Poznań Studies in Contemporary Linguistics 44(2), 2008, pp. 155–177

doi:10.2478/v10010-008-0009-8

THE SYLLABLE SHAPE OF AL-AHSA DIALECT: AN OT PERSPECTIVE*

ABDULLAH H. ALJUMAH
Ypsilanti
aaljumah@hotmail.com

ABSTRACT

This work discusses the syllable structure of Al-Ahsa dialect of Arabic. Al-Ahsa dialect is a variation of Arabic spoken in the eastern province of Saudi Arabia. An OT (Prince and Smolensky 1993, 2004) framework is adopted and provides a complete analysis of the syllable structure found in the dialect under discussion. It is concluded that Al-Ahsa dialect prohibits initial consonant clusters but allows complex codas. The geminates, however, are allowed medially and finally.

KEYWORDS: Optimality Theory; syllable shape; Arabic; Al-Ahsa dialect; phonology.

1. Introduction

Classical Arabic is the formal language of the media and of instruction in the Arab world. A more accurate form of classical Arabic is found in the Holy Book, the *Quran*, which is the source of the most formal type of classical Arabic. Centuries ago, Arabs living in the western part of the Arabian Peninsula were known for their eloquent and highly dense form of language in regular conversations, prose and poetry. Then came the *Quran*, which is believed to be the actual words of God by Moslems, to challenge their eloquent language. Since then, no one has ever been able to speak in a manner similar to the words of the holy *Quran*. Then the accurate form of classic Arabic started to fade away and variations existed. One, among others, is Al-Ahsa dialect.

Al-Ahsa dialect is a variation of Arabic spoken in the eastern part of Saudi Arabia. It is well-known for vowel lengthening and for the existence of many loanwords. Unlike Moroccan Arabic, the most interesting fact about this dialect is that it inhibits

* I would like to thank the anonymous *PSiCL* reviewers for their fruitful comments on the earlier version of this paper. Also, I would like to thank professor Beverley Goodman for her lovely discussions and encouragements. Lastly, a special thank to my friend, Kathryn Lincoln, whose editing skills helped me spot a few punctuation mistakes and a few grammatical ones, which I would have overlooked otherwise.

initial consonant clusters and initial geminates; yet it allows both medial and final geminates; it also allows consonant clusters in the coda as long as they are moraic. It is argued here that complex edges are not allowed in Al-Ahsa dialect and the only way to satisfy complex is to have minor syllables, and the only way to satisfy Foot-Binarity is to morify the consonants occupying the minor consonants.

Thus, the overall aim of this paper is to investigate the syllable shape of Al-Ahsa dialect and provide a comprehensive analysis of the syllable shape within the framework of Optimality Theory, hereinafter referred to as OT (Prince and Smolensky 1993, 2004). Whenever applicable, a few remarks will be made with reference to a contrastive dialect, specifically, Casablanca Moroccan Arabic.

The paper is organized as follows: This section introduces the topic under discussion; the second section presents general background information relevant to the study; the third section deals with the inventory of phonemes in Al-Ahsa dialect and presents pertinent information regarding the dialect discussed; the forth section presents the data under investigation; the fifth section introduces the OT framework and presents the analysis of the data presented; finally, the last section includes the conclusion and future directions.

2. Review of literature

The Arabic language and its varieties have received an extensive analysis within the OT framework. Relevant to this study are three significant facts about Arabic. These three facts are discussed briefly.

First, complex onsets are prohibited in Arabic. Several researchers, namely McCarthy (2005), Archibald (2003), Gafos (2003), Edzard (2000) and Haddad (2005), all agree that complex onset is prohibited in many varieties of Arabic. Archibald (2003), in particular, asserts that Egyptian Arabic does not allow initial consonant clusters. Haddad (2005) also states that complex onset is not allowed in both Cairene Arabic and modern standard Arabic. In addition, Gafos (2003) argues that while geminates are allowed in Arabic, the onset does not consist of geminates and that no complex syllable onsets are permitted in Arabic. McCarthy (2005), however, argues that initial consonant clusters are possible due to syncope of certain perfective verbs of Arabic. The examples which McCarthy cited are found in classic Arabic and Moroccan Arabic.

Second, no word starts with a vowel in Arabic. Several studies, e.g. McCarthy (2005), Haddad (2005), and Gadoua (2000), have all pointed out that words in Arabic do not start with initial vowels. Haddad (2005) asserts that words in Cairene Arabic and standard Arabic have no initial vowels. Furthermore, Carter (2004) points out that classic Arabic does not allow a syllable to start with a vowel.

Third, there are varieties of Arabic that allow geminates to occupy the onset. Boudlal (2004) points out several examples in Casablanca Moroccan Arabic. Moreover, Boudlal highlights that complex onsets exist in the onset without any process of syncope.

So, clearly, the syllable in most of the varieties of Arabic starts with an onset and does not start with a vowel. Based on the studies presented, the onset in most varieties of Arabic is not complex save for the case of Casablanca Moroccan Arabic. Geminates, on the other hand, exist in several varieties both word medially and finally. In Casablanca Moroccan Arabic, however, geminates are allowed even initially.

Al-Ahsa dialect retains many features of the varieties of Arabic cited by the previous studies, namely, no initial vowels, no complex onsets and geminates. Unlike Casablanca Moroccan Arabic, however, neither complex onsets nor geminates in the onset are allowed.

3. Inventory of phonemes in Al-Ahsa dialect

In this section, the inventory of phonemes of Al-Ahsa dialect is discussed. Consonants are presented in Section 3.1 below; vowels are presented in Section 3.2 below.

3.1. Consonants

As can be seen in Table 1 (overleaf), Al-Ahsa dialect has 31 consonants, as opposed to classic Arabic, which has only 28 consonants. The presence of the additional consonants in Al-Ahsa dialect seems to be a result of the linguistic influence of the neighboring countries and of word borrowing, as these additional consonants can mainly be seen in foreign words (for classic Arabic consonants, see Table 1 in the Appendix). Al-Ahsa dialect consonants can be categorized as follows: nine stops; fourteen fricatives; two affricates; one trill; one lateral; two nasals; two semi-vowels. Newman (2002: 66–67) points out that the fully voiced stops never aspirate. The voiceless consonants, on the other hand, aspirate, except the uvular /q/. Newman also points out that the absence of the consonant /g/ in classic Arabic is quite justifiable, as it exists in only 21 languages. Interestingly, the phoneme /g/ is present in Al-Ahsa dialect, and in most cases it replaces the uvular /q/, as in /qariib/ ‘near’. So, in Al-Ahsa dialect, this word is pronounced as /gariib/. As it can be noticed, there is no voiceless /p/ in the classical Arabic nor in Al-Ahsa dialect. According to Newman, the voiced stop /b/ is devoiced when it is next to a voiceless sound. In other words, the voiceless [p] exists in Arabic, but as an allophone of /b/, and not as a separate distinctive phoneme. In addition, the phoneme /tʃ/ does exist in Al-Ahsa dialect. In many cases, it replaces the phoneme /k/, as in the word /tʃalb/ ‘dog’, which in fact must be pronounced as /kalb/ in classic Arabic.

