Geminate representation in Arabic

Stuart Davis and Marwa Ragheb
Indiana University

Arabic dialects are characterized by the occurrence of geminate consonants in word-medial and word-final position. This article relates the patterning of Arabic geminates to the on-going controversy in phonological theory concerning the representation of geminate consonants. Two views are contrasted: the prosodic length analysis of geminates whereby a geminate is underlyingly a single consonant phoneme linked to two C-slots, and the moraic weight representation where a geminate is underlyingly a single consonant linked to a mora. We specifically argue that the patterning of geminate consonants in Arabic dialects largely supports the moraic weight representation. Evidence comes from phenomena such as the patterning of word-final geminates, the behavior of geminates with respect to stress, geminates in loanwords, and geminates in first language acquisition. We show that each of these phenomena supports the moraic weight representation of geminates.

1. Introduction

Arabic dialects are characterized by geminate consonants. These appear in intervocalic and word-final position.[1] Less common are dialects that allow for geminate consonants in word-initial position. The examples in (1) illustrates geminate consonants in Cairene Arabic where minimal pairs are shown contrasting singleton consonants and their corresponding geminates. The data in (1a–f) illustrate

1. We are not aware of Arabic dialects that lack geminate consonants, including word-final ones. Even peripheral dialects such as Maltese, Nigerian and Uzbeki Arabic have geminates. The Creole Nubi Arabic of Uganda is reported in Wellens (2005) to have geminates only marginally (and not word-finally) since in that variety most (not all) historical words with geminates have degeminated. Our discussion on geminates does not pertain to creole varieties. We note that while word-final geminates are common across Arabic dialects, they are unusual in other Semitic languages.

contrast in intervocalic position whereas (1g–h) shows the contrast with a final geminate. Cairene Arabic lacks word-initial geminates.

(1) Geminates in Cairene Arabic (stress is indicated by the acute accent marker)

a.	[kásar] 'he broke'	b.	[kássar] 'he smashed'
c.	[bátˁal] 'hero'	d.	[bátˁtˁal] 'he quit (something)'
e.	[ħamá:m] 'pigeons'	f.	[ħammá:m] 'bathroom'
g.	[ʔámal] 'hope'	h.	[ʔamáll] 'more/most boring'

The patterning of geminate consonants has been a somewhat neglected topic in Arabic phonology. While there has been important discussion on the role of geminates with respect to particular issues in specific dialects, such as Watson's (2002) discussion on geminates in the stress system of San'ani Arabic and Farwaneh's (2009) discussion of affixes triggering gemination in certain dialects, there has been less work that examines the phonology of geminates across a range of phenomena in a variety of dialects. Moreover, there has been little work that specifically relates the role of Arabic geminates to the on-going controversy within phonological theory regarding the representation of geminate consonants, where there are two major contrasting views: a prosodic length representation where a phoneme is underlyingly linked to two C-slots (see especially Ringen and Vago 2011) and a weight representation where a geminate is a phoneme that is underlyingly linked to a mora (as in Hayes 1989). Our primary goal in this article is to examine the phonological patterning of geminate consonants in Arabic dialects and bring the Arabic data to bear on the issue of geminate representation. We will specifically argue that the patterning of geminate consonants in Arabic largely supports a moraic weight representation of geminates. Evidence will come from geminate patterning in a variety of dialects, but especially from Cairene Arabic, and will be exclusively phonological. We will consider the following: the patterning of word-final geminates, the behavior of geminates with respect to stress, geminates in loanwords, and geminates in first language acquisition. We will show that each of these phenomena supports a moraic representation of geminates. Before presenting data showing the phonological patterning of geminates in Arabic, we first discuss in more detail the different representations that have been proposed for geminate consonants.

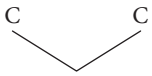
2. Background – The phonological representation of geminates

Within current work on phonological theory there is an on-going debate about the representation of geminate consonants. We will focus on two major views: the

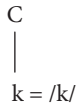
prosodic length representation and the moraic weight representation.[2] As discussed in Davis (2011), the prosodic length analysis of geminates goes back to Leben (1980) who posited an autosegmental representation of geminates in which a single phoneme is linked to two C-slots on a skeletal tier that encodes the prosody of the word in terms of C-slots and V-slots. Under this view, a geminate consonant would be represented as in (2a) while a nongeminate would be represented as a phoneme linked to a single C-slot as in (2b). In (3) we show the CV representation of the Arabic word [kassar] 'he smashed' displaying an intervocalic geminate. It should be understood that such a geminate comprises both the coda of the first syllable and the onset of the second.

