When Vowels are Empty and Consonants Collide:

Exploring the Consonant Clusters of Mukallawi Arabic in Strict CV

Husam Al-Gariri

11612304

BA Thesis Linguistics

UNIVERSITY
OF AMSTERDAM

Supervised by:

dr. B.M. van ’t Veer

June 22th 2020

Abstract

This study explores the consonant clusters of Mukallawi Arabic (MA), within the theoretical framework of Strict CV. This variety of Arabic differs from most other varieties in that it allows word-initial clusters, but prohibits word-final clusters, whereas most varieties of Arabic would exhibit the exact opposite. In Strict CV, consonant clusters are the result of an empty interconsonantal nucleus that does not have a phonetic interpretation. Furthermore, each nucleus must have a governor which allows it to be empty, otherwise it would obligatorily have a phonetic representation. From the perspective of Strict CV, the permissibility and prohibition of clusters in MA can be attributed to the government and licensing abilities of its nuclei and their relationship with their onsets as well as the onsets and nuclei of neighboring CV units. With the main difference between MA and most other varieties of Arabic lies in the capabilities of the word-final empty Nucleus (FEN). On one hand the FEN of MA cannot allow the preceding nucleus to be empty, while the FEN in most other varieties does have the ability to do so. On the other hand, the first nucleus of an MA word is allowed to be empty when followed by a full nucleus, while most other varieties never allow the first nucleus to be empty. Strict CV provides new insights to the specific properties that set MA apart from most varieties of Arabic in its tolerance of consonant clusters in certain positions and lack thereof in others.

Table of Contents

1. Introduction 1
1.1 The language 1
1.1.1 The syllable structure of MA 2
1.1.2 An Optimality Theoretic account of the clusters of MA 3
1.2 Theoretical background 4
1.2.1 Strict CV 4
1.2.2 P-licensing and consonant clusters 6
1.3 The aim 7
2. Methods 7
2.1 Data and observations 9
3. Analysis 14
3.1 Word-initial sequences 14
3.2 Word-final sequences 16
3.3 Triconsonantal sequences 17
3.4 The FEN of MA 18
4. Conclusion and discussion 19
5. Appendix 21
Appendix 1: Consonantal inventory of MA 21
6. Bibliography 21

1. Introduction

Modern Standard Arabic (MSA) is the variety used in literature, media and official settings in all Arabic speaking countries in the Middle East and North Africa. However, each of these countries has its own collection of spoken varieties that are similar to MSA in many aspects, but differ from it in many as well. One way most varieties are similar to MSA is in that they do not allow consonant clusters word-initially, but do allow them word-finally (Haddad, 1984; Haddad, 2006; Hayes, 1995; Watson 2002). On the other hand, there are varieties that do allow consonant clusters word-initially, such as, Moroccan Arabic (Dell & Elmedlaoui, 2002) and Mukallawi Arabic (Al Tairi, 2010). The latter is the focus of this paper. Al Tairi (2010) establishes that Mukallawi Arabic (henceforth, MA) does have initial consonant sequences that are not conditioned by sonority or assimilation. This study will further investigate the consonantal sequences of this variety, asking and answering questions such as; whether these sequences constitute complex onsets or rather the result of the adjacency of a syllabic consonant and a simplex onset, which was proposed to be the case in the initial consonant sequences of Casablanca Moroccan Arabic (Dell & Elmedlaoui, 2002). Furthermore, it will attempt to show that both types of consonant clusters as well as other related phenomena can be accounted for using more general concepts, without the need for external theoretical tools. Generally, this study will view the consonant sequences of MA from a new perspective, namely, Strict CV (Lowenstamm, 1996; Scheer 2004) with the goal of gaining a better understanding of what allows it to have consonantal sequences word-initially and not word-finally.

