l-guza ka-yzid-ha w lli bya izid ras l-hanut ka-yzid-u. (11) w s-skənžbər. (12) mn bšəd, ka-tqəl šəla hadik l-šina? dyal t-šin, ka-tqəl šəlih b ši... b ši kayit, kayit dyal qalb s-skk<sup>w</sup>ar, wahəd kayit zraq... (13) w ka-tšədd-u b wahəd l-qənba, mzyan... w ka-tddih l l-fərnatši. (14) l-fərnatši, ka-yxli-h šəndu l-lil küllu. (15) ka-ytiyb, həta l-yədd lli f š-šbah, ža huwa hadak l-hədaš wəlla l-šəšra ka-tmši šəndu. (16) w ka-txəlləs-u, ka-təšti-h ši baraka, imma šəšra drahəm wəlla xəməšəšər dərham wəlla miyəť dərham, šhal ma kan. (17) ka-thllə fih, w ka-tžib-ha l d-đar w ka-tak<sup>w</sup>əl-ha mša l-ša?ila dyalək, w b š-šəhha w l-šafiya!

### Gloss

(1) OK, now I'm going to tell you all how I make [a dish called] *tanjia* – Sefrou-style *tanjia*. (2) Now there are two *tanjias* in Morocco – there's Marrakech-style *tanjia* and Sefrou-style *tanjia*. (3) Sefrou-style *tanjia* is made with beef, or veal, and potato – one really small potato, a really tiny one. (4) Afterward... anyway we grab this pot made out of clay and put meat in it, we put a lot of onion in it, and olive oil. (5) A lot of olive oil. (6) And we add preserved butter to that. (7) We put in a bunch of cumin and red pepper, real saffron and imitation saffron, and put in garlic – we put in four or so garlic gloves – and we put in cabbage and potatoes. (8) And we put in a single cup of water, that's all. (9) We put in a garnish of coriander and parsley, a little garnish... (10) and whoever wants to add nutmeg adds it, and whoever wants to add *ras el hanout*<sup>26</sup> adds it. (11) And ginger. (12) Afterward you close off that clay pot, you close it off with some sort of course paper, the course paper from a block of sugar, this blue course paper... (13) and you grab it with a piece of burlap, grab it well... and you take it to the furnace operator<sup>27</sup>. (14) The furnace operator keeps it with him the whole night. (15) It cooks, until the next day in the morning, once ten or eleven comes around you go to him. (16) You pay him, give him some sort of tip, either ten dirhams or fifteen dirhams or a hundred dirhams, whatever you've got. (17) You take care of him, and you bring it to the house and eat it with your family, bon appetite!

<sup>26</sup> A blend of choice spices prepared by a spice seller and varying from shop to shop.

<sup>27</sup> Typically this would be the person who tends the furnace at the *hammam*, or public bath.

### Text 5: Making Tea

*female, Berber, 28, high-school education, Taznakht area*

#### Text

(1) salamu ʕlikūm, riḡ awn mldḡ... ʔanwərrikūm kra n... kifaš ka-ndiru atay, b l-lūya l-ʕərb... ɣanhḡər b l-ʕərbiya. (2) f l-luwl, ka-nhəzzu l-mqraš, ka-ndiru fih l-ma. (3) ka-n... nšʕlu [lly n] l-buṭa taʕ atay b l-ma. (4) ka-nhəzzu l-bərrad, ka-ndiru fih ḡbub d atay, ka-nšəlləlu-h, (5) f l-lu... f l-luwl ka-ndiru-h l-luwl f l-kas, mn bʕdət tani ka-nšəlləlu ḡbub atay... (6) ka-nkəbb<sup>w</sup>u l-ḡaṣa, zʕma ka-nḡiydu-h, ka-nšəlləlu-h, ka-nəržžʕu atay l-luwl l l-bərrad. (7) ka-nzidu ʕlih s-skk<sup>w</sup>ar, b š-šiba wəlla b ši aʕšab, kūll waḡəd aš ka-yšərb, kūll waḡəd w ašmən ḡaṣa ʕziza ʕndu. (8) mn bʕd ka-nəržžʕu l-bərrad l l-buṭa ḡta iyla, ḡta iṭiyb ʕawd tani. (9) ʕad ka-nḡṭtu-h, w nžərrbu-h, yak ṣa huwa hadak. (10) w ḡa atay waḡd. wa šukran.

#### Gloss

(1) Hello, I want to show you all...<sup>28</sup> I'm going to show you all how we make tea, speaking in Arabic. (2) First, we get a pitcher and fill it with water. (3) We light up the tea-sized burner and put water on it. (4) Then we get a teapot and put rolled tea leaves in it and rinse them. (5) At first we do that in a cup, then after that we rinse the tea leaves again. (6) We pour out a bit – I mean we remove it – we rinse it, then we put the original tea back in the teapot. (7) We add some sugar to it, with wormwood or some herbs – everyone drinks whatever [he likes], everyone and anything he likes. (8) Afterwards we put the teapot back on the burner until it boils, until it's steeped a little more. (9) Then we take it off, and we can give it a taste to see whether it's right. (10) And now the tea is ready. Thanks.

### Text 6: Moroccan Tea and Coffee

*female, Arab, 44, uneducated, Asfi area (same as in Text 2)*

#### Text

(1) šafi? ka-tḡir l-məqraš ḡtta ka-yṭiyb, w ta-tḡir l-ḡbub dyal atay f l-bərrad, w ka-txwa ʕlih l-ma ṭayb. (2) w ka-txəlli-ḡm waḡd šwiḡ, waḡəd žuž dqayq. (3) w ka-tšəffi f l-kas dak atay l-ləwwəl, w tʕawd tšəlləl... ʕawd tšəlləl žuž xəṭrat. (4) hadak t-tšlila dyal žuž xəṭrat, ḡadik yir txwi-ha. (5) ka-ttlaḡ. (6) w hadak atay l-ləwwəl tərədd-u l l-bərrad. (7) lli tərəddu l l-bərrad yaḡi tḡir ʕlih l-iqama, n-nəʕnaʕ. (8) ḡir s-skk<sup>w</sup>ar, w tʕəmmər l-bərrad. (9) llaḡ ila itʕəmmər yaḡi itšḡḡər ʕla l-buṭa. (10) ta-ytšḡḡər, ʕad thəṭṭəṭ, ʕad tk<sup>w</sup>əbb ʕad tšrəb... ʕad gul bismilla.

(11) l-qəḡwa rah... dyalək f ḡ-ḡar ʕərəfti l ḡad l-qəḡwa?

---

<sup>28</sup> First instance of this phrase is in Tashlhiyt Berber.



(12) l-qəhwa ta-tdirha f l-briq, ta-tdir ʕlah s-skk<sup>w</sup>ar, w ka-tdirha ttiyb. (13) htta ka-ttbəx mzyan, byiti tdir fiha l-ħlib diru, byiti tšerbha hakk<sup>w</sup>a kəhla šərbha. (14) walakin ka-thəbbəṭha ttiyb mzyan. šafi?