(2) Prosodic length analysis of geminates

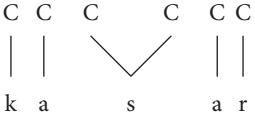
a. Geminate in UR



b. Single consonant in UR



(3) CV-tier representation of [kassar] 'he shattered'



The prosodic length analysis of geminates has been argued for in such works as Clements and Keyser (1983), Levin (1985) who uses an X-tier rather than a CV tier, Tranel (1991), Hume et al. (1997), and most recently by Ringen and Vago (2011).

A very different view of the representation of geminates is that posited by Hayes (1989) in which geminate consonants are considered to have inherent weight. This weight approach to geminates is couched within the theory of moraic phonology as developed in Hayes (1989), which characterizes the prosodic tier as being moraic rather than segmental as was shown in (3). Specifically, in Hayes's theory of moraic phonology, a short vowel is underlyingly monomoraic while a long vowel is bimoraic; a geminate consonant differs from a short consonant in that a geminate is underlyingly linked to a mora as in (4a) (where μ indicates a mora) whereas a nongeminate as in (4b) lacks a mora underlyingly. In (5) we show

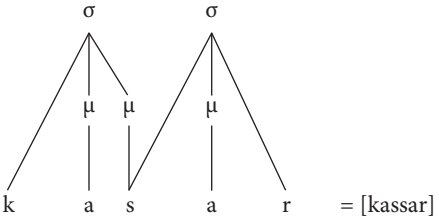
2. A third view that we will not discuss is the two root node theory of geminates advanced by Selkirk (1990). This theory is more applicable to cases of phonologically derived geminates as in the pronunciation of [ʔiʃʃams] 'the sun' from underlying /ʔil-ʃams/. Since such phonologically derived geminates are not our focus, we will not be considering the two root node theory. For a composite view of geminate representation see Curtis (2003).

the moraic representation of the Arabic word [kassar] 'he smashed' displaying an intervocalic geminate.[3]

(4) Moraic (weight) representation of geminates (Hayes 1989)

a. Geminate in UR b. Single consonant in UR

(5) Surface syllabification with moraic structure (σ = syllable)



In moraic theory, the formal distinction between a heavy syllable and a light syllable is that a heavy syllable is bimoraic (or greater) whereas a light syllable is monomoraic. Given this distinction, we observe in (5) that the first syllable is heavy (bimoraic) while the second is light. The diagram in (5) reflects the standard understanding of how intervocalic geminates syllabify in languages that have them: the geminate syllabifies in a way that adds weight to the first syllable where it is part of the coda, but does not add weight to the second syllable where it is part of the onset.[4]

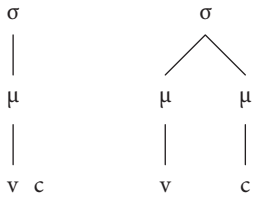
While in Hayes's (1989) version of moraic theory a geminate is underlyingly moraic, the weight of a (nongeminate) coda consonant is language specific. Such a coda can be made moraic by a rule (or constraint) that requires codas to be

3. As discussed in Davis (2011), there is an important difference between the moraic analysis in (5) and the prosodic weight analysis in (3). The prosodic length analysis is typically connected to phonetic duration. So, as seen in (3), even an initial onset consonant has its own C-slot. This contrasts with the moraic analysis shown in (5). A mora can be considered a unit of syllable weight, not necessarily tied to phonetic duration. Since the initial consonant in (5) does not affect syllable weight (e.g. its presence or absence does not influence weight-sensitive rules such as stress or vowel shortening), then it does not have a mora.

4. Topintzi (2008) points to the Micronesian language Marshalese as a possible rare exception to the observation regarding the syllabification of intervocalic geminates where she claims that intervocalic geminates syllabify entirely within the onset.

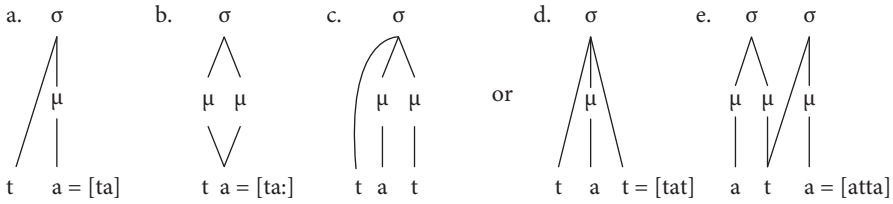
moraic. Hayes referred to this rule as 'Weight-by-Position' and his formulation of it is given in (6).