1.1 The language

Mukallawi Arabic (MA) is used to refer to the variety of Arabic spoken by the people of Mukalla city located in the southern coastal area of Yemen. Although we do not have an official number of speakers, based on the population of the city, one can estimate this variety to have about 300,000 speakers. Mukalla city belongs administratively to the province of Hadhramout, however, the term Hadhrami (meaning 'of Hadhramout') Arabic is usually used to refer to the variety or varieties spoken in the valley of Hadhramout (Al-Saqqaf, 1990). Most scholars make the distinction between the varieties spoken in the valley of Hadhramout as opposed to the varieties of the coast of Hadhramout. The MA variety belongs to the dialect continuum of the coast of Hadhramout which share many features among each other, such as their permissibility of consonant clusters, that are not present in the varieties of the valley.

1.1.1 The syllable structure of MA

The syllable structure of MA is different from that of MSA in that it does allow onset clusters, but prohibits coda clusters. The syllable patterns allowed in MA and MSA are listed in table (1) below. Note that the CVC_iC_i structure is allowed in MA only in the case of final geminates, which is beyond the scope of this study.

Table1. Possible syllable patterns in MA and MSA (Al Tairi, 2010: p.5, p.9).

	MA	MSA
CV	[la] ‘no’	[bi] ‘by’
CVC	[min] ‘from’	[man] ‘who’
CVV	[li:] ‘to me’	[li:] ‘to me’
CVVC	[d͡ʒa:d] ‘serious’	[ba:b] ‘door’
$CVC_{(i)}C_{(i)}$	[jadd] ‘hand’	[ʕɪlm] ‘knowledge’
CCV	[bka] ‘(he) cried’	
CCVC	[ktab] ‘(he) wrote’	
CCVV	[ħsa:.ba:t] ‘accounts’	
CCVVC	[fla:n] ‘person (unknown)’	

The sequences of consonants allowed word-initially in MA are similar to the syllable patterns of the Casablanca Moroccan as described in Boudlal (2001), except that the latter allows initial geminates whereas MA does not. In light of what Dell & Elmedlaoui (2002) propose about syllabic consonant in Moroccan Arabic, one can predict that the initial consonant sequences CCV, CCVC, CCVV and CCVVC, listed in Table (1), can better be described as C.CV, C.CVC, C.CVV and C.CVVC respectively. This would comply with the more common syllable patterns across Arabic languages, resembling the syllable structure of MSA (see Table 1).

1.1.2 An Optimality Theoretic account of the clusters of MA

Al Tairi (2010) provides an account for the consonantal sequences in MA within the framework of Optimality Theory (henceforth, OT) (Prince & Smolensky 1993, 2004). The constraints used in that analysis are listed along with their function in (1) below.

(1) List of OT constraints used in Al Tairi (2010: p.23, p.24, p.26, p.27)

a) *COM-ONSET: Onset clusters are prohibited.
b) MAX-IO: All segments in the input have a correspondent in the output.
c) DEP-IO: All segments in the output have a correspondent in the input.
d) SONRISE: Sonority must rise from C to N
e) *COM- CODA: Complex codas are prohibited.
f) SONFALL: Sonority must fall from N to C.

To account for the complex onsets, the markedness constraints *COM-ONSET and SONRISE are ranked below the faithfulness constraints MAX-IO and DEP-IO, as can be seen in tableau (1). This ranking accounts for the tolerance of sequences of consonants in the onset position, which are not restricted by any sonority hierarchy.

Tableau 1. MA Complex Onsets in OT (Al Tairi, 2010: p.25, Tableau 3)

/rfuuf/	MAX-IO	DEP-IO	*COM-ONSET	SONRISE
☞ a) rfuuf			*	*
b) ri.fuuf		*!		
c) fuuf	*!			

As for the complex codas, the markedness constraint *COM-CODA must be ranked above the faithfulness constraint DEP-IO, while SONFALL does not play any role in deciding which sequences are tolerated, as illustrated in tableaux (2a, 2b). Note that candidate (c) violates SONFALL, yet the constraint is irrelevant, as *COM-CODA is enough to predict the correct candidate. This accounts for the prohibition of consonantal sequences which are always separated by an epenthesized short vowel.