3.2. Vowels

As can be seen in Table 2 on p. 151, the vowel system is simple. It consists of six vowels. Those vowels are composed of three sets of pairs. The first set of pairs is realized as

Table 1. The inventory of consonants in Al-Ahsa dialect.

Major class	cons	son	voice	spr gl	constr gl	labial	round	coronal	ant	str	dors	nasal	cont	lateral	trill
b	+	–	+			+							–		
f	+	–	–			+				–			+		
θ	+	–	–					+	+	–			+		
ð	+	–	+					+	+	–			+		
t	+	–	–					+	+				–		
d	+	–	+					+	+				–		
s	+	–	–					+		+			+		
z	+	–	+					+		+			+		
ṭ	+	–	–		+			+					–		
ḍ	+	–	+		+			+					–		
ṣ	+	–	–		+			+					+		
ð̣	+	–	+		+			+					–		
ʃ	+	–	–					+	–	+			+		
x	+	–	–								+		+		
ɣ	+	–	–								+		+		
k	+	–	–								+		–		
ʒ	+	–	–					+	–	+			–		
g	+	–	+								+		–		
q	+	–	–										–		
ħ	+	–	–										+		
ʕ	+	–	+										–		
h	+	–	–	+									+		
ʔ	+	–	–		+								–		
r	+	+	+					+	+				+		+
l	+	+	+					+					+	+	
n	+	+	+					+				+	–		
m	+	+	+			+						+	–		
w	+	+	+			+	+				+		+		
j	+	+	+						–				+		
ʧ	+	–	–						–	+			–		
ʤ	+	–	+						–	+			–		

Table 2. The inventory of vowels in Al-Ahsa dialect.

Features	i	ii	a	aa	u	uu
[high]	+	+	−	−	+	+
[low]	−	−	+	+	−	−
[back]	−	−	−	−	+	+
[round]	−	−	−	−	+	+
[tense]	−	+	−	+	−	+

[+high] [−back] [+tense] [ii] and [+high] [−back] [−tense] [i]. The second set of pairs is realized as [+high] [+back] [+tense] [+round] [uu] and [+high] [+back] [−tense] [+round] [u]. The third set is realized as [+low] [−back] [+tense] [aa] and [+low] [−back] [−tense] [a]. While the central mid vowel [ə] does not exist in Al-Ahsa dialect, according to Gadoua (2000), it does exist in Quranic Arabic due to certain phonological processes. (See Table 2 in the Appendix for the inventory of vowels in classic Arabic.)

One last point to mention about vowels of Al-Ahsa dialect, in particular, is the diphthongs [eɪ] and [ia]. Al-Ahsa dialect exhibits a diphthong similar to that of English [eɪ]. It is assumed here that it has entered this dialect through some loanwords, as it can be mostly found in them. The English words *radiator*, *light* and *gear* have been borrowed into this dialect and are pronounced as [radeɪtar], [leɪt] and [geɪr] respectively. Smeaton (1973) states that the dipthongs [au] and [ai] also exist in Al-Ahsa dialect. He relies on *laun* 'color' and *mai* 'water'. It is argued here that the words he relied on are purely Arabic and must be transcribed as *lawn* and *maj*, respectively. Therefore, these diphthongs do not exist in Al-Ahsa dialect.

3.3. The syllable structure of Al-Ahsa dialect

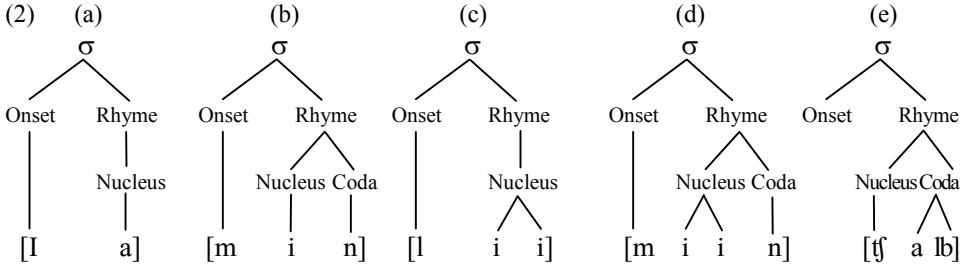
Al-Ahsa dialect exhibits six syllable patterns. These are organized in (1a–e):

(1) (a) CV
(b) CVC
(c) CVV
(d) CVVC
(e) CVCC

The syllables in (1) are organized from the super-light syllable to the super-heavy one. (a) is the lightest. (e), on the other hand, is the heaviest. From the syllable shapes presented in (1a–e), it is obvious that no syllable starts with a vowel. Onsetless syllables are prohibited in most, if not all, Arabic dialects. Also, similar to classical Arabic, while complex codas are allowed in Al-Ahsa Arabic, complex onsets are prohibited. Complex onsets exist in a few dialects of Arabic, the chief among them being Moroccan Arabic,

as it has been shown earlier in this paper. The representations of (1a–e) are demonstrated below. (The examples provided are the author's.)

(2)



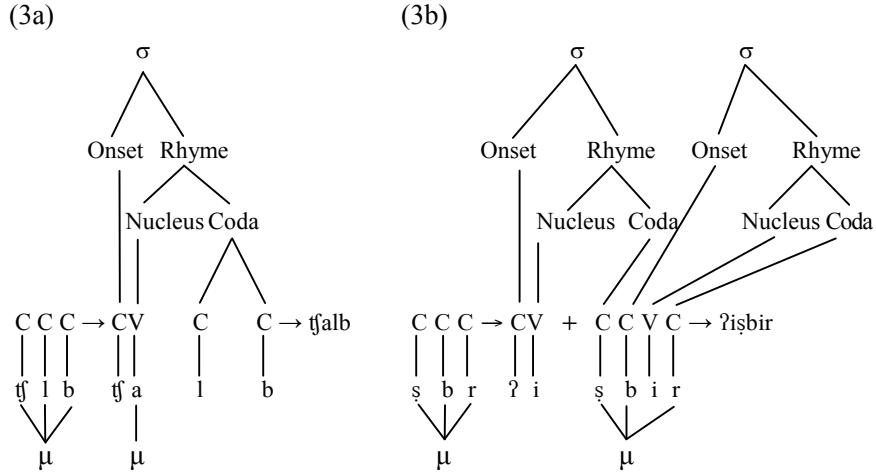
[Ia] 'no' [min] 'of, from' [lii] 'to me' [miin] 'Who's it?' [ʧalb] 'dog'

3.4. The morphology of Al-Ahsa Arabic

A few words in Al-Ahsa Arabic may be very simple and can be accounted for by the theories of concatenative morphology. Other words can be very complex and better analyzed using the Autosegmental Approach. I will start with the Autosegmental Approach, sketching it out very briefly.