(6) Weight-by-Position (Hayes, 1989)



This rule, which adds a mora to a coda consonant, applies in many Arabic dialects, but usually not in word-final position. The effect of this rule can be seen in the stress pattern of many dialects where it is common for stress to be attracted to a penultimate syllable that ends in a consonant, as in [ka.táb.ti] 'you (f.) wrote' but not to a final closed syllable, for example, [ká.tab] 'he wrote'. That is, a word-final consonant is extrametrical in many dialects. In (7) we show examples of mora structure on different syllable types noting that Weight-by-Position has applied in (7c) and that geminates are moraic as in (7e).

(7) Surface syllabification with moraic structure



With this as background, we will now turn to the phonological evidence from Arabic dialects that supports the moraic weight representation of geminates (4a), as opposed to the length representation in (2a).

3. Arabic evidence for the moraic representation of geminates

In this section we will present evidence from four different phonological phenomena that strongly supports the moraic weight representation of geminate consonants in Arabic: the patterning of word-final geminates, the behavior of geminates with respect to stress, geminates in loanwords, and geminates in first language acquisition. For each phenomenon we will compare the length analysis of geminates in (2a) with the moraic weight analysis in (4a). We first consider the patterning of word-final geminates.

3.1 Word-final geminates

Virtually all Arabic dialects are characterized by the occurrence of word-final geminate consonants. A number of recent studies, such as Al-Tamimi et al. (2010), show that final geminates are perceptually distinct from their singleton counterparts. While we will not address specific phonetic issues regarding Arabic geminates in this article, we do note that word-final stops in Arabic dialects such as Cairene Arabic tend to be released so that the durational cue of a geminate stop is often preserved even in word-final position.

With respect to the phonological patterning of geminates, the length analysis in (2a) tacitly assumes that geminates should have a patterning that is similar to a sequence of two consonants since under that analysis a geminate consonant is linked to two C-slots. Crucially, the weight representation in (4a) does not make such an assumption. While in intervocalic environments it would often be hard to distinguish the two representations since Arabic dialects allow for both geminates and word-medial consonant clusters, there are dialects that have word-final geminates but lack word-final clusters. These dialects are important for distinguishing between the representations in (2a) and (4a) because they show that a geminate can pattern differently than a consonant cluster. One such dialect is Hadhrami Arabic as spoken in the town of Ghayl Bawazir near the south coast of Yemen (Bamakhramah 2009, personal communication). Consider the representative data in (8).

(8) Hadhrami Arabic (Bamakhramah 2009, personal communication).

a.	/gird/ – [gírid] 'monkey'	([gírd-i] 'my monkey')
b.	/bint/ – [bínit] 'girl'	([bínt-i] 'my girl')
c.	[rább]	'Lord'
d.	[ʔaxáff]	'lighter/lightest'

The data items in (8c)–(8d) show that Hadhrami Arabic allows for word-final geminates. However, as (8a)–(8b) illustrate, lexical items that underlyingly have a final consonant cluster do not surface with that final cluster intact; rather, vowel epenthesis breaks up the potential word final cluster. The examples in (8a)–(8b) are compelling since the final clusters in these lexical items have falling sonority; such final clusters are permitted to surface in many other dialects.

The vowel that surfaces between the final two consonants in (8a)–(8b) is indeed epenthetic; evidence for the epenthetic status of this vowel comes from the possessive form given in parentheses in (8a)–(8b) where no vowel occurs between these two consonants, as seen in the example [bint-i] 'my girl' which preserves the underlying consonant cluster. Further, one cannot argue that the underlying forms of the lexical words in (8a)–(8b) contain a vowel between the two final consonants (e.g. /binit/ 'girl' in (8b)) because Hadhrami Arabic words that do underlyingly

have the pattern /CVCVC/ surface as [CCVC] with the first vowel deleted. (See Bamakhramah (2009) for details.) Consequently, we observe from (8) that Hadhrami Arabic does not permit words to surface with a final consonant cluster though final geminates are allowed.

Similar to Hadhrami Arabic is Baghdadi Arabic where Blanc (1964) reports that Muslim and Christian varieties avoid final consonant clusters, but final geminates are common as indicated in (9).

(9) Baghdadi Arabic – Blanc (1964)

a. /ward/ – [wáred] 'flowers'	b. /sabt/ – [sábet] 'Saturday'
c. [sádd] 'he shut'	d. [yendázz] 'he will be sent'

The occurrence of final geminates without the occurrence of final consonant clusters is unexpected given the length representation in (2a). Although languages may restrict final clusters to those of level (or nonrising) sonority, it is surprising under (2a) for a language to allow for final geminates without allowing any other clusters. The expectation is for geminates to pattern like other consonant sequences. This expectation does not arise under the moraic weight representation in (4a) in which geminates are not represented as linked to two C-slots. Moreover, as seen by the specific examples in (8d) and (9d), a final geminate always attracts word stress to the final syllable. This is consistent with the weight representation of geminates in (4a) as we will discuss in the following subsection.