### Gloss

(1) OK? You put the pot [of water] on until it boils, and you put rolled tea leaves in the teapot, and you pour boiling water on it. (2) You leave them for a little bit, two minutes or so. (3) You strain that original tea into a cup, and keep rinsing... you rinse it two times. (4) You can just pour out what you've rinsed. (5) It gets tossed out. (6) You return the original tea to the teapot. (7) You add mint to what you put back in the pot. (8) Put in some sugar, and fill up the pot. (9) Once it's full it's going to steep on the stove. (10) It steeps, then you take it off, then you pour it, then you drink it... then you say "bismillah."

(11) That coffee... you make in the house, do you know how to make it?

(12) [As for] coffee, you put it in the coffeepot, you add sugar to it, and you make it boil. (13) Once it's cooked well, if you want to add milk add it, if you want to drink it black like that drink it. (14) But set it down to boil well. OK?

### Text 7: Ramadan in South Morocco

*male, Berber, 16, HS student, Taznakht*

### Text

(1) salamu ʕlikūm, yaḍi nḥḍər mʕəkūm b l-ʕərbiya, nwərrikūm kifaš ka-ykun ramḍan f l-mayrib – t-taqalid dyalhūm. (2) f l-luwl, mn s-šur ka-nsəhr... (3) nsəhru f r-rbʕa d š-šbaḥ wəlla r-rbʕa w nəšš, l-mūhim ʕəla ḥsəb təwqit kūll mdina. (4) r-rbəʕ w nəšš, mmkn nsəhru... kūll waḥd ka-y... baš ka-yshər ka-ykun bḥal daba ka-ykun lli ka-yshər b t-tažin, w l-xūbz, l-mūhəm w atay... (5) lli ka-y... lli ka-yshər ʕawd tani b l-ḥələwiyat w l-ʕašir w l-limun baš zəʕma ikun ʕəndu t-taqa baš iṣbər kūll n-nhar htta l... ʕawd tani htta l s-sbəʕ dyal l-ʕəšiya, (6) w ʕawd tani kaynin daba bḥal l-qūra w dakši dyal l-badiya ka-ysəhru b ši tažin w t-tmər w l-ħlib baš nit hit ḥūma ka-ykunu bzzaf tkərfəs f d-dwawər, (7) zəʕma baš nit ikun... zəʕma ikun ihss... b... dɣiya zəʕma ihss b l-ʕəṭaš wəlla... wəlla ɖəɣru ž-žuʕ.

(8) l-mūhim faš kay-tsəhru ka-ymšiu ʕawd tani, lli ka-y... kaynin ka-ymšiu l l-xdma mʕa t-tmənya d š-šbaḥ, wəlla l-ʕəšra, (9) w ʕawd tani f... f... f ma-ka-ykun-š, ma-ka-ykun-š l-ftur ḥətta l s-sbəʕ w nəšš. (10) s-sbəʕ w nəšš ta... l-mūhim ḥna ka-n... ka-nftəru s-sbəʕ w nəšš (11) w ʕawd tani f... l-ftur ka-ykun bzzaf f l-ftur, l-ftur xaššətan ka-ykun bzzaf fih dyal l-ʔanwaʕ dyal l-makla. (12) ka-ykun bzzaf, xaššətan waḥd n-nuʕ dyal l-ḥərira ka-tkun, kūll waḥd ka-ydir l-ḥərira f l-ftur. (13) w ka-ykun l-msmmən w l-bayrir, ʕəla ḥsəb lli bya i... waḥd. (14) w ʕawd tani ḥna ka-ndiru l-xūbz b š-šəhma, lli ka-y... ka-y... ka-ygul... ka-ygul lih... ka-ygul lih mm<sup>w</sup>alin š-šəlha aɣxšiš. (15) l-mūhim w... ka-ngul

ʕəla ʔsəb kʷll waħd ka-y... aš ka-yakʷəl f l-ftur. (16) ka-ykun l-qəhwa, atay, l-ħlib, l-ʕašir, žamiš l-ʔanwaš dyal l-ħələwiyat ka-tkun fih, w ka-tkun fih ʕawd tani ka-ngul lik l-ħərirə, w ka-tkun fih... ayy waħd. (17) ayy waħd – [...] had l-ftur ka-ykun zəfma kan ka-ydərɣ-k ž-žuš n-nhar kʷllu kaml dyal l-xdma, w dyal l-ʕəya, wəlla dyal... (18) l-mūhim, ila ma-ʕndək-š dik t-ʔaqa baš tšbər kʷll n-nhar ka-ykun fik waħd l-ʕəya dyal l-makla, ka-takʷəl htta tšəbš.

(19) ʕawd tani faš ka-takʷəl l-ftur, ka-y... ka-txʷərž l s-saħa txʷərž mʕa l-ʕaʔila dyalk, wəlla šhabk, ka-txʷərž l s-saħa... (20) ka-ykun, l-mūhim ka-yduwru n-nas baš inqqšu... l-mūhim yduwru baš dik l-makla dyalhūm zəfma ma-iži dirikt ma-ynʕəs ma-ydərɣ-u ž-žuš. (21) ah, [...] l-makla ttəhdəm, l-mūhim... ka-tnzəl b šwiya, zəfma, ma-y... hit ila kaynin... kaynin ši waħdin ka-yftəru, ka-ymšiu dirikt s<sup>29</sup> n-nʕas. zəfma, ka-ydərɣhūm... ila mərɖ... ka-ymərɖu wəlla ši ھاڙا. (22) ھنا l-mūhim ka-nftəru, ka-nxʷəržu l s-saħa. (23) faš ka-nxʷəržu l s-saħa ka-nləšbu mʕa š-šadiqa dyalna, mʕa l-mūhim l-ʕaʔila dyalna, ka-nmši mʕa l-walid, l-walida... ka-nduwru f s-saħa, htta dik t-ʔnaš w nəšš, t-ʔnaš w nəšš dyal l-layl, l-lil. (24) t-ʔnaš w nəšš, l-mūhim l-insan ʕawd tani f š-šbaħ lli kayn... lli ma-ʕndu-š l-xədma ka-ymši ɣir inʕəs zəfma... ka-ynʕəs ʕawd tani baš zəfma ma-ydərɣb l... wlla ž-žuš wəlla l-ʕəš.