Perhaps the most prominent and widely cited account of the Arabic morphology is the Autosegmental Phonology applied by McCarthy (1979) on Semitic languages as cited in Spencer (1998: 133–173). Undoubtedly, the same approach can be applied to account for the Arabic of Al-Ahsa. Briefly, like most of the dialects in Arabic, the words in Al-Ahsa Arabic are composed of a consonantal root (two, three or four consonants), and vowels are sprinkled in between those consonants. So, according to McCarthy's theory, the representation of a word or a morpheme is a string of consonants and vowels but no indication of the concise identity of those consonants and vowels is given. According to McCarthy, the specification of the CV templates – "prosodic templates" – is part of the Arabic grammar. The detail of McCarthy's theory and the derivational steps are beyond the scope of this paper. The interested reader can refer to McCarthy (1979) and Spencer (1998). However, there are few points that need to be stressed. What concerns us are the following: Every CV template must be associated with at least one melody element, and every melody element must be associated with at least one appropriate C or V slot; that is to say, C must be associated to a C slot and V must be associated to a V slot. Association of lines must not cross, and the process of association takes place from left to right, complying with the "Well-Formedness Condition". One last point is the assumption that every morpheme that makes up a word is assigned a separate tier known as the "morphemic tier" – this morphemic tier also represents the consonantal root of a word. To illustrate, let us consider *ʧalb* 'dog'. *ʧalb* has the triliteral root ʧ-l-b, and the vowel segment [a] is inserted between [tʃ] and [lb]. This is illustrated below in

(3a). (3b), on the right below, has the imperative *ʔiṣbir* 'wait.' It illustrates a somewhat complicated example. I do not intend to explain this example in detail, but it suffices to say that it consists of *s-b-r* as its consonantal root. Moreover, we have cases of an epenthetic glottal and a vowel. The purpose of the glottal is to start with an onset. The purpose of *-i* is to mark the grammatical property *imperative* and to break the consonantal cluster. (More will be mentioned about the glottal insertion later in the paper; see Table 3 in the Appendix for the list of abbreviations.)



Concatenative theories of morphology can also be geared to account for some regular patterns in Al-Ahsa Arabic. A word such as *ʧalb* can carry the inflectional accusative suffix *-u* and becomes *ʧalbu*. Also, it can carry the definite prefix *ʔal-* 'the' and becomes *ʔal-ʧalbu* acc.[1] 'the dog.' Some words have a regular plural pattern e.g. *ʔintʃ*/ *ʔintʃ-aat* fem.[2]pl.[3] 'inches' and *ʃarab*/*ʃarab-aat* 'socks.'

To sum up, we have seen cases where regular word patterns in Al-Ahsa Arabic can be explained by the concatenative morphological theories. In other cases, however, complex word patterns are better analyzed using the Autosegmental Approach.

4. Investigation

In Section 4.1 and 4.2, the data will be presented. Two data sets will be provided: A set of data for Al-Ahsa dialect and, for the sake of comparison, another dialect of Arabic, namely Casablanca Moroccan Arabic. Section 4.3 includes preliminary observations in regard to the data sets presented.

4.1. Al-Ahsa dialect data[1]

The following data contain a list of words exclusively pronounced by Hasawi speakers. Hasawi speakers reside in the eastern part of Saudi Arabia and are well-known for the following words. The data set is provided by the author of the present work who is a native speaker of the dialect discussed and, hence, is a native speaker of Arabic. (See Table 3 in the appendix for the list of abbreviations).

(4) [fukk] imp.[4] 'open'
(5) [ṣukk] imp. 'close'
(6) [baṭṭil] imp. 'open'
(7) [bannid] imp. 'close', usually of a tap or a valve
(8) [ʔintʃ] 'inch', [ʔintʃaat] 'inches'
(9) [xið] imp. 'take'
(10) [sahil] 'easy'
(11) [ṣaʕib] 'hard'
(12) [ħagaa] 'Is it real?'
(13) [ras] 'head'
(14) [daxal] v.[7] 'entered'
(15) [ṭalaʕ] 'went out'
(16) [ħabil] 'rope'
(17) [ṣabir] 'be patient'
(18) [ɣaziir] 'deep'
(19) [nuur] 'light'
(20) [xubiz] 'Arabian round bread'
(21) [niḍiif] adj.[5] 'clean'
(22) [daxaan] 'smoke'
(23) [gufil] n.[6] 'lock'
(24) [kaatib] 'clerk'
(25) [ʔiṣbir] 'wait'
(26) [ħajwan] 'animal'
(27) [qarʃa] 'bottle'
(28) [ʃarabaat] 'socks'
(29) [kabat] 'closet'
(30) [maʃɣuul] 'busy'
(31) [badʒi] or very informally [dʒaj] 'coming'
(32) [kabuut] 'hood/condom'
(33) [raħ] 'he went'
(34) [baɖja] 'bowl'
(35) [farħan] 'happy'
(36) [ʃaɣil] imp.'turn on'
(37) [bas] 'only'
(38) [dʒadaħa] 'lighter'
(39) [kal] 'ate'
(40) [zigara] 'cigarette'
(41) [tidʒuuri] 'treasury'
(42) [dʒalanṭ] 'handbrake'
(43) [giriib] 'close/near'
(44) [ʃarʃaf] 'light blanket',
(45) [diriiʃa] 'window'
(46) [naxalna] 'our farm'
(47) [ʒidir] 'pot'
(48) [dig] imp. 'call'
(49) [xaraaba] 'junkyard'
(50) [xirga] 'wash rag'
(51) [ʕeɪʃ] 'rice'
(52) [fiiʃ] 'plug'
(53) [diθwi] 'stupid'
(54) [ʃarabaat] 'socks'
(55) [diriiʃa] 'window'
(56) [dulaab] 'closet'
(57) [xirga] 'wash rag'
(58) [ʔiðbaħ] imp.'slaughter'
(59) [ʔinħaʃ] imp.'run away'
(60) [tʃalb] 'dog'
(61) [galb] 'heart'

[1] I thank the anonymous *PSiCL* reviewer for being a bit dubious in regard to examples (17 and 20) – which made me notice that there is a missed phoneme in the IPA transcription in both of them.

4.2. Casablanca Moroccan Arabic Data

The list provided is adopted from Boudlal (2004). The data set is very small but is provided for comparative purposes. (The interested reader can find an adequate and complete OT analysis of Casablanca Moroccan Arabic in Boudlal (2004).)