3.2 Word stress and geminate consonants

Word stress in most Arabic dialects is quantity-sensitive. That is, stress is attracted onto a heavy syllable, specifically one with a long vowel or ending in a consonant (as long as it is not at the end of the word). For example, if we focus on syllables ending in a consonant in Cairene Arabic, a CVC syllable in penultimate position will attract stress as seen in (10a)–(10b), while such a closed syllable in final position does not do so as in (10c). However, if the final syllable ends in two consonants then stress is on that syllable as exemplified by the Cairene words in (10d)–(10e).[5]

5. Note that in Cairene Arabic, words can end in a sequence of two consonants with no restrictions on what those two consonantal phonemes can be. One subpattern of stress found in Cairene that we do not discuss here is the well-known preference to stress the penultimate syllable when the antepenultimate is heavy as in [mar.tá.ba] 'mattress', [mad.rá.sa] 'school' and [kas.sá.ru] 'he smashed it'. This should be compared with words like [ká.ta.bu] 'he wrote it' which take antepenultimate stress when both the penultimate and antepenultimate syllables are light. See Hayes (1995) for an analysis that references quantity-sensitivity and mora structure.

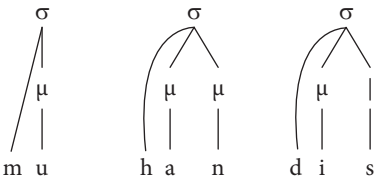
(10) Stress on some Cairene Arabic words

a.	[ka.táb.ti]	'you (f.) wrote'	b.	[mu.hán.dis]	'engineer'
c.	[ká.tab]	'he wrote'	d.	[ka.tábt]	'I wrote'
e.	[maf.hímʃ]	'he didn't understand'			

A standard analysis of the Cairene pattern shown here (Hayes 1995, among others) is that a coda consonant is normally moraic unless it is in word-final position, in which case it is extraprosodic.

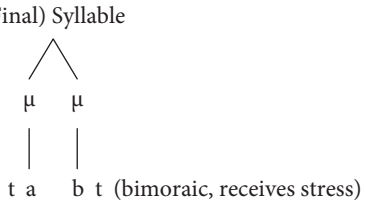
For example, the syllables in the word [mu.hán.dis] 'engineer' in (10b) would have the moraic representation shown in (11) where the closed penultimate syllable is bimoraic but the closed final syllable is monomoraic.

(11) Moraic representation of [mu.hán.dis] 'engineer'



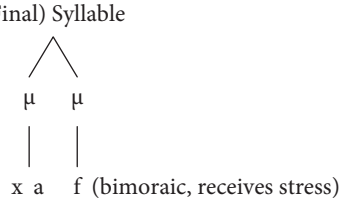
A final syllable ending in two consonants, such as [ka.tábt] 'I wrote' in (10d) would have the moraic representation shown in (12) where the first of the two consonants is moraic but the final is not. The final syllable receives stress since it is bimoraic.

(12) [ka.tábt] 'I wrote'



What (11) and (12) show is that the rule (or constraint) Weight-by-Position in (6), which normally makes a coda consonant moraic, does not apply to a word-final consonant. However, a word ending in a final geminate consonant always receives stress on the final syllable in Cairene Arabic (and in most other dialects). For example, the word [ʔa.xáff] 'lightest' that we saw in (8d) for Hadhrami Arabic also occurs in Cairene Arabic. In both dialects the word has final stress. In (13) we show the mora structure of the final syllable of [ʔa.xáff].

(13) Moraic structure of the final syllable of [ʔa.xáff] 'lightest'



Since we have seen that Weight-by-Position does not apply to a final consonant, the weight of a word final geminate can be understood as reflecting its inherent underlying weight as in (4a). This is perhaps clearer in dialects such as Hadhrami (8) and Baghdadi (9) that do not otherwise allow for final consonant clusters and in which final CVC syllables do not normally attract stress. The fact that a word-final geminate attracts stress to the final syllable in such dialects strongly supports the underlying weight analysis of geminates in (4a) given that consonant clusters are not normally allowed at the end of the word, and that a geminate clearly adds weight to the final syllable.

San'ani Arabic (Yemen) presents a different case in which syllables closed by a geminate (CVG) attract the stress even when other closed syllables (CVC) do not. This has been specifically observed by Watson (2002), though the discussion below is based on Davis (2011). Consider the data in (14) in which none of the words display a geminate. (In the transcription below [g*] represents a voiced post-velar dorsal fricative.)