(25) f l-ʔnaš ʕawd tani ka-nrəžšu l d-dar, kaynin ši waħdin ka-ytʕəššaw. (26) ila kaynin... zəfma kaynin ši waħdin lli ma-y... lli... ka-ydərɣhūm zəfma ž-žuš waxxa ila klau l-ftur, ka-ytʕəššaw. (27) yəʕni l-mūhim ʕawd tani kʷll waħd iʕəš... ka-ytʕəšša, ytʕəšš... ayy waħd lli ka-ybɣi [...] ši ھاڙا ka-y... ka-ytʕəšša biha, w ka-ynʕəs. (28) ka-ynʕəs ʕawd tani htta l r-rbəš d š-šbaħ ʕawd tani d s-šur. (29) kʷll šhur, kʷll waħd ka-ybɣi ʕawd tani aš ka-yakʷəl fih w hakkʷa t-taqalid dyalna f l-mayrib. (30) l-mūhim kʷll waħd baš ka-ygul htta daba ž-žəzaʔir wəlla mšər wəlla tuns kʷll waħd mʕa t-taqalid dyalu... (31) l-mūhim t-taqalid dyalna bħal hadi. (32) gul... ʕawd tani gul ka-ytfərɣqu... l-fərɣ bin l-mudun. (33) kʷll mdina, baš ʕawd tani ka-ttmiz b l-mudun. (34) bħal... maši bħal š-šamal, bħal... maši bħal ž-žanub, maši bħal ʕawd tani l-ʕərəb... (35) kaynin... b l-ʕərbiya w kaynin ʕawd tani š-šəlha u... kaynin mūxtalfa b l-ʕərbiya w kaynin ʕawd tani mūxtalfa b š-šəlha. (36) baš ikun... kaynin l-mūhim ixtilaf f l-lūɣa w dakši, w ʔsəb l-lūɣa ta-ykunu ʕawd tani taqalid, kʷll waħd w t-taqalid dyalu iduwz ramɖan. (37) l-mūhim hadši lli ya-n... ka-nxdəm l-mūhim hadši l-məʕlumət lli kafiyin, ka-ntuffər biha.

## Gloss

(1) Hello, I'm going to talk to you all in Arabic, going to tell you how Ramadan is in Morocco – about their customs. (2) The first thing we do is have *shour*<sup>30</sup>. (3) We eat it at four or four-thirty in the morning, depending on each city's schedule. (4) At four-thirty we might have *shour*... for *shour*, there are people who have a tagine, and bread, and tea... (5) and have sweets, and juice, and oranges in order to have the energy to get

<sup>29</sup> Likely the Berber particle *s* 'to' in an impromptu code-switch.

<sup>30</sup> A meal eaten in the early morning, prior to beginning the daily fast during Ramadan.

through the whole day, until seven in the evening. (6) There are [people] in villages and rural places that have some sort of tagine, and milk, and dates for *shour* because they really have a tough time in the villages, (7) and that makes them get thirsty or hungry really quickly.

(1) Anyway, once they've eaten *shour* they go out. There are some people that go to work at eight in the morning, or ten, (2) and there won't be the *iftar*<sup>31</sup> until seven-thirty. (10) Between seven-thirty and... well we eat it at seven-thirty. (11) The *iftar*, in particular, involves a lot of different types of food. (12) There's a lot – especially this certain type of *harira*<sup>32</sup>; everyone makes *harira* for *iftar*. (13) And there's *mesemmen* and *baghrir*<sup>33</sup>, depending on what one wants. (14) And we also make fried fat-bread, which Berber speakers call “*arkhsis*.” (15) Anyway, as we say, everyone has something different for *iftar*. (16) There's coffee, tea, milk, juice, there are all sorts of sweets, and of course there's *harira*, there's... anything. (17) Anything – by the time this *iftar* comes around, you've been hungry throughout the whole long day of work, and of exhaustion, and of... (18) anyway, if you don't have that energy [you need] to get through the day you get exhausted from not eating, so you eat until you're full.

(19) Once you've eaten *iftar* you go out to the [public] square – perhaps with your family, or your friends – you go out to the square. (20) People walk around in order to lessen... I mean they walk around so that food of theirs... I mean so they won't go straight to sleep feeling hungry. (21) The food has to digest... it has to go down slowly, because some people, if they go straight to sleep, might get sick or something. (22) As for us, we eat *iftar* and then go out to the square. (23) When we go out to the square we kill time with our friends, or with our family, until twelve-thirty or so, twelve-thirty at night. (24) At twelve-thirty... and in the morning whoever doesn't have work in the just goes back to sleep... he goes to sleep so he won't get hungry or thirsty.

(25) Around twelve we go back to the house; there are some people who eat dinner. (26) If there are... I mean there are some people who... they feel hungry even though they ate *iftar*, so they eat dinner. (27) Anyone who wants something for dinner has it, and then sleeps. (28) Once again, he'll sleep until four-thirty in the morning, the time for *shour*. (29) For each *shour*, everyone eats whatever he wants, and that's how our customs go in Morocco. (30) Anyway everyone says something different, like in Algeria or Egypt or Tunisia, everyone has his own customs. (31) But ours are like this. (32) You can also say that there's a difference between cities. (33) Every city is distinct from the others. (34) Like... it's not like the North, it's not like the South, it's not like [those of] the Arabs. (35) There are [customs] in Arabic and there are likewise [customs] in Berber, some that are different in Arabic and some that are different in Berber. (36) Anyway, there's the

---

<sup>31</sup> The fast-breaking meal.

<sup>32</sup> Moroccan soup.

<sup>33</sup> Similar to crepes and pancakes, respectively.

difference in language and that sort of thing, and according to the language there are customs – everyone spends Ramadan with his own customs. (37) Anyway, that’s what I have to work with, that’s as much relevant information as I’ve got.

### **Text 8: The Baccalaureate Exam**

*male, Arab, 20, HS student, Taznakht*

#### **Text**

(1) bəsmilla r-rəḥman r-rəḥim, ana təlmid... ṭalib b s-sana t-tanya ḥakaḷuṛya. (2) ana ka-nqra hna f taznaxt. (3) l-... l-ḥakaḷuṛya yəṣni... šhadət l-ḥakaḷuṛya ka-təṣni bzzaf dyal l-ḥwayṣ. (4) b n-nəsba l l-ḡimtiḥan ḥakaḷuṛya, ka-ndəwwzu-h f axər sana f l-ṣi... f l-b... t-tanawi, yəṣni l-ḥak. (5) b n-nəsba l had l-ḡimtiḥan, imtiḥan f ṣiḥa kan yəṣni sahl – mn ṣiḥa nṣufu [ʔ]nn l-mūḥtawa dyalu sahl, walakin l-ḡistiṣdad dyalu ka-ykun šwiya ṣṣib. (6) had l-ḡimtiḥan yəṣni ka-ykun f axər sana, ka-ykunu fih xəmsa d l-mwadd. b n-nəsba l l-inḡliziya, r-riyaḍiyat, l-fizik, l-ṣulum, mwadd uxra. (7) b n-nəsba l l-ḡistiṣdad d l-ḡimtiḥan, ka-nstaṣddu lih, ngulu šḥər qbəl l-ḡimtiḥan. (8) ka-n... ka-nraṣṣu l-mwadd lli ṣndna, ka-nraṣṣu d-duṛus, yəṣni ka-nstaṣddu ṣəqliyan w nəṣsaniyan l had l-imtiḥan.