(62) [ʕməʃ] 'sleep'
(63) [ħməd] 'Ahmed (proper noun)'
(64) [ħnəʃ] 'snake'
(65) [smən] 'preserved butter'
(66) [nməl] 'ants'
(67) [ɡməl] 'lice'
(68) [nnʕəs] 'we sleep'
(69) [ħəbs] 'jail'
(70) [mədd] 'give'
(71) [sərʒ] 'saddle'
(72) [bəqləl] 'he gazed at'
(73) [ʃəmʃ] 'sun'

4.3. Preliminary observations

Observing the data presented above in 4.1 reveals the following facts about the syllable shape of Al-Ahsa dialect. The first observation is that all words begin with onsets which, in its turn, entails that no word begins with a vowel in this dialect. The second observation is that complex onsets do not exist, whereas complex codas are allowed. The last observation is that geminates are allowed in Al-Ahsa dialect in a syllable-medial position and in a syllable-final position only.

Like the dialect of Al-Ahsa, all the words of Casablanca Moroccan Arabic presented in 4.2 above start with onsets and, hence, no word starts with a vowel. Also, similar to Al-Ahsa, geminates and final consonantal clusters are allowed. However, unlike the dialect of Al-Ahsa, onsets can be complex consisting of more than one consonant – which is not seen, and, therefore, is not allowed in the dialect of Al-Ahsa. Another difference is that geminates can appear in the onset position. Let us see how those observations are captured by OT in the following sections.

5. Analysis within the Optimality Theory framework

In Section 5.1, a very brief summary of OT framework is presented. Section 5.2 contains the analysis of data presented within the OT framework. Finally, section 5.3 discusses the geminate consonants.

5.1. The Optimality Theory framework in a nutshell

In this section, I will first provide a very brief summary of OT (Prince and Smolensky 1993). An outline of the "classic" OT theory looks like Figure 1 below. Figure 2 contains the explanation for the abbreviations used in Figure 1.

Input → GEN → candidate set → EVAL (constraints) → Optimal Output

Figure 1. The structure of OT.

GEN: From a given input, the GEN generates a potential set of candidate outputs. **EVAL:** From the set of candidates, the EVAL chooses the optimal output for the input. There is also another module called **CON**. **CON** works in hand with **EVAL** and manipulates a language specific ranking of constraints from what is called the UNIVERSAL SET OF CONSTRAINTS.

Figure 2. Explanation.

There are a few points that I would like to stress in regard to OT. Both inputs and outputs are levels of representation. Unlike the theory of Generative Phonology, OT is nonderivational and mapping from inputs to outputs takes place directly without intermediate steps. GEN generates candidate outputs. EVAL compares the candidates selecting the most optimal according to the language-specific constraint hierarchy. In the latest version of Prince and Smolensky (2004), the system of candidate generation (GEN) of OT is completely separate from the system of candidate evaluation (EVAL). With this framework in mind, let us see how OT captures the facts about the syllable of Al-Ahsa Arabic and, whenever applicable, a very few similar or contrastive points are highlighted regarding the syllable of Casablanca Moroccan Arabic.

5.2. Analysis within the OT framework

According to Prince and Smolensky (1993, 2004), a syllable structure in OT is generated in a way similar to any other grammatical feature. GEN generates a set of candidates for unsyllabified inputs. EVAL chooses the optimal candidate which should comply with the constraints imposed by the universal grammar (UG) and ranked on a language-specific basis. Taking the previously stated observations into consideration and considering the syllable patterns presented in 3.3 (2a–e) above, we are led to formulate the following constraints in regard to the syllable shape of Al-Ahsa dialect. These are presented in (74) below.

(74a) ONSET: A syllable must have an onset. A syllable must begin with an onset as all words in Al-Ahsa dialect begin with a consonant and none starts with a vowel. Similarly, all words in Moroccan Arabic start with a consonant.

(74b) NO-CODA: A syllable must not have a coda. Clearly this constraint is very frequently violated in both Al-Ahsa and Moroccan Arabic.

According to Prince and Smolensky, ONSET and NO-CODA are universally unmarked constraints. From the data presented in Section 4.1 together with the observations presented earlier, it is clear that all words start with a consonant and no word starts with a vowel. As a result, we will assume that ONSET must be ranked very high. To illustrate, let us consider the input (31) /badʒi/ 'coming' from the data presented in Section 4.1 above. Together ONSET and NO-CODA constraints enable GEN to generate the following candidates presented in Tableau 1 below. (The dot "." marks the syllable boundary.)

Tableau 1.

/badʒi/	ONSET	NO-CODA
☞ a) .ba.dʒi.		
b) .badʒ.i.	*!	*

Tableau 1 shows two possible candidates generated by GEN, of which only (a) is optimal. (b) is eliminated by the fatal mark "*!" caused by the second syllable starting with a vowel. /badʒi/ is of a CVCV shape and seems be straightforward. Let us consider another example, preferably, of a CVC shape. Considering input (33) /raħ/ 'went' together with ONSET and NO-CODA constraints, GEN generates the following possible candidates which are presented in Tableau 4 below.

Tableau 2.

/raħ/	ONSET	NO-CODA
a) .raħ.		*
☞ b) .ra.ħ		

Tableau 2 shows two possible candidates generated by GEN, of which only (a) is optimal. However, the choice (b) is forced by the ranking ONSET>>NO-CODA. Moreover, we should expect GEN to generate a candidate such as the one in (b) and selects it as optimal based on the hierarchy of constraints and the violations they incur. Reasonably good questions to ask are: What forces the violation of NO-CODA? And is there a constraint ranked higher than NO-CODA and lower than ONSET? And, most importantly, how is (b) eliminated and how is (a) chosen as optimal? According to Prince and Smolensky (1993), (b) can be pronounced fatal by the constraint PARSE-segment (hereinafter PARSE-seg). PARSE-seg is stated in (75).

(75) PARSE-seg: Every segment must belong to a syllable.

With the constraint (75), (b) in Tableau 2 above is pronounced dead. Tableau 2 can be reformulated as in Tableau 3 below. From Tableau 3, it is clear that PARSE-seg outranks NO-CODA and, therefore, NO-CODA is better violated.

Tableau 3.

/raħ/	ONSET	PARSE-seg	NO-CODA
☞ a) .raħ.			*
b) .ra.ħ		*!	

According to McCarthy and Prince (1995, 1999), there is another set of the "faithfulness family of constraints" related to the syllable structure. These are stated below in (76). (The constraints (76a–b) correspond to PARSE and FILL in Prince and Smolensky (1993)).

(76a) MAX-IO: Every segment of the input has a correspondent in the output.
(76b) DEP-IO: Every segment of the output has a correspondent in the input.