(14) San'ani Arabic (Watson 2002: 81–82)

a.	da.rást	'I/you m.s. learnt'
b.	mak.tú:b	'letter'
c.	sá:.fa.rat	'she travelled'
d.	mig*.sá:.lih	'launderette'
e.	mi.gám.bar	'sitting'
f.	mád.ra.sih	'school'
g.	mak.tá.ba.ti:	'my library'
h.	lí.bi.sat	'she wore/put on'
i.	ká.tab	'he wrote'
j.	rá.ga.ba.tih	'his neck'

As seen in (14) word stress normally falls on one of the last three syllables of the word: it falls on a final CVCC or CVVC if there is one as in (14a–b); it falls on the rightmost nonfinal heavy syllable (CVC or CVV) up to the antepenultimate as in (14c–f); otherwise, stress falls on the leftmost CV syllable as in (14g–j). The data in general show that the word-final segment does not play a role in the computation

of weight so that the final syllable can only be stressed if it is bimoraic, absent the final segment. The word in (14g) illustrates that a CVC syllable in pre-antepenultimate position fails to attract stress. This suggests that Weight-by-Position (6), which assigns a mora to a coda consonant, is restricted to one of the last three syllables of the word. (That is, Weight-by-Position does not apply to a pre-antepenultimate syllable.)

Now let us consider the data in (15) illustrating words possessing geminate consonants. Note that words ending in a final geminate receive stress on that syllable (Janet Watson, personal communication), though we do not show such words here.

(15) San'ani Arabic (Watson 2002: 81–82) – stress on words with geminate consonants

a.	yi.ħíb.bu	'they m. love/like'
b.	mit.ʔáx.xi.ra:t	'late f. pl.'
c.	mu.sáj.ji.la.ti	'my recorder'
d.	dáw.wart	'I/you m.s. looked for'

The comparison between (15) and (14) indicates the priority of syllables closed by a geminate (henceforth, CVG) to receive stress (in comparison to other CVC syllables) in that CVG syllables always attract stress even when in pre-antepenultimate position as in (15c). The word in (14g), in contrast, shows that a CVC syllable does not receive stress in pre-antepenultimate position. The difference between CVG and CVC syllables in San'ani Arabic can be readily understood on the inherent weight analysis of geminates (4a). If a geminate is underlyingly moraic, it contributes weight to the syllable regardless of its location in the word. In contrast, it is not clear under the length analysis of geminates in (2a) why the (pre-antepenultimate) syllable closed by a geminate as in (15c) would pattern differently than the (pre-antepenultimate) syllable closed by the nongeminate in (14g) since both constitute a final consonant of a word-internal syllables ending in a C-slot. Weight-by-Position does not apply in (14g) since the initial syllable is in pre-antepenultimate position. Moreover, (15d) shows that a syllable closed by a geminate has priority for stress over other CVC syllables within the same word. Specifically, in (15d), the (penultimate) CVG syllable has a priority of stress over a final superheavy syllable and should be compared with (14b) where a penultimate CVC syllable is devoid of such priority. This comparison suggests that Weight-by-Position in San'ani Arabic only applies to words that would not otherwise have bimoraic syllables (CVG or CVV). That is, there is no necessity for Weight-by-Position to apply in (15d). The priority of syllables closed by a geminate in attracting stress strongly supports the moraic weight analysis of geminates as in (4a).

3.3 Geminates in loanwords (Cairene Arabic)

Loanword phonology has been an important focus of research during the past decade. However, there has been very little systematic phonological research on the loanword phonology of Arabic dialects especially as it relates to issues of syllable structure and gemination. Nevertheless, two recent (unpublished) studies highlight the role of syllable augmentation through vowel lengthening and gemination in loanwords, that of Abu-Guba (2011) for Jordanian Arabic and Reynolds (2011) for Cairene Arabic. Our focus will be on Cairene Arabic. In particular, we will consider loanwords in which the source word has stress on the final syllable. Such words often induce gemination of the final consonant when borrowed into Cairene. Our claim, then, is that this is consistent with the moraic representation of geminate consonants. We divide our loanword data into three groups. The first group in (16) comprises English monosyllabic words ending in a consonant cluster; the second group in (17a–e) consists of monosyllabic English words ending in a single consonant, and the third group in (18) contains French loanwords with final stress. Some of the data here have been discussed by Reynolds (2011).