(9) mnin ka-yṣi l-ḡimtiḥan, yəṣni ḍaruri xṣṣ l-ḡinsan inuḍ mṣa ṣ-ṣbaḥ, yəṣni ḍaruri ikun waḥd t-təyḍiya mtwazna, mn hit t-təyḍiya, mn hit küḷṣi, w yəṣni xəṣṣ l-insan ikun mstaṣdd l yəṣni... mn ṣamiṣ ḍ-ḍuruf. ikun i... ikun müḥaḥhal yəṣni bəḍəniyan w küḷṣi. (10) mnin ka-nmṣiu nduwzu l-ḡimtiḥan, yəṣni ka-nduwzu-h mṣa t-tmənya d ṣ-ṣbaḥ. (11) ka-ndəxlu mṣa t-tmənya d ṣ-ṣbaḥ, ka-ndəwwzu saṣṭayn. (12) matalan, təlt... tlat... təlt swayṣ f l-fizik – ka-nduwzu təlt swayṣ ḥəṭta l-ḥəḍaṣ. (13) mnin ka-nkəmmulu, yəṣni ka-nxwərṣu... kayn lli fərḥan, w kayn lli ka-ybki, ka-nmṣiu yəṣni ka-nərtahu waḥəd šwiya, w lli ḃya iraṣṣ – (14) yəṣni ṣəḥqaṣ kayna waḥəd l-mwadda xwəra ka-nduwz-ha f l-ṣəṣiya, hiya l-anglay, *english*, ka-nduwz-ha f l-ṣəṣiya mn t-tlata, ḥəṭta l-xəmsa. (15) w b n-nəsba l... kima gət lik kaynin lli f... ka-yxwərṣ fərḥan kayn lli ka-yxwərṣ ka-ybki. (16) w ka-n... yəṣni ṣafi ka-nkəmmulu l-ḡimtiḥan, xəṣṣna ṣawd tani nrəṣṣu n-nḥar t-tani. (17) n-nḥar t-tani, ka-ndəwwzu... ṣawd tani ka-ndəwwzu l-... l-... l-fiz... l-ṣulum... ka-ndəwwzu l-ṣulum w... mwadda xwəra nsit-ha. (18) l-muḥim, ka-ndəwwzu l-ṣulum w waḥəd l-mwadda xwəra. (19) yəṣni l-... fiha təlt yyam.

(20) mnin ka-ndəwwzu təlt yyam, yəṣni küḷṣ waḥəd aṣ ka-ydir. (21) kayn lli ka-ymṣi igəḷs f ḍ-ḍar, kayn lli ka-ysafər, kayn lli ka-ydəwwz l-wəqt mṣa ṣḥabu, ka-ymṣi r-rəḥalat. (22) ana b n-nəsba li ka-nmṣi ngəḷs f ḍ-ḍar hnaya, ka-ntsənnan n-ntaṣiṣ. (23) yəṣni ka-tduz waḥəd simana, w ka-yxwərṣu n-ntaṣiṣ. (24) haduk n-ntaṣiṣ, tamma ka-tṣərf b rask waṣ nṣəḥti wəlla ma-nṣəḥti-ṣ. (25) yəṣni kayn l-... ntaṣiṣ l— ka-ybiynu fihūm lli nṣəḥu, w kaynin n-ntaṣiṣ xwəra ka-ybiynu l-... lli ma-ṣabu-ṣ n-nūqṭa mzyan ṣəndḥūm l-ḡistidraki. (26) iqdər idəwwzu l-ḡistidraki. (27) l-ḡistidraki yəṣni... ka-t... yəṣni tṣawd tərṣṣəṣ n-nūqṭa lli mṣat lik, fḥəmti? (28) ka-tdəwwz waḥəd xəmsṣaṣər yum, ka-tdəwwz l-... ka-tṣawd l-ḡimt... nəṣs l-ḡimtiḥan. (29) mnin ka-tṣawd nəṣs l-ḡimtiḥan, ka-tsənnan waḥəd

təlt yyam, bima ytšhhu l-wraq, w dik s-saṣa ka-tx<sup>w</sup>ərž n-natiža. (30) mnin ka-tx<sup>w</sup>ərž n-natiža, ka-tšənt w... ila nžəhti, raha šhad... haṣalta ʕla šhadət l-ḥakaḷuṛya. (31) w ma-nžəhti-š, kayn lli ka-ymši ka-yxdəm, imši ixdəm ši blaša, ši... ši məʕməl wəlla ka-ymši idir l-mikanik wəlla... (32) yəʕni maši... küll waḥəd aš ka-ydir, yəʕni mxəddmat ʕadiyin. (33) w kayn lli mnin ka-yšədd... lli šədd l-ḥakaḷuṛya, kayn lli ka-ymši l l-madaris – kayn l-fərq bin l-madaris w kayna l-žamiṣa. (34) yəʕni küll waḥəd baš ntšəbbət küll waḥəd aš dar, yəʕni dakši lli bya.

(35) had šhadət l-ḥakaḷuṛya ka-təʕtabər mühimma, ɖaruriya. (36) yəʕni ila ma-ʕndək-š šhadət l-ḥaṛuṛya... l-ḥakaḷuṛya rah ma-ʕndək walu. (37) f... f had l-ḥala, ila šədditi l-ḥakaḷuṛya rah ka-tkun fərhan w l-ʕaʕila dyalk ka-tfərḥ mṣak, w ka-ykun... ka-ykun waḥəd s-sṣada kbira wəšt l-ʕaʕila. (38) w ila ma-nžəhti-š rah l... l-ʕaʕila dyalk yəʕni ka-ybqa fiha l-ḥal walakin ka-tṣawnk, yəʕni ka-tsaṣdək. (39) ka-tsaṣdək tžawzk l... tžawzk l-muškil, w lli bya itṣawd l-ṣam yəʕni ši haža ṭabiṣiya ma-kayn ḥtta muškil. (40) w f had l-ḥala ka-ngul lik bižanna raha qimət had šhadət l-ḥakaḷuṛya mühimma. (41) ila ma-ʕndək-š šhadət l-ḥakaḷuṛya ka-təʕtabər ummi – baqi ma-qriti. (42) ka-təʕtabər bḥal yalla ḥarbt l-ʕumiyya. (43) w hadši lli kayn.

### Gloss

(1) In the name of God... I'm a student in the second year of the baccalaureate. (2) I study here in Taznakht. (3) The baccalaureate is... the baccalaureate degrees means a lot of things. (4) As for the baccalaureate exam, we take it during the last year of high school, the “bac” year. (5) As for the exam itself, on one hand the exam is easy – in the sense that its content looks easy, but preparing for it is a bit tough. (6) The exam is during the last year, and it covers five subjects – English, mathematics, physics, sciences, other subjects. (7) As for the preparation for the exam, we start preparing for it, say, a month before the exam. (8) We review our subjects, review our lessons, I mean we prepare ourselves mentally and psychologically for this exam.

(9) When the exam comes around, a person has to get up in the morning, has to eat a balanced meal – in terms of nutrition and everything – and has to be prepared from all angles, to be physically ready and everything. (10) When we go to take the exam, we take it at eight in the morning. (11) We go in at eight, and spend a couple hours. (12) For example, three hours for physics – we spend three hours, until eleven. (13) When we finish, I mean get out... there are some people who are happy, there are some who are crying, we all go and have a rest for a bit, and whoever wants to review – (14) because there's this other subject we'll take in the afternoon, which is English – we take it in the afternoon from three to five. (15) As for... as I said there are people who come out happy and there are those who come out crying. (16) And we... I mean that it's it, we finish the exam and then we have to study for the second day. (17) [On] the second day, we take... again we take.. phys... sciences... we take sciences and... one other subject which I'm forgetting. (18) Anyway, we take sciences and this other subject. (19) I mean... it takes three days.