Observing the data in Section 4.1, those who are familiar with Arabic may notice that items (25, 58, 59) contain an epenthetic glottal stop to satisfy ONSET and, thus, it is evident that DEP-IO is violable and must be outranked by MAX-IO and ONSET. Therefore, it follows that ONSET outranks DEP-IO since (25, 58, 59) will parse and pass ONSET and incur a violation of DEP-IO. Furthermore, according to McCarthy (2005), the constraint ONSET is not dominated in Arabic. Therefore, we can safely assume that ONSET outranks both MAX-IO and DEP-IO in Al-Ahsa dialect. In addition, it is a fact about Arabic agreed upon by McCarthy (1979), Boudlal (2004) and Gadoua (2000), among others, that no vowel starts a syllable in Arabic, and when this situation arises and an epenthetic glottal stop is inserted to satisfy ONSET and, hence, DEP-IO is unavoidably violated. Cases of epenthesis in Arabic are many and beyond the scope of this paper, but it suffices to hypothesize that MAX-IO outranks DEP-IO but ONSET outranks both.

Regardless of what has been mentioned above, let us see whether or not DEP-IO is actually outranked by MAX-IO. Let us take (25) *ʔiṣbir* 'wait'. Having said the glottal is a case of epenthesis, the given input must be /iṣbir/. The relevant possible candidates are shown in Tableau 4 below.

Tableau 4.

/iṣbir/	ONSET	DEP-IO	MAX-IO	NO-CODA
a) .iṣ.bir.	*!			*
b) .ʔiṣ.bir.		*!	*	*
☞ c) .iṣ.bi.			*	*

In Tableau 4 above, (a) is fatal for not having an onset. (b) is punished for having a segment in the output that does not have a correspondent in the input, but this is forced by the ranking given. (c) is chosen as the wrong optimal candidate based on the hierarchy of the constraint violation shown above. Therefore, MAX-IO must outrank DEP-IO; ONSET must outrank both.

Now, let us consider *ʔiṣbir* 'wait' again and see how the constraints in (76a–b) interact with ONSET, PARSE-seg and NO-CODA. The input /iṣbir/ and the relevant possible candidates are illustrated in Tableau 5 below.

Tableau 5.

/iṣbir/	ONSET	MAX-IO	PARSE-seg	DEP-IO	NO-CODA
a) .iṣ.bir.	*!				*
☞ b) .ʔiṣ.bir.				*	*
c) .ʔiṣ.bi.		*!		*	*
d) .ʔiṣ.bi.r			*!	*	*

In Tableau 5, (a) is penalized for not having an onset – a very high-ranked constraint that cannot be dominated. The candidate in (c) is grounded by MAX-IO for not having a correspondent for the segment [r] in the output. Finally, (d) is eliminated because of having a segment that does not belong to a syllable. We are left with (b) with a minimal acceptable violation.

Let us summarize what we have developed thus far. ONSET is undominated and outranks MAX-IO; MAX-IO outranks both PARSE-seg and DEP-IO. NO-CODA is the lowest ranking constraint and is very violable. The hierarchy of constraints can be summarized in (77) as follows:

(77) ONSET>>MAX-IO>>PARSE-seg>>DEP-IO>>NO-CODA

Notice that in all the examples given earlier the onset in Al-Ahsa dialect consists of a single consonant. No word starts with a complex onset. Complex onsets in Moroccan Arabic, on the other hand, are possible as some forms listed earlier in Section 4.2 contain complex onsets e.g. /ʕməʃ/ 'sleep', /ħnəʃ/ 'snake' and /nməl/ 'ants', among others. However, from the dialect of Al-Ahsa, consider (34) /badja/ 'bowl' from Section 4.1. The relevant candidates are shown in Tableau 6 below.

Tableau 6.

/badja/	ONSET	MAX-IO	PARSE-seg	DEP-IO	NO-CODA
a) .badj.a.	*!				*
☞ b) .ba.dja.					
c) .bad.ja.					*

(a) is pronounced dead for violating ONSET. (c) violates a very low-ranked constraint, namely, NO-CODA. (b) is chosen as the optimal candidate because it incurs no violation at all. However, a close observation of the data in 4.1 reveals that a complex onset does not exist in Al-Ahsa Arabic and, thus, a complex onset is not allowed. It is preferred to violate a very low-ranked constraint than to have a complex onset. So, in order to ensure that EVAL selects (c) not (b) as optimal, another constraint must be formulated and must be of a high rank. The constraint is formulated in (78) below.

(78) COMPLEX (hereafter COM): A syllable edge must consist of a single consonant only.

It is obvious from Tableau 6 that a complex onset will pass MAX-IO. Thus, COM must be ranked higher than MAX-IO and lower than ONSET. Tableau 6 is reformulated in Tableau 7 with the constraint COM.

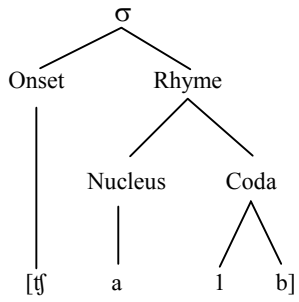
Tableau 7.

/badja/	ONSET	COM	MAX-IO	PARSE-seg	DEP-IO	NO-CODA
a) .badj.a.	*!					*
b) .ba.dja.		*!				
☞ c) .bad.ja.						*

Candidate (a) is eliminated for violating ONSET. (b) is also eliminated for having a complex onset. Because of the constraint stated in (6), the dilemma is resolved and we are left with (c) as the optimal candidate.

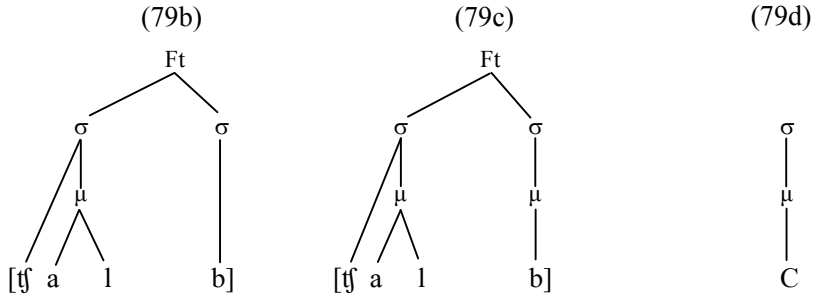
So far we have considered examples with no final consonant clusters. Words such as /tʃalb/ ‘dog’ ends with two consonants in the coda. Our previous syllabification of /tʃalb/ is repeated in (79a) below for convenience.

(79a)



For a reason that will be shown later as the paper progresses, we need to modify the syllabification presented in (79a). According to Hayes (1989), Al Ghadi (1994) and Boud-

lal (2004: 62–68), under moraic theory, two other possible syllabifications are available in the literature, each of which has consequences to consider. Those syllabifications are presented in (79b–c).