Tableau 2. MA Complex Codas in OT (Al Tairi, 2010: p.27-28, Tableau 6)

a	/ramz/	SONFALL	*COM-CODA	DEP-IO
	a) ramz		*!	
	☞ b) ramiz			*

b	/ʃakl/	SONFALL	*COM-CODA	DEP-IO
	c) ʃakl	*!	*	
	☞ d) ʃakil			*

This constraint-based approach provides a good formalization of the positions in which consonant sequencing is allowed in MA, showcasing that there is no sonority hierarchy nor assimilation involved in deciding what sequences of consonants are allowed or not. However, there is a lot to be discovered when it comes to what makes the sequences allowed in one position and prohibited in another. For instance, constraints such as SONRISE and SONFALL are rather descriptive than informative, so much so that an analysis using them yields very little insight into why sonority considerations do not hold value for the language. Simply put, we have an idea of what occurs, but we still lack any deeper understanding of why or how it occurs and what conditions do govern it. In order to do so, one would need to investigate the different tiers of the phonological structure of these syllables (e.g. skeletal structure). It would be worthwhile to inspect the syllables of MA from the perspective of a different theoretical framework, one that is specifically concerned with syllables and syllable structures.

1.2 Theoretical background

1.2.1 Strict CV

Strict CV phonology is a theoretical framework that assumes that a number of tiers are involved in phonological representations (Lowenstamm, 1996; Scheer 2004). These tiers are the skeletal tier, the timing tier, the stress tier and the tone tier. The tier most relevant to this study is the skeletal tier, also known as the autosegmental tier, which has two items: C (onset) and V (nucleus). According to Strict CV, the one and only constituent of this tier is the CV unit. That is to say, every C is followed by a V and, likewise, every V is preceded by a C. Based on this, consonant clusters are analysed to be separated by an empty V slot that does not have a phonetic interpretation, which complies with the Empty Category Principal (ECP) proposed by Kaye, Lowenstamm & Vergnaud (1990). The ECP is defined as follows “A position may be uninterpreted phonetically if it is properly governed” (Kaye et al., 1990: p.219). Government is a lateral relationship between a governor and a governee, where former allows the latter to not have a phonetic realization. These two components of a government relationship constitute what is known as a government domain. As for “proper government”, it is described as a stronger form of government where the governor may not be governed itself and the domain of this government may not include a governing domain. Essentially, this means that the phonological structure of all words ends with a V unit, where the absence of the phonetic representation for this unit indicates that the unit is empty. It is important to note that strictly describing segments in terms of C’s and V’s leaves no room for the traditional notion of a ‘coda’. Within Strict CV, every C unit precedes a V unit, whether either is full or empty. Subsequently, any C unit to the right of a V unit, which would traditionally be described as a coda, is parsed with the following V unit instead. This approach is favored over ‘codas’ and ‘onsets’ especially in analyzing clusters, because it provides generalizable insights into the permissibility of these clusters and lack thereof (Kaye et al., 1990).

The illustration presented in (2) highlights an example of a CV view of clusters in Moroccan Arabic originally provided by Kaye et al. (1990).

(2) C and V final words end in nuclei in Moroccan Arabic (Kaye et al., 1990: p.220)

```
a.   k     t  ɨ  b          consonant-final   word: [ktɨb]   ‘I write’

b.   k  ɨ  t     b  u       vowel-final       word: [kɨtbu]  ‘we write’
     |  |  |  |  |  |
     C  V1 C  V2 C  V3
```

In this example, there are multiple empty units. In (2a) V1 and V3 are empty, while in (2b) V2 is empty. Notice that the emptiness of V3 in (2a) results in the word being consonant-final phonetically. The permissibility of final empty units is accounted for by prosodic licensing (better known as “p-licensing”), which will be explained in the following subsection.

1.2.2 P-licensing and consonant clusters

In Strict CV, consonant clusters are caused by word-internal empty nuclei, which are permitted not to have any phonetic representation, by means of government. However, word-final empty nuclei do not have governors and yet they are not phonetically realized, thanks to prosodic licensing (p-licensing). Final empty nuclei (FEN) such as the one in example (2a) are p-licensed in languages that allow consonant final words. Some languages do not allow consonant-final words, which is proposedly attributed to the fact that these languages do not p-license FENs, meaning that their FENs must have a phonetic interpretation (Faust & Ulfsbjorninn, 2018).