(20) Once we get though the three days, everyone does his own thing. (21) Some people stay at home, some travel, some spend time with their friends, go on trips. (22) As for me, I'm staying here at home, waiting for the results. (23) One week or so passes, and the results come out. (24) From those results you'll know for sure whether or not you passed. (25) I mean there's... the results make it clear who passed, and there are these other results that show the... those who didn't get a good score have the remedial exam. (26) They can take the remedial exam. (27) The remedial exam means you can get back the score you missed, you know? (28) You spend fifteen days or so, you spend the... you repeat the... the same exam. (29) Once you repeat the same exam, you wait three days or so, so the papers can be graded, at then the result comes out. (30) When the result comes out, you get all worked up and... if you passed, there it is... you've obtained the baccalaureate degree. (31) And if you didn't pass, there are some people who go work, go work somewhere, some sort of factory or they become mechanics or... (32) I mean it's not... everyone does his own thing, I mean normal vocations. (33) And when some people get the baccalaureate, they go off to private schools – there's a different between private schools, and there's the university. (34) Everyone does something different, whatever he wants.

(35) This baccalaureate degree is considered important, indispensable. (36) I mean if you don't have a baccalaureate degree, you don't have anything. (37) At this point, if you get the baccalaureate you're really happy and your family's happy with you, and there's this huge happiness in the family. (38) And if you don't pass... your family is upset but they help you, I mean they support you. (39) They help you get through the problem, and if anyone wants to repeat the year, I mean it's something natural, there's no problem at all. (40) At this point I'll say that the value of this baccalaureate degree is really important. (41) If you don't have a baccalaureate degree you're considered illiterate – you haven't ever studied. (42) You're seen as if you've just learned to read. (43) And that's how it is.

### **Text 9: Weddings in Asfi**

*female, Arab, 22, college education, Asfi*

#### **Text**

(1) as-salamu ʕalikūm – b n-nəsba l l-badiya dyalna hna qribin mn l-mdina; ʕa-tkun bʕida mnha b ši *maximum* səbʕa kiʕumiṭər. (2) b n-nəsba l z-zwaʕ ka-yduz ʕəndna sbəʕ yyam. (3) b n-nəsba l n-nhar l-ləwwl, ʕa-n... ka-yʕibu l-... kima ta-ngulu hna l-ʕədul, lli ka-yktəb l-ʕəqd bin l-ʕəris w l-ʕərusa – (4) ta-ykun ʕadi *just*... ʕa l-ʕaʕila l-qriba hiya lli hədəra, ma-ta-ykun ši hədd bərrani bʕid.

(5) b n-nəsba l n-nhar l-ləwwl, ʕa-tt... iktəbu l-ʕəqd. (6) b n-nəsba l n-nhar t-tani, n-nhar taʕ l-ħfəl lli məʕruf. (7) ka-yʕiu ʕamiʕ ʕaʕilət l-ʕərusa, l-ʕərusa ka-tlbəs, ʕadi ttəwwəq, qrab l-ʕəris kadalika l-miṭəl ilbəsu l-libas t-təqlidi wəlla, ʕla hšəb... (8) ta-ygəlsu f wahəd

l-kürsi müfəyyən f wəhəd ş-şala kbira, l-muşıqa xəddama, bnadm ta-yştəh... (9) *normalement* fəddna... ka-ykunu ya l-fayalat. r-rəzal ta-ykunu bərra. ta-ykun ya l-fayalat w l-fəris. (10) *so*... l-fəris ta-ykun haşl bin l-fayalat.

(11) l-fayalat ka-yştəhu, ta-yğənniu, ta-yduz n-nhar b xir w fəla xir, ka-tbda hadik lli l-luwla b l-fərusa w l-fəris, ka-yduz fadi... (12) n-nhar l-ax<sup>w</sup>ər, l-fəris ka-yhəwwəd l l-mdina huwa w... hu... w l-wəzir. (13) l-wəzir huwa f l-... fəndkum təmma ka-tgulu l-*best man*<sup>34</sup>. (14) l-wəzir ka-yddi-h l l-mdina, ka-ybqa işəyyl-u təmma. (15) b n-nəsba l l-fərusa ka-tdir wəhəd t-təqlid smitu təhəmmamt. (16) fəlahqaş f l-fərubə ma-kayn-ş... f l-fərubiya ma-kayn-ş l-həmmam bhal hna fəddna hna f l-mdina, *like, public steam house*, təhəmmamt ka-ydiru bhal l-... *hut*, (17) kay-bn... ka-ybni-u b l-gşab w dakşi, ka-ydəxxlu lih l-ma sxun başh isxən. (18) [...] ta-tdxəl l-fərusa təmma ta-ttysəl w ta-tx<sup>w</sup>ərž. (19) ta-yduz n-nhar fadi; ka-tdəwwz-u mfa... ka-tkun təbiət l-hal mfa l-fa?ila taş z-zuž dyalha, ka-tdəwwz n-nhar fadi ka-yşali.

(20) n-nhar l-ax<sup>w</sup>ər lli huwa n-nhar r-rabəş ta-yzi l-fəris. (21) ka-tnuđ ş-şbañ fadi, ka-tlbəs, ka-ttgədd, ka-tsnna ražəha ha yzi f l-fəşiya. (22) ražəha ka-yzi f l-fəşiya, hu w l-*best man* dyalu, hu w l-wəzir. (23) l-wəqt lli ka-yzi, ka-yzib tlata mn l-ħaža. l-ħaža lli ka-yzib liha hiya ta-yzibha l mm<sup>w</sup>ha, w ta-yzibha l mm<sup>w</sup>u. (24) yəfni ida žab liha š-şərbil, ta-yzib l mm<sup>w</sup>u š-şərbil ta-yzib l mm<sup>w</sup>ha š-şərbil. (25) ka-yzib l-luz, ka-yzib l-gərgaş, ka-yzib t-tmə, w ka-yzib ħuta – *like, big fish*. (26) l-wəqt lli ka-yzi l-fərusa ka-tx<sup>w</sup>ərž ka-tlaqa lih. (27) ka-tx<sup>w</sup>ərž ka-tlaqa lih, ka-tkun hiya labsa t-təqşi, ka-thəzz ž-žlal dyalha w ta-ybda ihəzz dakşi lli žab, ka-yhəzz-u ka-ythətt-u fəndha. (28) mnin ka-thəzz ž-žlal dyalha, axər ħaža ka-yzbəd hiya dik l-ħuta. (29) ka-tax<sup>w</sup>əd hiya fəşa, ka-yax<sup>w</sup>əd huwa fəşa, w ta-ytsab aşkun lli yadi inəqqi l-ħuta, inəqqi-ha mn l-qşur dyalha. (30) lli sbəq zəfma f t-tənqiya, huwa lli yadi ikun ħakəm f l-ħayat z-zəwžiya. (31) waxxa huwa maşi fadəl anna l-məra ka-tkun hazzə mnəqha mn şhal mn ħaža w xəşşa thəzz žlalha baş ma-ttiyhş, (32) w mfa dalika thawl tsabəq w r-ražəl ma-ta-ykun hazz walu. (33) *normalement* r-ražəl yadi irbəh. (34) şafi, l-wəqt lli ka-yrbəh dakşi l-mühm n-nas ka-yfərhu ta-yfərrqu. (35) [...] dak l-luz, w dak l-gərgaş, w t-tmə lli žab, ka-nfərrqu-h fəla d-dyaf, ta-yak<sup>w</sup>əlu, ka-ydəhku.