The representation in (79b–c) makes a distinction between two types of syllables: A major syllable whose nucleus is a vowel and a light syllable known as a "degenerate syllable". The degenerate syllable consists of a single consonant only. A representation of the degenerate syllable is shown in (79d). Notice that the syllable node is adjoined directly to the phoneme [b] in (79b), whereas it is adjoined to a mora, then, to the phoneme [b] in (79c).

What concerns us here are the consequences of adopting either one in our analysis. It should be pointed out that adopting (b) does not constitute a violation of the constraint (75) PARSE-seg stated earlier. Rather, it is Selkirk's (1980, 1981) Strict Layer Hypothesis that is violated – which states that every prosodic constituent must be dominated by a constituent of the immediately superior type. That is to say, the mora is dominated by the syllable, and the syllable is dominated by the foot which is, in its turn, is dominated by the prosodic word. Moreover, according to Prince and Smolensky (1993), (79b) poses another problem in regard to the word minimality: They highlight that "the foot is subject to binarity which requires the prosodic word (PWd) to have at least two moras if the language under study is quantity-sensitive or two syllables if the language is quantity-insensitive". Clearly, then, the syllabification in (7b) poses several issues and incurs two violations: Strict Layer Hypothesis and foot binarity.

In this paper, we adopt the syllabification in (79c). The advantages are many. First, neither Strict Layer Hypothesis nor foot binarity is violated.[2] In addition, recall what is

[2] Recall that Arabic words are formed from a consonantal root called a morphemic tier. Also, according to Prince and Smolensky (1993), a member of the morphological category corresponds to a prosodic word (Lexical (LX) ≈ Prosodic word (PWd)), which, as a result, corresponds to a foot. Thus, by adopting (79c) above, we not only solve the issue of foot-binarity violation but also solve problems regarding LX ≈ PWd and satisfy their constraints.

presented in Section 3.4 in regard to the consonantal root of Arabic. McCarthy (1979) points out that the consonants of the root constitute a morphemic tier. In other words, they are adjoined to moras. Thus, the syllabification (79c) is unarguably better than the one in (79b) and conforms to Arabic in general and to Al-Ahsa dialect in particular. However, it should be stated that the representation in (79c) is not entirely without flaws. The morification of the consonants does incur a violation of Nuclear Harmony (Prince and Smolensky 1993). However, as Boudlal (2004) highlights satisfying foot-binarity is very common and is found cross-linguistically.

Having developed the previous argument, three relevant constraints must be formulated accordingly. One concerns the minor syllable – of the representation (79d) above. The second concerns the Foot-Binarity. The third concerns Nuclear Harmony. Those are stated in (80–82), respectively.

(80) Degenerate Syllable (hereafter DEG-σ): A degenerate syllable is prohibited.
(81) Foot-Binarity (hereafter Ft-BIN): Foot is binary under syllabic or moraic analysis.
(82) Nuclear Harmony (henceforth NUC-H): A higher sonority nucleus is more harmonic than one with lower sonority.

With those constraints in mind, let us consider [tʃalb] 'dog'. The candidates are in Tableau 8 below.

Tableau 8.

/tʃalb/	Ft-BIN	COM	DEP-IO	DEG-σ	NUC-H	NO-CODA
a) .tʃal.b	*!			*		*
☞ b) .tʃal$^{\mu}$.b$^{\mu}$				*	*	*
c) .tʃalb.		*!				*

In Tableau 8, candidate (a) incurs a fatal violation of Ft-BIN for having a minor syllable with (an) unmorified consonant(s). It should also be noted that if the constraint DEG-σ is not hypothesized, then we assume that [b] is an unparsed segment violating PARSE-seg – which entails that Ft-BIN dominates PARSE-seg.[3] Candidate (c) is eliminated for having a complex coda, whereas (b) is chosen as optimal with one minimal acceptable violation in DEG-σ, NUC-H and NO-CODA.[4]

[3] In other words, if I use a syllabification such as [tʃal.b] without marking DEG-σ with an asterisk, then it is assumed to be an unparsed segment not belonging to any syllable or having a minor syllable of its own. In such a case, it is to be handled and taken care of by PARSE-seg.

[4] It should be noted that the only way to pass COM is to have a degenerate syllable, and the only way to satisfy Ft-BIN is to morify the consonants which, in its turn, incurs a violation of NUC-H.

It should be clear from what has been said thus far and from the constraints stated in (80–82) that the only way to satisfy Ft-BIN is to morify the consonant which, in its turn, incurs a violation of NUC-H. So, clearly, Ft-BIN must dominate both DEG-σ and NUC-H. Note also that DEP-IO must outrank DEG-σ and NUC-H. Henceforth, in order not to be repetitive, any form of the representation (79c) will be written with moras as in [tʃal$^{\mu}$.b$^{\mu}$]; otherwise, it is written as [tʃal.b]. Let us take another example and see how all the constraints developed interact. The example we consider is (61). The input is /galb/ 'heart.' The candidates are presented in Tableau 9.

Tableau 9.

/galb/	Ft-BIN	ON-SET	COM	MAX-IO	PARSE-seg	DEP-IO	DEG-σ	NUC-H	NO-CODA
a) .galb.			*!						*
b) .gal.b	*!						*		*
☞ c) .gal$^{\mu}$.b$^{\mu}$							*	*	*
d) .lgal$^{\mu}$.b$^{\mu}$			*!			*	*	*	*
e) .al.gal$^{\mu}$.b$^{\mu}$		*!				**	*	*	*
f) .ʔal.gal$^{\mu}$.b$^{\mu}$						***!	*	*	*
g) l.gal$^{\mu}$.b$^{\mu}$					*!	*	*	*	*
??h) l$^{\mu}$.gal$^{\mu}$.b$^{\mu}$						*	**	*	*

From Tableau 9 above, candidate (a) is eliminated for having a complex coda. (b) is grounded for violating Ft-BIN by having an unmorified minor syllable. Candidates (d–h) contain segments of definite article "the" and, thus, are very possible candidates, each of which violates several constraints. (d) violates a very high-ranked constraint, namely COM; therefore, it is punished. Likewise, (e) violates a very high-ranked constraint, namely ONSET; therefore, it is grounded for starting with a vowel. While candidate (f) is very possible, it is punished based on the hierarchy of the constraints and on the number of violations in incurs – it violates DEP-IO three times in addition to violating the last three constraints. Thus, in terms of quantity, candidate (c) is much better than the one in (f).[5] (g) is punished for having an unparsed segment that does not belong to a syllable nor does it have a minor syllable of its own. It should be noted that the first [l] does not violate Ft-BIN; it is assumed here to be just a floating segment. However,

[5] There had been debates in regard to the underlying form of the definite article *ʔal-* in Arabic literature. Logically, there are three possible forms: *ʔal-*, *al-*, and *l-*. The debate centered on the first form and the third one as *al-* is rejected based on the fact that Arabic words do not start with a vowel. *ʔal-* denotes that the glottal is epenthetic, whereas *l-* means that both the glottal and the vowel are epenthetic. Further information in regard to this matter can be found in Gadoua (2000) among others. Future research should put those points in mind and consider the possible hierarchy and interaction of the constraints regarding the definite article in Arabic.

candidate (h) passes the high-ranked undominated constraints by having all the three degenerate syllables with moraic consonants. Unlike the Casablanca Moroccan Arabic, a minor syllable even with a moraic consonant is not allowed in Al-Ahsa dialect.[6] Accordingly, a new constraint must be formulated. This is stated in (83).