It is important to make the distinction between governed empty nuclei and p-licensed empty nuclei, with the main difference being that p-licensing occurs on the prosodic level, whereas government occurs on the autosegmental level. Another important distinction to make is between p-licensing and regular (government) licensing, as the former allows the constituent that it licenses to remain unrealized, whereas the latter allows it to be realized.

Kaye et al. (1990) indicate that word-internal empty nuclei are permitted, by means of government, when they are followed by a full nucleus. In other words, a nucleus is allowed to be empty if it precedes a nucleus that is full. This is well illustrated by Faust & Ulfsbjorninn (2018: p. 566) who provided an illustrative example from Arabic of a word that has both types of relationships; government and licensing. This example is presented in figure (3), where the solid arrow indicates that the filled V3 governs the empty V2 and the dotted arrow indicates that V3 licenses its onset. Note that we added the hashtag symbol (#) indicating a word boundary, to illustrate the prosodic licensing of the FEN, while remaining agnostic to its epistemological status; i.e. whether it is a proper phonological symbol on par with C and V or not. The hashtag symbol is used because it is a convenient way to represent the effect that word boundaries exert on final constituent of a word.

(3) The skeletal structure of an Arabic word illustrating government and licensing (Faust & Ulfsbjorninn, 2018: p.566)

k	a	t		b	a	t		
|	|	|		|	|	|		
C	V1	C	V2	C	V3	C	V4	#

→ Government

→ P-licensing

1.3 The aim

This study aims to provide theoretical explanations to the permissibility of word-initial and word-final consonantal sequences of the MA variety by investigating them from the perspective of Strict CV. Al Tairi (2010) has given an OT account for these sequences, explaining what sequences are allowed and in which positions. However, there is still a lot to be discovered regarding the rules governing these sequences and what causes them to be permitted in certain positions and what limits them from being permitted in others.

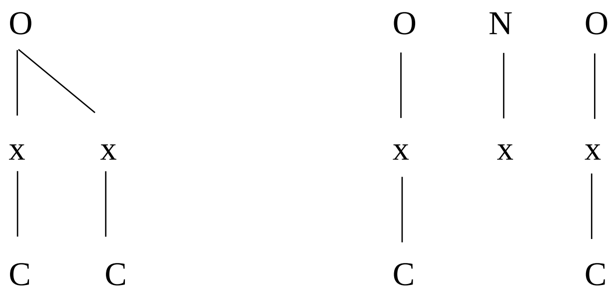
In this paper, we will explore the skeletal structure of the CC sequences of MA, in both word-initial and word-final positions, in an effort to provide a deeper understanding of the permissibility of the consonantal sequences in this variety of Arabic.

2. Methods

As stated previously, this study takes a novel approach towards investigating the CC sequences of Mukallawi Arabic (MA). Where previous studies provided formal description of the phenomenon in this variety within the framework of Optimality Theory (Al Tairi, 2010), this study is concerned with showing how the permissibility of such consonantal sequences can be predicted from general principles, such as; Government, Licensing and the Empty Category Principle, which were introduced in previous sections. Since we are investigating adjacent consonants and whether they form real or bogus clusters, we use the term consonantal sequence to refer to all instances of adjacency between two consonants on the surface level. The distinction between real and bogus clusters was first introduced by Harris (1997) as a way to formulate whether a consonantal sequences forms a true complex clusters such as that in (4a) or rather a fake cluster belonging to two different syllables and separated by an empty nucleus as in (4b). While we do not necessarily subscribe to Harris' (1997) analysis in this paper, we find in it the motivation to approach the consonants clusters of MA from the perspective of a theoretical framework that views all clusters as bogus, which is Strict CV.

(4) Illustration of the difference between complex onsets and and bogus clusters (Harris, 1997: p. 397)

a. Complex onset b. Bogus cluster



To investigate whether two occupied consonantal positions are truly adjacent or not, we need to look at the phonotactic dependencies between the two. According to Harris (1997; p.331), "lack of phonotactic dependency can be taken as a sign of lack of adjacency''. That is to say, the two positions need to adhere to syllable contact laws, where a coda has a narrower distributional latitude than the onset following it; fewer sets of consonants are permitted in a coda position in comparison to the following onset. If that is not the case in a sequence of consonants, and the first consonant has a wider distributional latitude than the following one, then there is no reason to assume the two consonants to be underlyingly adjacent to one another. After all, surface-adjacency could result from the combination of the end of one word and the beginning of another, meaning that surface-adjacency does not entail underlying-adjacency. The idea is that the rules governing the distributional properties of syllables should not be expected to extend their influence beyond their own syllable. Similarly, syllables should not be expected to be influenced by the phonotactics of other syllables.