(36) n-nhar l-ax<sup>w</sup>ər, ka-tduz l-?umur ya fadiya. (37) n-nhar l-ax<sup>w</sup>ər smitu nhar l-ħəzam. (38) ka-yzi l-... xu l-fəris, [...] – ka-yzi, ka-yfiyt fəla l-fərusa, w ka-yhətt qəşfa taş l-xşəb, ta-yfəmmər-ha b l-ma w ta-yhətt fiha d-dəbliyiž taş n-nəqra -- *silver*. (39) l-fərusa ka-thətt gdəmha wəşt l-gəşriya lli fiha l-ma w fuq d-dəbliyiž. w ka-tfyyət fəla tlatət l-... l-fəzzara – *like, single guys*. (40) tlat l-fəzzara, ka-tşərrəb-ħüm mn dak l-ma lli kan f l-qşi. (41) mnin ka-tşərrəb r-ražəl, *like* d-dərri l-axər, mnin ka-tşərrəb-u xəşşa thrəb l-yurfə dyalha fəlahqaş ila bqat haşla təmma yadi iqtəlu-ha b l-fəşa. (42) *so like*, xəşşa

<sup>34</sup> The speaker, who is highly proficient in English, uses a number of insertional code-switches, almost certainly due to my presence.

tšərrəb tūmma thrəb. l-wəqt lli ka-thrəb, şafi ka-tduz l-ʔumur b xir w ʕəla xir, hadak n-nhar l-xaməs ta-ykun tsala.

(43) n-nhar l-ax<sup>w</sup>ər ʔadi tnuḍ, ka-tšqa f ɖarhūm ʕəla asas annəha şafi raha wəllat mərə ka-ttʕətamd təmmaya. (44) f l-ʕəšiya, ta-yžibu l ɖarhūm qəş... mn qəşʕət l-ʕərus. (45) l-qəşʕa, fiha k<sup>w</sup>əksu ʕadi, walakin fiha š-šklaṭ, fiha l-biḍ, fiha l-məska, fiha... qəşʕət l-ʕərus ka-tkun həluwa. (46) ta-yžibu žuž qşaʕi, ka-yšəṭhu, ka-yak<sup>w</sup>əlu-ha, ka-yšərbu, ka-ytfərtəṭ l-ʕərs. (47) w şafi – sbəʕ yyam ka-tkun kəmməlat. *that's it.*

### Gloss

(1) Hello – as for our village, we're close to the city; it's going to be 7 kilometers or so away at most. (2) As far as [our] weddings go, they last seven days. (3) On the first day... they bring the... as we say, the *adoul* [justice], who writes up a [marriage] contract between the groom and bride – (4) normally it's just close family who's in attendance, there's not anyone distant or from outside [the family].

(5) As for the first day, they write the contract. (6) As for the second day, it's known as the party day. (7) The bride's entire family comes, the bride gets dressed up, she's going to get made up, on the same note the groom's relatives wear traditional outfits or, it depends... (8) they [the bride and groom] sit in this special chair in this big hall, music is playing, people are dancing... (9) normally we have... it's just women. The men are outside. It's just women and the groom. (10) So... the groom is surrounded by women.

(11) The women dance, sing, the day goes by wonderfully, that first [dance] starts off with the bride and groom, it goes by normally... (12) the next day, the groom goes down to the city, him and... him and the *wezir*. (13) The *wezir* is... over there you all say “the best man.” (14) The *wezir* takes him to the city and keeps him occupied there. (15) As for the bride, she does this tradition called *tahemmamt*. (16) Because in the countryside there's no *hammam* like we have in the city, like, public steam house, [for] *tahemmamt* they make like... a hut... (17) they build it with bamboo reeds and all that, and fill it with hot water so it'll warm up. (18) The bride goes in there, washes herself, and comes out. (19) The day goes by normally; she spends it with... naturally, she's with her husband's family, she spends the day normally and it's over.

(20) The next day, which is the fourth day, the groom comes. (21) She gets up in the morning like normal, gets ready, and awaits her husband, who will be there in the afternoon. (22) Her husband comes in the afternoon, him and his best man, him and the *wezir*. (23) When he comes, he brings three things. Anything he brings to her, he [also] brings to her mother, and to his mother. (24) Meaning if he brings her slippers, he brings his mother slippers and brings her mother slippers. (25) He brings almonds, brings walnuts, brings dates, and brings a single fish – like, big fish. (26) When he comes, the bride goes out and meets him. (27) She goes out and meets him, and she's wearing a



*tekchita*<sup>35</sup>, she takes her gifts and he starts taking out what he brought, they take it out and hand it to her. (28) As she's taking her gifts, the last thing that he pulls out is that fish. (29) She takes one stick, he takes another, and it gets figured out who's going to clean that fish, clean it of its scales. (30) Whoever wins the right to clean the fish, they're the one who's going to be in control of marital life. (31) Even though it's not really fair that the woman is carrying a bunch of things, and she has to hold the gifts and still not fall over, (32) but even still she tries to compete whereas the man isn't carrying anything. (33) Normally the man wins. (34) OK, anyway once he wins that people are happy and distribute [the food]. (35) Those almonds, and those walnuts and the dates that he brought – they distribute them among the guests, and they eat and they laugh.

(36) The next day, things just go normally. (37) The next day [after that] is called “belt day.”<sup>36</sup> (38) The groom's brother comes, [...] – he comes and calls for the bride, and sets out a wooden basin, he fills it with water and puts a bangle made of silver in it. (39) The bride puts her foot in the middle of the bowl of water and on top of the bangle. And she calls over three... bachelors – like, single guys. (40) Three bachelors, and she gives them the water that was in the basin to drink. (41) Once she's given the man a drink, like the last guy, once she's given him a drink she has to flee from her room because if she stays there surrounded they're going to beat her up with a stick. (42) So like, she has to make them drink then run away. When she runs away, that's it and everything goes wonderfully, and that fifth day<sup>37</sup> is over.

(43) The next day she'll get up and do housework, with the logic that she's now become a woman and is being depended upon there. (44) In the afternoon, they bring her a wedding platter. (45) In the wedding platter there's normal couscous, but there's also chocolate, there are eggs, there's gum, there's... the wedding platter is [full of] sweets. (46) They bring her two platters, they dance, they eat it, they drink, and the wedding comes to a close. (47) That's all – seven days have passed. That's it.

---

<sup>35</sup> A dress-like garment worn on special occasions.

<sup>36</sup> This refers to a tradition, not discussed here, where the bride wears a special colored belt throughout the wedding week. On the day she takes it off she is considered officially married.