(83) NO ONS-DEG-σ: No degenerate syllables are allowed in the onset.

Tableau 9 is repeated in Tableau 10; irrelevant constraints are deleted. The same analysis presented earlier applies here. The only difference is that candidate (h) is eliminated for having a minor syllable starting the word. We assume that ONS-DEG-σ ranks lower than COM and outranks MAX-IO since segments with moraic consonants will pass MAX-IO as it is shown in Tableau 9.

Tableau 10.

/galb/	Ft-BIN	ON-SET	COM	ONS-DEG-σ	MAX-IO	PARSE--seg	DEP-IO	DEG-σ
a) .galb.			*!					
b) .gal.b	*!							*
☞ c) .gal$^{\mu}$.b$^{\mu}$								*
d) .lgal$^{\mu}$.b$^{\mu}$			*!				*	*
e) .al.gal$^{\mu}$.b$^{\mu}$		*!					**	*
f) .ʔal.gal$^{\mu}$.b$^{\mu}$							***!	*
g) l.gal$^{\mu}$.b$^{\mu}$						*!	*	*
h) l$^{\mu}$.gal$^{\mu}$.b$^{\mu}$				*!			*	**

To conclude this section, let us recapitulate what we have developed thus far. First, it should be noted that the constraints DEP-IO, DEG-σ, NUC-H and NO-CODA are highly violable in Al-Ahsa dialect, as Tableau 9 and 10 show. The constraints Ft-BIN, ONSET, COM, ONS-DEG-σ, MAX-IO and PARSE-seg, on the other hand, are very high-ranked constraints. Hence, the constraints at our disposal and their ranking are listed in (84) as follows:

(84) Ft-BIN >> ONSET >> COM >> ONS-DEG-σ >> MAX-IO >> PARSE-seg >> DEP-IO >> DEG-σ >> NUC-H >> NO-CODA

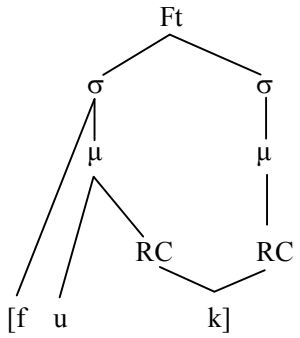
[6] Boudlal (2004) assumes that a degenerate syllable in the onset is allowed in Casablanca Moroccan Arabic. Therefore, the constraint ONS-DEG-σ does not apply for Casablanca Moroccan Arabic.

5.3. Analysis of the geminates in Al-Ahsa dialect

The theory we adopt in regard to the analysis of geminates is the Two-Root Theory of Selkirk (1990, 1991). This theory views geminates as a representation of two root nodes that share stricture and place features.

According to Selkirk, the advantage is that it allows for a clear distinction between full and partial geminates. Full geminates involve full feature sharing, whereas partial geminates are structures where specifications for laryngeal features or nasality may differ in halves. What concerns us here are the full geminates. Selkirk's representation of full geminates is demonstrated in (85) below. The example we consider is /fukk/ 'open' from section 4.1.

(85)



However, it should be pointed out that the Two-Root Theory does not specify the moraic structure of geminates. This is a property that should be supplied from the language. In Arabic and in Al-Ahsa dialect in particular – as it has already been noted earlier in regard to the root consonants – the final consonants of the word in (85) is associated to a morphemic tier that is represented here as two root nodes, hence producing a final geminate, both of which is associated to a mora to satisfy Ft-BIN. In other words, the Two-Root Theory treats the final geminate as a final consonantal cluster. Note that the representation in (85) somewhat resembles the syllable representation presented earlier in (79c) save for the fact that the RC nodes are adjoined together in (85).

Thus, similar to the final consonantal cluster, it follows that the structure of geminates is represented as [.fuk$^{\mu}$.k$^{\mu}$] or else it will be punished by the constraint COM or the relevant constraints stated earlier. All the examples of final geminates presented in 4.1 have no problem since they are all treated as final consonant clusters. So, the same constraints that handle consonant cluster can take care of the geminates. However, examples of medial geminates such as [bannid] 'close – usually of a tap or a valve' may be subject to a syllabification such as [.ban.nid.] in which the first [n] is a coda of the first syllable while the second [n] is an onset of the second syllable, thus, passing the

constraint COM by splitting the geminates. This is not what we desire to have. Therefore, a new constraint is formulated in (86) as follows.

(86) No Geminates Splitting (Hereafter $^{\mu}]^{\sigma}$ Gem. $^{\sigma}[^{\mu}$): Geminates must not split.

I assume that the function of $^{\mu}]^{\sigma}$ Gem. $^{\sigma}[^{\mu}$ is to ensure not only that the geminates do not split but also that no vowel is inserted between them. Now, let us see how this constraint interacts with the previously presented ones. The underlying input is /bannid/. The candidates are shown in Tableau 11.

In Tableau 11, the dotted lines mean we assume no ranking between PARSE-seg and $^{\mu}]^{\sigma}$ Gem. $^{\sigma}[^{\mu}$. Candidate (a) is punished for splitting the geminates. (b) is eliminated for the fact that the second syllable starts with a vowel. The first syllable in (b) is also punished for having a complex coda. (c) is left as an optimal candidate.[7]

To summarize, in this section we have looked at geminates both finally and word medially. Regarding the structure of geminates, we adopt Selkirk's Two-Root Theory. It is concluded that final geminates are treated as final consonant clusters. However, in order to account for geminates occurring in medial positions, a new constraint has been formulated to prevent geminate splitting. The hierarchy of constraints at our disposal is listed below in (87).

(87) Ft-BIN >> ONSET >> COM >> ONS-DEG-σ >> MAX-IO >> $^{\mu}]^{\sigma}$Gem.$^{\sigma}[^{\mu}$, PARSE-seg >> DEP-IO >> DEG-σ >> NUC-H >> NO-CODA

Tableau 11.