An example of lack of adjacency is the disobedience of sonority sequencing constraints. In general, onsets require a rise in sonority where the first constituent of a consonantal sequence in onset position is less sonorant than the second one. Failure to adhere to this constraint is an indicator that these consonants are not adjacent in the higher levels of representation (Harris, 1997).

This is supported by further evidence in languages that exhibit vowel-zero alternations, where consonantal sequences are broken up by a vowel in certain derived forms. According to the resyllabification analysis, first proposed by Clements and Keyser (1983) and cited by Harris (1997), in the case of CC-sequences in word-final position, the consonants are parsed heterosyllabically. The first of the two consonants is syllabified as the coda of the preceding syllable, while the second one is labelled as extraprosodic. Then, an epenthesized vowel invites the first consonant to its onset and the second one to its coda. An example of derivation in Turkish demonstrates this, where the accusative form [kabr-i] 'tomb' has a CC-sequence while the nominative form [kabir] has an epenthesised vowel separating the final two consonants. How the resyllabification is indulged by the vowel epenthesis is illustrated in (5) below:

(5) Resyllabification by means of vowel epenthesis in Turkish

$$kab.<r>_{ep} \rightarrow kab.i<r>_{ep} \rightarrow ka.bir.$$

(Harris, 1997: p. 331-332)

A similar analysis can be applied to the word-intial CC-sequences, where the first consonant of the sequence is extraprosodic, while the second one is syllabified as the onset of the following syllable. Afterwards, the epenthesis of the vowel between the two consonants invites the first consonant (previously extraprosodic) to its onset and the second one (previously onset) to its coda.

The following subsection will introduce the MA data that will be used in the analysis as well as a number of preliminary observations about the possible consonant combinations in word-initial CC sequences in the variety. Furthermore, it presents predictions about the phonotactics governing the permissibility of these sequences.

2.1 Data and observations

Most of the MA data used in this study were presented first in Al Tairi (2010), while some of the data were provided and transcribed by the author, who is a native speaker of MA himself, and they were validated to be grammatical words of MA by three other native speakers of the variety, who were naive to the aims of the study to avoid the influence of bias.

The categories of consonants that can form a word-initial CC sequence with one another in MA do not seem to have any restrictions based on latitude or sonority sequencing, as can be seen in (6). A consonant from any category (voice, place or manner), in initial position, can be followed by a consonant from any other category, indicating that combinations are symmetrical (see appendix 1 for the full consonantal inventory of MA). That is to say, relative order of the consonants does not pose any prohibition on the word-initial consonantal sequences.

(6) Permitted consonant sequences word-initially in MA.

(a) Initial liquids

Followed by a nasal	[rmaħ] '(he) kicked'
Followed by a voiced fricative	[lʕɪb] '(he) played'
Followed by a voiceless fricative	[rsɪb] '(he) failed'
Followed by a voiced stop	[lgam] '(he) was fed'
Followed by a voiceless stop	[lkad] '(he) stacked'

(b) Initial nasals

Followed by a liquid	[mrax] '(he) slapped'
Followed by a nasal	[mnaʕ] '(he) prohibited'
Followed by a voiced fricative	[nzal] '(he) went down'
Followed by a voiceless fricative	[msɪk] '(he) held'
Followed by a voiced stop	[ngal] '(he) moved'
Followed by a voiceless stop	[nkaʃ] '(he) dug up'

(c) Initial voiced fricatives

Followed by a liquid	[ʁraf] '(he) scooped'
Followed by a nasal	[ʕnɪd] '(he) persisted'
Followed by a voiced fricative	[zʕɪl] '(he) got upset'
Followed by a voiceless fricative	[ʁsal] '(he) washed'
Followed by a voiced stop	[ʕbar] '(he) walked by'
Followed by a voiceless stop	[ʕkar] '(he) annoyed'