<sup>37</sup> In the light of the rest of the narration, the speaker was probably actually referring to the sixth day.

## References

- Abdel-Massih, E. T. (1971). *A Reference Grammar of Tamazight; a Comparative Study of the Berber Dialects of Ayt Ayache and Ayt Seghrouchen*. Ann Arbor: Center for Near Eastern and North African Studies, University of Michigan.
- Aguadé, J. (2010). On vocalism in Moroccan Arabic dialects. In *The Arabic Language Across the Ages* (pp. 95–105). Reichert Verlag.
- Belyayeva. (1997). Definiteness Realization and Function in Palestinian Arabic. In *Perspectives on Arabic Linguistics: Papers from the Annual Symposium on Arabic Linguistics, Salt Lake City, 1996* (pp. 47–67). John Benjamins Publishing.
- Bentahila, A., & Davies, E. E. (1983). The syntax of Arabic-French code-switching. *Lingua*, 59(4), 301–330.
- Boumans, L. (Ed.). (1998). *The Syntax of Codeswitching: Analysing Moroccan Arabic/Dutch Conversation*. Tilburg Univ Press.
- Brunot, L. (1949). Emprunts dialectaux arabes à la langue française dans les cités marocaines depuis 1912. *Hespéris*, 36, 347–430.
- Brustad, K. (2000). *The Syntax of Spoken Arabic: A Comparative Study of Moroccan, Egyptian, Syrian, and Kuwaiti Dialects*. Washington, D.C: Georgetown University Press.
- Caubet, D. (1983). Quantification, Negation, Interrogation: Les Emplois De La Particule Si En Arabe Marocain. *Arabica*, 30(3), 227–245. doi:10.1163/157005883X00526

- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and Topic* (pp. 159–182). New York Academic Press.
- Chafik, M. (1999). *al-Darijah al-Maghribiyah : majal tawarud bayna al-Amazighiyah wa-al-`Arabiyyah*. al-Rabat: Akadimiyat al-Mamlakah al-Maghribiyah.
- Chekayri, A. (2011). *An Introduction to Moroccan Arabic and Culture* (Pap/DVD Bl.). Georgetown University Press.
- Chtatou, M. (1997). The influence of the Berber language on Moroccan Arabic. *International Journal of the Sociology of Language*, 1997, 101–118.
- Corriente, F. (1977). *A grammatical sketch of the Spanish Arabic dialect bundle*. Madrid: Instituto Hispano-Arabe de Cultura.
- Corriente, F. (2008). *Dictionary of Arabic and Allied Loanwords: Spanish, Portuguese, Catalan, Galician and Kindred Dialects*. BRILL.
- Croft, W. (1990). *Typology and Universals*. Cambridge University Press.
- Cukor-Avila, P. (2000). Revisiting the Observer's Paradox. *American Speech*, 75(3), 253–254.
- De Mulder, W., & Carlier, A. (2011). The grammaticalization of definite articles. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Grammaticalization*. Oxford University Press.
- Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht : Boston: Kluwer Academic.

- Destaing, E. (1937). *Textes arabes en parler des Chleuhs du Sous Maroc*. Paris: P. Geuthner.
- Dorian, N. (1978). The fate of morphological complexity in language death: evidence from East Sutherland Gaelic. *Language*, 54, 590–609.
- Dryer, M. S. (1989). Article-noun order. *Chicago Linguistic Society*, 25, 83–97.
- El Aissati, A. (2006, 2009). Berber Loanwords. In (K. Versteegh, Ed.) *Encyclopedia of Arabic Language and Linguistics*. Leiden; Boston: Brill.
- Erwin, W. M. (1963). *A short reference grammar of Iraqi Arabic*. Georgetown University Press.
- Gazzah, M. (2008). *Rhythms and Rhymes of Life: Music and Identification Processes of Dutch-Moroccan Youth*. Amsterdam University Press.
- Givón, T. (1978). Definiteness and referentiality. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language: Vol. 4, Syntax* (pp. 291–330). Stanford University Press.
- Golston, C., & Wiese, R. (1995). Zero morphological and constraint interaction: subtraction and epenthesis in German dialects. In G. E. Booij & J. van Marle (Eds.), *Yearbook of Morphology*. Springer.
- Greenberg, J. H. (1978). How does a language acquire gender markers? In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Word Structure* (Vol. 3, pp. 47–82). Stanford: Stanford University Press.

- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, 69(2), 274–307.  
doi:10.2307/416535
- Hachimi, A. (2007). Becoming Casablancon: Fessis in Casablanca as a case study. In *Arabic in the City: Issues in Dialect Contact and Language Variation* (pp. 97–122). Hoboken: Taylor & Francis.
- Hardy, H. K., & Montler, T. R. (1988). Alabama Radical Morphology. In W. Shipley (Ed.), *In Honor of Mary Haas: From the Haas Festival Conference On Native American Linguistics* (pp. 379–409). Berlin: Walter de Gruyter.
- Harrell, R. S. (1962). *A Short Reference Grammar of Moroccan Arabic*. Washington: Georgetown University Press.
- Harrell, R. S. (1966). *A dictionary of Moroccan Arabic: Arabic-English*. Washington: Georgetown Univ. Pr.
- Haut-Commissariat au Plan du Royaume du Maroc. (2004). Recensement général de la population et de l’habitat 2004. *Site institutionnel du Haut-Commissariat au Plan du Royaume du Maroc*. Retrieved May 13, 2012, from [http://www.hcp.ma/Recensement-general-de-la-population-et-de-l-habitat-2004\\_a633.html](http://www.hcp.ma/Recensement-general-de-la-population-et-de-l-habitat-2004_a633.html)
- Heath, J. (1984). Language Contact and Language Change. *Annual Review of Anthropology*, 13, 367–384.
- Heath, J. (1987). *Ablaut and Ambiguity: Phonology of a Moroccan Arabic Dialect*. Albany: State University of New York Press.

- Heath, J. (1989). *From code-switching to borrowing: foreign and diglossic mixing in Moroccan Arabic*. Kegan Paul International.
- Heath, J. (1997). Moroccan Arabic Phonology. In *Phonologies of Asia and Africa: (including the Caucasus)* (pp. 205–217). Eisenbrauns.
- Heath, J. (2002). *Jewish and Muslim dialects of Moroccan Arabic*. Routledge.
- Heine, B. (2012). On polysemy copying and grammaticalization in language contact. In C. Chamoreau & I. Léglise (Eds.), *Dynamics of Contact-Induced Language Change* (pp. 152–166). Walter de Gruyter.
- Heine, B., & Kuteva, T. (2003). On contact-induced grammaticalization. *Studies in Language*, 27(3), 529–572.
- Ingham, B. (1994). *Najdi Arabic: central Arabian*. John Benjamins Publishing Company.
- International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge University Press.
- Ionin, T. (2006). This is definitely specific: Specificity and Definiteness in Article Systems. *Natural Language Semantics*, 14, 175–234.
- Jastrow, O. (2005). Uzbekistan Arabic: A Language Created by Semitic-Iranian-Turkic Linguistic Convergence. In É. Á. Csató, B. Isaksson, & C. Jahani (Eds.), *Linguistic convergence and areal diffusion: case studies from Iranian, Semitic and Turkic* (pp. 133–139). Psychology Press.
- Kaufman, S. (1974). *The Akkadian Influences on Aramaic*. Chicago: University of Chicago Press.