(16) Monosyllabic English loanwords with final consonant clusters

	English	Cairene
a.	film	fílm
b.	bank	bánk
c.	prince	bi.ríns
d.	stable	ʔis. tˤábl
e.	clutch	kilátʃ

(17) Monosyllabic CVC and CCVC loanwords from English

	English	Cairene
a.	(seven) up	ʔabb
b.	book (purse)	búkk
c.	watt	wátˤtˤ
d.	L	ʔíll
e.	stock	ʔis.túkk
f.	elastic	ʔás.tik
g.	thermos	túr.mus

(18) French loanwords (source has final stress)

	French	Cairene
a.	manchette	manʃítt 'headline'
b.	claquette	kilakítt (as in [kilakítt ʔawwil marra] 'take one')
c.	crepe	kiríbb 'crepe'
d.	parachute	baraʃútt 'parachute'

In the words in (16), (17a–e), and (18) the monosyllabic word or the final syllable needs to have a stress when borrowed into Cairene Arabic since these words keep their source stress on the final syllable when borrowed. In order to fit into the Cairene Arabic stress pattern, a stressed final syllable must be bimoraic. While in (16) the final borrowed syllable is already bimoraic since it is CVCC (exactly like the structure shown in (12)), the final syllable of the loanwords in (17a–e) and (18) need to acquire a second mora in order to preserve the source language stress. The preference shown in (17a–e) and (18) is that when the source vowel of the final (C) CVC syllable is short, the second mora comes about through gemination of the final consonant. Thus, the final consonant of all the words in (17a–e) and (18) geminates so that the final syllable can be bimoraic. If the final consonant were not to geminate then the final syllable would be monomoraic (CVC), given that a word final (nongeminate) consonant does not add weight to the syllable; such a syllable is unstressable in Cairene Arabic. The data items in (17f–g) show that when a loanword has a final closed syllable (CVC) but where source stress is on a prior syllable, no final gemination occurs. That is, final gemination does not occur if the source word lacks final stress as shown by the examples in (17f–g). The strategy of final gemination for adding weight to the syllable is very consistent with a weight representation of geminates, as seen in (4a). While one could view this strategy as adding an extra C-slot to the end of the word in examples (17a–e) and (18) under the length analysis of geminates in (2a), this would still need to reference a weight representation of the final consonant so as to distinguish these loanwords from (17f–g) where final gemination does not occur. The loanword data is consistent with the moraic representation of geminates.

3.4 L1 acquisition of final clusters in Cairene Arabic

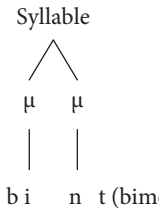
Further support for the moraic weight representation of geminates comes from first language acquisition data originally reported in the work of Ragheb (2010) and discussed by Ragheb and Davis (2010). These works provide an account of the acquisition of Cairene Arabic final consonant clusters by a child referred to as MG who was growing up in a monolingual setting in Cairo. Data were gathered from him at the age of 2 years 8 months over a one-month period, using pictures eliciting responses from a pre-designed word list (focused on consonant clusters) and from spontaneous speech. As far as we are aware, this was the first study that examined the acquisition of final consonant clusters for any Arabic dialect. A main finding is that if a final cluster is not pronounced target appropriately, then it will be pronounced as a geminate consonant. Consider the data in (19) from Ragheb and Davis (2010) where the target pronunciation is indicated in the first column

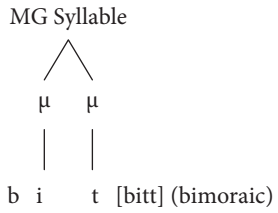
and MG's pronunciation is in the second column. (We ignore here the interesting occurrence of glottal replacement in onset position.)

(19) Examples from MG (2;8) acquiring Cairene Arabic in a monolingual setting

Target pronunciation	MG's pronunciation	Gloss
a. ward	wadd	Flowers
b. bint	bitt	Girl
c. kalb	kabb	Dog
d. ʃiribt	ʔitt	I drank
e. miʃtˤ	ʔitt	Comb
f. rigl/riglu/rigli	ʔill/ʔillu/ʔilli	Leg/his leg/my leg
g. naml	ʔall	Ants
h. ħabl	ʔall	Rope
i. nusˤsˤ	ʔusˤsˤ	Half
j. malħ	ʔall	Salt
k. ʔamħ	ʔamm	Wheat

From the data in (19) we can observe that MG employs a consistent strategy in words that end in a final cluster with all types of sonority profiles. As seen in (19a–h) he typically geminates the final consonant of the cluster, deleting the first one. However, if a pharyngeal consonant is a final member of the consonant cluster as in (19j–k), the first of the two consonants geminates. Note at this stage, as shown in (19i) (/nusˤsˤ/), MG correctly produces target word-final geminates (i.e., MG has already acquired geminate consonants). Geminates are acquired by MG early in the acquisition process, though it should be noted that MG did not have geminate pharyngeals when the data were collected. In Ragheb and Davis (2010) we discussed why the child geminates in a final cluster, rather than, for example, deleting one of the consonants without gemination, which commonly happens in the acquisition of English word final clusters. We concluded that by gemination, MG is able to preserve the prosodic structure (i.e., mora structure) of the target word without having to pronounce two adjacent consonants that have two different articulations. This is exemplified in (20).