(d) Initial voiceless fricatives

Followed by a liquid	[frag] '(he) divided'
Followed by a nasal	[smar] '(he) stayed up'
Followed by a voiced fricative	[xzan] '(he) stored'
Followed by a voiceless fricative	[fʃɪl] '(he) failed'

Followed by a voiced stop	[ʃbak] ‘(he) connected’
Followed by a voiceless stop	[ftˤar] ‘(he) had breakfast’

(e) Initial voiced stops

Followed by a liquid	[blaʕ] ‘(he) swallowed’
Followed by a nasal	[gmar] ‘(he) scammed’
Followed by a voiced fricative	[bzag] ‘(he) spitted’
Followed by a voiceless fricative	[dfar] ‘(he) pushed’
Followed by a voiced stop	[gdɪr] ‘(he) managed to do’
Followed by a voiceless stop	[btar] ‘(he) amputated’

(f) Initial voiceless stops

Followed by a liquid	[kraf] ‘(he) swooped’
Followed by a nasal	[knaf] ‘(he) cleaned’
Followed by a voiced fricative	[tʕɪb] ‘(he) got tired’
Followed by a voiceless fricative	[kʃaf] ‘(he) revealed’
Followed by a voiced stop	[kbər] ‘(he) grew big’
Followed by a voiceless stop	[ktab] ‘(he) wrote’

The example words used in (6) are all in the past tense in its masculine form. When we compare those forms to their feminine equivalents, as we do in (7) below, we observe that the resulting pairs exhibit vowel-zero alternation. In these examples, the symbol (-) denotes a morpheme boundary, while the symbol (.) denotes a syllable boundary.

(7) Vowel-zero alternation in feminine and masculine forms of the past tense in MA.

	i. MASC	ii. FEM	
(a)	[rmaħ]	[ram.ħ-at]	‘kicked’
	[lʕɪb]	[lɪʕ.b-ɪt]	‘played’
	[rsɪb]	[rɪs.b-ɪt]	‘failed’
	[lgam]	[lag.m-at]	‘was fed’
	[lkad]	[lak.d-at]	‘stacked’
(b)	[mrax]	[mar.x-at]	‘slapped’
	[mnaʕ]	[man.ʕ-at]	‘prohibited’
	[nzal]	[naz.l-at]	‘went down’
	[msɪk]	[mɪs.k-ɪt]	‘held’
	[ngal]	[nag.l-at]	‘moved’
	[nkaʃ]	[nak.ʃ-at]	‘dug up’
(c)	[ʁraf]	[ʁar.f-at]	‘scooped’
	[ʕnɪd]	[ʕɪn.d-ɪt]	‘persisted’
	[zʕɪl]	[zɪʕ.l-ɪt]	‘got upset’

	[ʁsal]	[ʁas.l-at]	‘washed’
	[ʕbar]	[ʕab.r-at]	‘walked by’
	[ʕkar]	[ʕak.r-at]	‘annoyed (with smell)’
(d)	[frag]	[far.g-at]	‘divided’
	[smar]	[sam.r-at]	‘stayed up’
	[xzan]	[xaz.n-at]	‘stored’
	[fʃɪl]	[fɪʃ.l-ɪt]	‘failed’
	[ʃbak]	[ʃab.k-at]	‘connected’
	[ftˤar]	[fatˤ.r-at]	‘had breakfast’
(e)	[blaʕ]	[bal.ʕ-at]	‘swallowed’
	[gmar]	[gam.r-at]	‘scammed’
	[bzag]	[baz.g-at]	‘spitted’
	[dfar]	[daf.r-at]	‘pushed’
	[gdɪr]	[gɪd.r-ɪt]	‘managed to do’
	[btar]	[bat.r-at]	‘amputated’
(f)	[kraf]	[kar.f-at]	‘swooped’
	[knaf]	[kan.f-at]	‘cleaned’
	[tʕɪb]	[tɪʕ.b-ɪt]	‘got tired’
	[kʃaf]	[kaʃ.f-at]	‘revealed’
	[kbər]	[kəb.r-ɪt]	‘grew big’
	[ktab]	[kat.b-at]	‘wrote’