- Khan, G. A. (1984). Object Markers and Agreement Pronouns in Semitic Languages. *Bulletin of the School of Oriental and African Studies, University of London*, 47(3), 468–500.
- Kossmann, M. (2007). Berber Morphology. In *Morphologies of Asia and Africa* (pp. 429–446). Eisenbrauns.
- Kossmann, M. (2012). Berber. In Z. Frajzyngier & E. Shay (Eds.), *The Afroasiatic Languages*. Cambridge University Press.
- Kossmann, M., & Stroemer, H. J. (1997). Berber Phonology. In *Phonologies of Asia and Africa: (including the Caucasus)* (pp. 461–475). Eisenbrauns.
- Kusters, W. (2003). The Fate of Complex Languages: Classical Arabic and Old Norse in the Age of Globalisation. *Nordlyd : Tromsø University Working Papers on Language & Linguistics / Institutt for Språk og Litteratur, Universitetet i Tromsø*, 31(2), 275–289.
- Labov, W. (1972). *Sociolinguistic Patterns*. University of Pennsylvania Press.
- Maas, U. (2000). Moroccan: a language in emergence. In *Arabic as a minority language* (pp. 385–404). Walter de Gruyter.
- Maas, U. (2011). *Marokkanisches Arabisch. Die Grundstrukturen*. LINCOM publishers.
- Maas, U., & Procházka, S. (2012). Moroccan Arabic in its wider linguistic and social contexts. *STUF - Language Typology and Universals*, 65(4), 329–357.
- Marçais, P. (1977). *Esquisse grammaticale de l'arabe maghrébin*. Librairie d'Amérique et d'Orient.

- Marouane, M. (2009). Gradual Integration of Moroccan Arabic Loanwords in Ayt Souab Tashlhit. *Asinag*, 2, 111–129.
- Matar, `Abd al-`Aziz. (1966). *Lahn al-`ammah fi daw al-dirasat al-Lughawiyah al-hadithah*,. al-Qahirah: al-Dar al-Qawmiyah lil-Tiba`ah wa-al-Nashr.
- Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4), 530–555.
- Moscoso García, F. (2003). *El Dialecto Árabe De Chauen (Norte De Marruecos): Estudio Lingüístico Y Textos*. Cádiz, Spain: Universidad de Cádiz, Área de Estudios Árabes e Islámicos.
- Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes* (1st Edition.). Oxford University Press, USA.
- Nortier, J. (1990). *Dutch-Moroccan Arabic code switching among Moroccans in the Netherlands*. Foris.
- Nortier, J. (1995). Code switching in Moroccan Arabic/Dutch vs. Moroccan Arabic/French language contact. *International Journal of the Sociology of Language*, 112, 81–95.
- Owens, J. (1985). The Origins of East African Nubi. *Anthropological Linguistics*, 27(3), 229–271.
- Owens, J. (2005). Bare forms and lexical insertions in code-switching: A processing-based account. *Bilingualism: Language and Cognition*, 8(01), 23–38.
- Owens, J. (2006). *A Linguistic History of Arabic*. Oxford: Oxford University Press.



- Pat-El, N. (2009). The Development of the Semitic Definite Article: A Syntactic Approach. *Journal of Semitic Studies*, 54(1), 19–50. doi:10.1093/jss/fgn039
- Pat-El, N. (2010). The origin and function of the so-called correlative in Classical Syriac. *Folia Orientalia*, 45-46, 125–130.
- Pereira, C. (2011). Arabic in the North African Region. In S. Weninger (Ed.), *The Semitic Languages : An International Handbook* (pp. 954–969). Berlin: De Gruyter Mouton.
- Post, R. (2013). Arancia fil Internet: A syntactic comparison of written Arabic-French code-switching. Presented at the Jil Jadid Conference in Arabic Literature and Linguistics, Austin, TX.
- Ratcliffe, R. (2005). Bukhara Arabic: A Metatypized Dialect of Arabic in Central Asia. In *Linguistic convergence and areal diffusion: case studies from Iranian, Semitic and Turkic* (pp. 141–151). Psychology Press.
- Rubin, A. D. (2005). *Studies in Semitic grammaticalization*. Winona Lake, Ind: Eisenbrauns.
- Sayahi, L. (2011). Spanish in Contact with Arabic. In nuel Díaz-Campos (Ed.), *The Handbook of Hispanic Sociolinguistics* (pp. 473–489). Wiley-Blackwell.
- Stroomer, H. (2001). *An Anthology of Tashelhiyt Berber Folktales (South Morocco): Berber-English*. (H. Stroomer, Ed.) (1st ed.). Rüdiger Köppe.
- Stroomer, H. (2008). Three Tashelhiyt Berber Texts from the Arsène Roux Archives. In A. Lubotsky, J. Schaeken, & J. Wiedenhof (Eds.), *Evidence and Counter-*

- Evidence: Essays in Honour of Frederik Kortlandt. Volume 2: General linguistics* (pp. 389–397). Rodopi.
- Thomason, S., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
- Tilmatine, M. (2011). Berber and Arabic Language Contact. In *The Semitic Languages : An International Handbook* (pp. 1001–1014). Berlin: De Gruyter Mouton.
- Tura, S. S. (1986). Definiteness and Referentiality in Turkish Nonverbal Sentences. In D. I. Slobin & K. E. Zimmer (Eds.), *Studies in Turkish Linguistics* (pp. 165–194). John Benjamins Publishing.
- Watson, J. (2011). Arabic Dialects (general article). In *The Semitic Languages : An International Handbook* (pp. 851–895). Berlin: De Gruyter Mouton.
- Weinreich, U. (1953). *Languages in contact : findings and problems*. New York: Weinreich.
- Winford, D. (2003). *An Introduction to Contact Linguistics* (1st ed.). Wiley-Blackwell.
- Wright, W. (1896). *A Grammar of the Arabic Language* (Third Edition., Vols. 1-2, Vol. 1). Cambridge: The University Press.
- Youssi, A. (1992). *Grammaire et lexique de l'arabe marocain moderne*. Casablanca (Maroc): Wallada.
- Youssi, A. (1995). The Moroccan triglossia: facts and implications. *International Journal of the Sociology of Language*, 1995(112), 29–44.

## **Vita**

Michael Lee Turner grew up in Elizabeth, City North Carolina. After graduating from Northeastern High School in Elizabeth City in 2004, he enrolled at the University of North Carolina at Chapel Hill, from which he graduated in 2008 with a Bachelor of Arts degree. In the following years Michael served as a Peace Corps Volunteer in rural southern Morocco, working as a teacher and activities coordinator in a local youth center and completing his assignment in 2010. In August 2011 he enrolled as a graduate student at the University of Texas at Austin, where his interests include Arabic and Berber dialectology and developing materials for second-language learners.

Email: [mlturner@utexas.edu](mailto:mlturner@utexas.edu)

This thesis was typed by the author.