(20) a. 

b. 

The strategy of word-final consonant gemination seen with MG can be understood as a means of preserving the prosodic moraic structure of the bimoraic final syllable in words that end in two consonants without the need for making two distinct consonantal gestures. While one could give an account of final gemination in terms of the two C-slot analysis of geminates as in (2a), the moraic representation of geminates in (4a) more directly captures the preservation of phonological weight between the target pronunciation and MG's actual pronunciation as shown in (20). Also, it is worth pointing out that gemination as a strategy in L1 acquisition for the pronunciation of final clusters has not been witnessed in languages such as English where children often delete consonants or insert vowels in final clusters (cf. McLeod et al. 2002) or even in a language like Turkish where a consonant in a final cluster will be deleted with subsequent compensatory lengthening in the preceding vowel (cf. Topbas and Kopkalli-Yavuz 2008). However, given that Arabic has final geminates that clearly add weight to the syllable, gemination as a replacement for final clusters in Arabic child phonology is not surprising.

4. Conclusion and remaining problems

In this paper we have shown that there is strong evidence for the moraic weight representation of Arabic geminates. We believe that each type of evidence presented in Section 3 that included data from the patterning of word-final geminates, stress, loanwords, and acquisition is most consistent with this representation (4a) and more problematic for the length representation of geminates in (2a). Moreover, when all the evidence is taken together they converge on an analysis in which geminate consonants have a moraic weight representation. Nonetheless, there are at least two potentially problematic issues for the moraic representation of geminates that come up in certain dialects. One issue concerns potential cases where geminates syllabify entirely within a word-internal coda and the other issue is the neglected topic of the phonological patterning of word-initial geminates in dialects that have them.

With respect to the first issue, the Palestinian Arabic dialect described by Abu Salim (1980) and mentioned in Rose (2000) is said to distinguish a singleton coda

consonant in a word like [bit.na] 'our house' from an internal coda geminate in a word like [sitt.na] 'our grandmother'. On a strictly moraic view of geminates, the coda [t] in [bit.na] which acquires a mora by the rule of Weight-by-Position (6) is representationally indistinguishable from the mora on the geminate in [sitt.na] which has an underlying mora. Unless a CV length tier is assumed as in (3), there is no obvious way to distinguish the two cases. However, Farwaneh (2009) notes that forms similar to /sitt-na/ undergo degemination in Levantine dialects, as in [kull] 'all' vs. [kulhum] 'all of them'. We suspect that this is the case too for Palestinian Arabic so that there is really no surface length contrast between the [t] in /bitna/ and the geminate in /sittna/, but a phonetic study would need to be undertaken. It should be noted that in many dialects this situation of a possible geminate/singleton contrast entirely within a coda does not arise. In Cairene Arabic, for example, underlying /sittna/ would surface as [sit.ti.na] with vowel epenthesis so that a word-internal geminate entirely within a coda is not permitted.

A more interesting problem for the nature of geminate representation in Arabic concerns dialects with initial geminates such as Moroccan Arabic and various Gulf and Levantine varieties. As far as we are aware, the weight properties of initial geminates have not been systematically explored in Arabic dialects (though see Muller (2001) for initial work on Moroccan Arabic). Our sense is that in dialects with initial geminates, stress is not attracted onto a syllable containing them. Kiparsky (2003), however, represents Arabic initial geminates as having an unsyllabified mora where the mora links to the word node (not the syllable node). Under Kiparsky's view, an initial geminate would add an extra mora to the word but not to the syllable. This would perhaps predict that initial geminates would not attract stress, but could play a role in certain other weight-sensitive processes such as a bimoraic minimal word requirement. Further, it should be noted that dialects that allow for initial geminates usually also allow for initial consonant clusters of all types. Such initial consonant clusters do not seem to add weight to the syllable either because they do not typically attract stress. Since all dialects have final geminates that are weight bearing, we suspect that in those dialects that have initial geminates, there is an asymmetry in that final geminates are underlyingly moraic while initial geminates are not. Further, we do not seem to find dialects that are, in a sense, the opposite of what is attested in (8) and (9) for Hadhrami and Baghdadi Arabic, namely a dialect with initial geminates but with no other initial consonant clusters. The lack of such dialects suggests that the nature of initial geminates is quite different (and distinct) from that of final geminates, perhaps reflecting different historical origins of initial versus final geminates. Initial geminates are almost always heteromorphemic derived phonologically by assimilation, and thus may have a different (i.e. nonmoraic) representation. We leave for future research issues

regarding initial geminates, their phonological patterning, and their implications for geminate representation in Arabic dialects.