/bannid/	Ft-BIN	ON-SET	COM	ONS-DEG-σ	MAX-IO	$^{\mu}]^{\sigma}$ Gem. $^{\sigma}[^{\mu}$	PARSE--seg	DEP-IO	DEG-σ	NUC-H	NO-CODA
a) .ban.nid.						*!					*
b) .bann.id.		*!	*!								*
☞c) .ban$^{\mu}$.n$^{\mu}$id.									*	*	*

6. Conclusion and future directions

This work discusses the syllable structure of Al-Ahsa dialect. Al-Ahsa dialect is a variation of Arabic which prohibits initial consonant cluster but allows complex codas and

[7] It is highlighted in the literature that an input without geminates such /banid/ will pass the constraints and is chosen as the optimal utterance – a situation which made Boudlal (2004), among others, posit a constraint in regard to the geminates, namely, MAX-RC which states that consonants of the root must be preserved. We assume here that MAX-IO takes care of such an input and eliminates it; a constraint as MAX-RC is necessary if non-derived words are considered.

geminates. The OT framework (Prince and Smolensky 1993, 2004) is adopted and provides a complete analysis of the syllable structure in Al-Ahsa dialect. A detailed set of constraints is hypothesized to account for what is permissible and what is prohibited. It is shown that Al-Ahsa dialect, unlike Morocan Arabic, exhibits no initial consonant clusters but allows complex codas and geminates. The complex codas are analyzed as dominated by moras which are dominated by syllables. In regard to final geminates, the Two-Root Theory is adopted and, thus, final geminates are analyzed as final consonant clusters. Medial geminates, however, demand another constraint, namely, no geminate splitting. This paper, however, has not discussed the non-derived forms of words and how they are generated nor has it discussed the underlying form of the definite article *ʔal-* 'the' in Arabic and its complete assimilation to the following coronal consonants. Future research should have those points in mind.

REFERENCES

Al Ghadi, A. 1994. An OT account of Moroccan Arabic prosody. [Ms., Delaware University, and Université Mohamed V, Rabat.]

Archibald, J. 2003. "Learning to parse second language consonant clusters". *Canadian Journal of Linguistics/Revue Canadienne de Linguistique* 48(3–4). 149–177.

Boudlal, A. 2004. Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic. [Unpublished Ph.D. dissertation, Université Mohamed V, Rabat.] (Rutgers Optimality Archive 650.)

Carter, M.G. 2004. *Sibawayhi*. London: Tauris.

Edzard, L. 2000. "Sibawayhi's observations on assimilatory processes and re-syllabification in the light of optimality theory". *Journal of Arabic and Islamic Studies* 3. 48–65.

Gadoua, A.H. 2000. "Consonant clusters in Quranic Arabic". *Cahiers Linguistiques d'Ottawa* 28. 59–85.

Gafos, A.I. 2003. "Greenberg's asymmetry in Arabic: A consequence of stems in paradigms". *Language* 79(2). 317–355.

Haddad, Y.A. 2005. "Etymological itineraries in second language phonology: The case of Arabic". Paper presented at the Second Language Research Forum, Columbia University. (Rutgers Optimality Archive 784–1105.)

Hayes, B. 1989. "Compensatory lengthening in moraic phonology". *Linguistic Inquiry* 20(2). 253–306.

McCarthy, J. 1979. Formal problems in Semitic phonology and morphology. [Unpublished Ph.D. dissertation, MIT.] (Distributed by IULC; published by Garland, New York, 1982.)

McCarthy, J. 2005. "Optimal paradigms". In Downing, L., T. Alan Hall and R. Raffelsiefen (eds.), *Paradigms in phonological theory*. Oxford: Oxford University Press. 170–210.

McCarthy, J. and A. Prince. 1995. "Faithfulness and reduplicative identity". In Beckman, J., L. Walsh Dickey and S. Urbanczyk (eds.), *University of Massachusetts Occasional Papers in Linguistics* 18*: Papers in Optimality Theory*. Amherst, MA: GLSA Publications. 249–384. (Rutgers Optimality Archive 60.)

McCarthy, J. and A. Prince. 1999. "Faithfulness and identity in prosodic morphology". In Kager R., H. van der Hulst and W. Zonneveld (eds.), *The prosody–morphology interface.* Cambridge, MA: MIT Press.

Newman, D.L. 2002. "The phonetic status of Arabic within the world's languages: The uniqueness of the lughat al-daad". *Antwerp Papers in Linguistics* 100. 65–75.
Prince, A. and P. Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. (Technical Report no. RuCCS-TR-2.) New Brunswick, NJ: Rutgers University Center for Cognitive Science.
Prince, A. and P. Smolensky. 2004. *Optimality theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.
Selkirk, E.O. 1980. "The role of prosodic categories in English word stress". *Linguistic Inquiry* 11. 563–605.
Selkirk, E.O. 1981. "Epenthesis and degenerate syllable in Cairene Arabic". In Borer, H. and J. Aoun (eds.), *Theoretical issues in the grammar of the Semitic languages*. Cambridge, MA: MIT Press. 111–140.
Selkirk, E.O. 1990. "A two-root theory of length". In Padgett, J. and E. Dunlap (eds.), *Papers in phonology*. (*University of Massachusetts Occasional Papers* 14.) Amherst, MA: GLSA Publications. 123–177.
Selkirk, E.O. 1991. "On the inalterability of geminates". In Bertinetto, P.-M., M. Kenstowicz and M. Loporcaro (eds.), *Certamen Phonologicum II. Papers from the 1990 Cortona Phonology Meeting*. Torino: Sellier. 187–209.
Smeaton, B.H. 1973. *Lexical expansion due to technical change*. Bloomington: Indiana University Press.
Spencer, A. 1998. *Morphological theory*. Malden, MA: Blackwell Publishing.

APPENDIX

Table 1. The inventory of consonants in Classic Arabic.

	Bilabial	Labiodentals		Interdental		Dental		Pharyngealized Dentals		Alveolar	Palatal	Velar		Uvular	Pharyngeal		Glottal
Voicing		−	+	−	+	−	+	−	+			−	+		−	+	
Stop	b ب					t ت	d د	ṭ ط	ḍ ض			k ك		q ق			ʔ ء
Fricative		f ف		θ ث	ð ذ	s س	z ز	ṣ ص	ð̣ ظ		ʃ ش	x خ	ɣ غ		ħ ح	ʕ ع	h ه
Affricate											dʒ ج						
Trill										r ر							
Lateral										l ل							
Nasal	m م									n ن							
Semi-Vowel	w و										j ي						

Table 2. The inventory of vowels in Classic Arabic.

	front		Central	Back	
High	ii	i		uu	u
Mid			ə		
Low	aa	a			

Table 3. Abbreviations.

#	Abbreviations	Explanation
1	acc.	accusative case
2	fem.	feminine
3	pl.	plural
4	imp.	imperative form of the verb
5	adj.	adjective
6	n.	noun
7	v.	verb