All the past tense words in (7) have triconsonantal roots separated by one vowel, between the first two consonants in the feminine form and between the latter two in the masculine form. Notice that the feminine form has a suffix, resulting in it containing one more syllable than the masculine form, which makes the final two consonants of the root heterosyllabic, where the final consonant of the root is the onset of the second syllable. For example, the masculine form of (8a) [rmaħ] has its first (and only) vowel after the second consonant, whereas the feminine form [ramħat] has its first vowel after the first consonant and no vowel after the second consonant. This provides evidence to the speculation that each consonantal position is followed by a vowel position that is filled in certain contexts and left empty in others.

If we compare the MA and MSA masculine forms of the past tense, for example [lʕɪb] and [laʕɪba] ‘(he) played’ respectively, we see that in MSA all vowel positions are filled, whereas in MA only one is filled and two are empty. This suggests that MA favors empty nuclei wherever possible. We already know that MA prohibits word-final ‘coda’ sequences (Al Tairi, 2010), which is why it makes sense that when there are no suffixes, as in the masculine form, the vowel separates the final two consonants of the root, instead of the first two. Whereas in forms that do have a suffix, such as the feminine forms, separating the final two consonants of the root by a vowel is not necessary anymore, as the suffix ensures that the consonants are heterosyllabic. For instance, where the full vowel position in [lʕɪb] '(he) played' is essential for preventing a word-final sequence, it is not important in [lɪʕ.bɪt] '(she) played', because in this form, /ʕ/ belongs to a different syllable than /b/ which is now the onset of the second syllable. However, it is noteworthy that the full vowel maintains its position, in the sense that it assumes the position separating the last two consonants before the word boundary in both cases, for the same purpose of avoiding the word-final CC-sequence.

Word-final CC-sequences are not the only prohibited sequence of consonants in the syllables of MA. Assuming that MA allows only one vowel in a suffix-less triconsonantal root, with a prohibition on word-final CC-sequences, one would expect the hypothetical form [lʕbɪ] to be an acceptable one. However, that is not the case, MA seems to pose a prohibition on the so-called "tautosyllabic" triconsonantal sequences. If we look at the masculine form of the past tense, e,g, [lʕɪb] '(he) played' and compare it to the same form with the causative case marker, e.g. [laʕ:ab] '(he) caused to play'. Ignoring the difference in vowel quality, notice that the causative form displays gemination of the second consonant of the root, meaning that a hypothetical form of the causative would be [lʕ:ab]. However that form is incorrect, despite having only one full vowel position. Instead, the correct form [laʕ:ab] breaks up the triconsonantal sequence by filling the vowel position after the first consonant. Therefore, it is concluded that triconsonantal sequences are prohibited in MA.

Closer inspection of the data provides further evidence to the lack of influence of a sonority hierarchy in the sequencing of consonants in MA. If you compare the pattern of the vowel-zero alternation in any of the words in (8a), which do not follow the sonority hierarchy, to the pattern in any of the words in (8f) which do comply with it, you will notice that the pattern is the exact same. This suggests that the sonority hierarchy is entirely irrelevant to deciding which vowel position will be filled and which position will be empty. The decisive factors of the empty or full vowel positions are the number of 'adjacent' consonantal constituents, as no more than two consonants are allowed in a sequence, as well as the position of the sequence, as sequences are never allowed in word-final positions.

In terms of government and licensing, this can be generalized by the claim that, in MA, each full vowel position can govern one and only one empty vowel to its left side, as will be further illustrated in the next chapter.