Acknowledgements

This research has been partially supported by an Indiana University Grant-in-Aid for Research. The acquisition data reported in Section 3.4 was collected under Indiana University IRB study number 08–13285. A preliminary version of this paper was presented at the 25th meeting of the Arabic Linguistics Symposium held at the University of Arizona in March 2011. We thank the audience at that meeting for their comments. We would also like to thank Daniel Dinnsen, Samira Farwaneh, and Reem Khamis-Dakwar as well as the anonymous reviewers for their helpful suggestions.

References

Abu-Guba, M.N. (2011). Gemination in English loanwords in Jordanian Arabic. Paper presented at the International Conference on Phonetics and Phonology 2011, Kyoto, Japan, December 2011.

Abu Salim, I. (1980). Epenthesis & geminate consonants in Palestinian Arabic. *Studies in the Linguistic Sciences*, 10, 1–11.

Al-Tamimi, F., Abu-Abbas, K. H., & Tarawnah, R. (2010). Jordanian Arabic final geminates: An experimental clinical phonetic study. *Poznan Studies in Contemporary Linguistics*, 42, 111–125.

Bamakhramah, M. (2009). *Syllable structure in Arabic varieties with a focus on superheavy syllables*. Ph.D. dissertation, Indiana University.

Blanc, H. (1964). *Communal dialects in Baghdad*. Cambridge, Mass.: Harvard University Press.

Clements, G. N. & Keyser, S. J., (1983). *CV phonology*. Cambridge, Mass.: MIT Press.

Curtis, E. (2003). *Geminate weight: Case studies and formal models*. Ph.D. dissertation, University of Washington.

Davis, S. (2011). Geminates. In M. van Oostendorp, C. Ewen, E. Hume; & K. Rice (Eds.), *The Blackwell companion to phonology* Vol. 2 (pp. 873–897). Oxford: Wiley-Blackwell.

Farwaneh, S. (2009). Towards a typology of Arabic dialects: The role of final consonantality. *Journal of Arabic and Islamic Studies*, 9, 82–109.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry,* 20, 253–306.

Hayes, B. (1995). *Metrical stress theory*. Chicago: University of Chicago Press.

Hume, E., Muller J., & van Engelenhoven, A. (1997). Non-moraic geminates in Leti *Phonology*, 14, 371–402.

Kiparsky, P. (2003). Syllables and moras in Arabic. In C. Féry & R. van de Vijver (Eds.), *The syllable in optimality theory* (pp. 147–182). Cambridge: Cambridge University Press.

Leben, W. (1980). A metrical analysis of length. *Linguistic Inquiry*, 11, 497–509.

Levin, J. (1985). *A metrical theory of syllabicity*. Ph.D. dissertation, MIT.

McLeod, S., van Doorn, J., & Reed, V. A. (2002). Typological description of the normal acquisition of consonant clusters. In Windsor, F., Kelly, L., & Hewlett, N. (Eds.), *Themes in clinical phonetics and linguistics* (pp. 185–200). Hillsdale, NJ: Lawrence Erlbaum.

Muller, J. (2001). *The phonology and phonetics of word-initial geminates*. Ph.D. dissertation, The Ohio State University.

Ragheb, M. (2010). The phonological acquisition of word-final consonant clusters in Cairene Arabic. Ms., Indiana University.

Ragheb, M. & Davis, S. (2010). The acquisition of word-final clusters in Cairene Arabic: An OT analysis. Paper presented at the Arabic Linguistics Symposium, Austin, Texas, April 2010.

Reynolds, E. (2011). Multiple scansion of Egyptian Arabic monosyllabic loanwords: Coda gemination in underweight syllables. Ms., Indiana University.

Ringen, C. & Vago, R. (2011). Geminates: heavy or long?. In C. Cairns & E. Raimy (Eds.), *handbook of the syllable* (pp. 155–169). Leiden: Brill.

Rose, S. (2000). Rethinking geminates, long-distance geminates, and the OCP. *Linguistic Inquiry*, 31, 85–122.

Selkirk, E. (1990). The two root node theory of geminates. *University of Massachusetts Occasional Papers*, 14, 123–171.

Topbas, S. & Kopkalli-Yavuz, H. (2008). Reviewing sonority for word-final sonorant+obstruent consonant cluster development in Turkish. *Clinical Linguistics and Phonetics*, 22, 871–880.

Topintzi, N. (2008). On the existence of moraic onsets. *Natural Language and Linguistic Theory*, 26, 147–184.

Tranel, B. (1991). CVC light syllables, geminates, and moraic theory. *Phonology*, 8, 291–302.

Watson, J. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Wellens, I. (2005). *The Nubi language of Uganda: An Arabic creole in Africa*. Leiden: Brill.