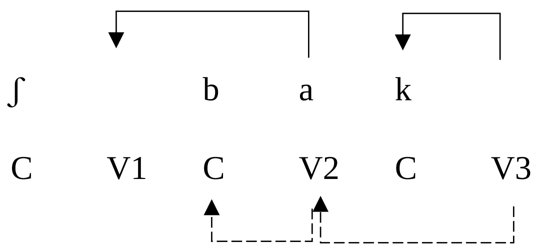
3. Analysis

3.1 Word-initial sequences

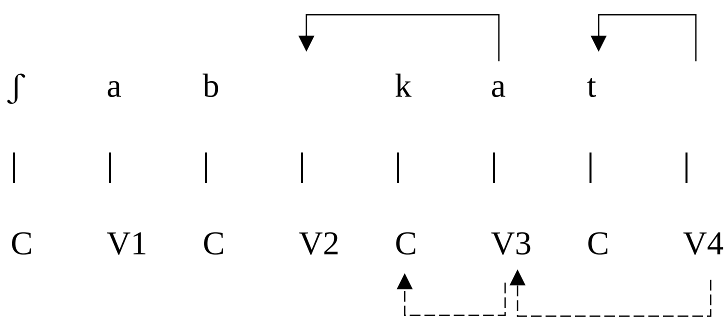
As introduced in previous sections, the word-initial consonant sequences of Mukallawi Arabic (MA) exhibit vowel-zero alternations, which is an indicator of lack of phonotactic dependency between the two consonants of the sequence. The figures of (8) below take the feminine and masculine forms of the past tense, [ʃbak] and [ʃab.kat] ‘he/she connected’ respectively, as an example. Both forms stem from the triconsonantal root [ʃbk], where the masculine form exhibits a consonantal sequence followed by a vowel, while in the feminine form the first two consonants are broken up by a vowel.

(8) Skeletal structures of MA words

(a) Masculine form of past tense, with word-initial consonantal sequence



(b) Feminine form of past tense, with no word-initial consonantal sequence



→ Government

⇢ Licensing

As mentioned previously, MA has a tendency to allow consonantal sequences word-initially, in environments where these sequences would usually be split up by a vowel in most other varieties of Arabic. If we view things from the perspective of these other varieties, one can go as far as to claim that MA 'omits' the first interconsonantal vowel of its words. For example, the Modern Standard Arabic (MSA) word [rama] '(he) threw' has an equivalent in MA without the first vowel [rma] '(he) threw'. However, from the point of view of Strict CV, which is a more neutral one, MA does not necessarily 'omit' the vowel, but rather has an empty nucleus that is governed by the following vowel, leading to the vowel not being realised in the variety, as illustrated in (9) below:

(9) Full nucleus governs preceding nucleus in MA

```
r         m    a
|    |    |    |
C    V1   C    V2
```

→ Government

⇢ Licensing

On the other hand, the nuclei of MSA do not exhibit a similar government relationship, as shown in (10) where the first vowel of the word is required to have a phonetic realization:

(10) Full nucleus does not govern preceding nucleus in MSA

```
r    a    m    a
|    |    |    |
C    V1   C    V2
```

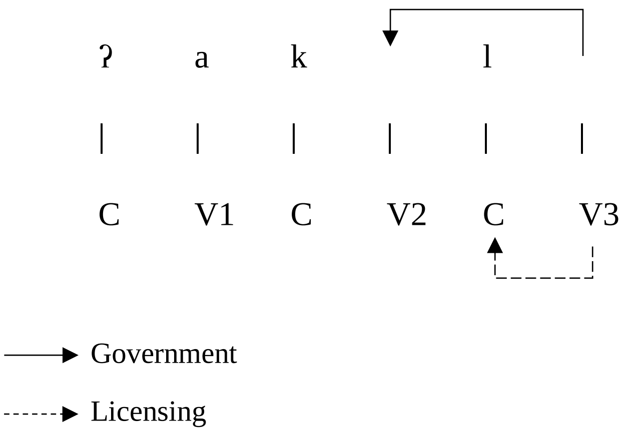
→ Government

⇢ Licensing

3.2 Word-final sequences

Contrary to word-initial sequences, word-final sequences are prohibited in MA and permitted in other varieties, such as MSA. Again, from the perspective of MSA, MA has the tendency to 'epenthesize' a vowel in between the two consonants of a word-final sequence. For instance, the MSA word [ʔakl] 'food' has an equivalent in MA where the final two consonants are separated by a vowel [ʔakəl] 'food'. When we view things from the perspective of Strict CV, however, it is clear that the Final Empty Nucleus (FEN) in MSA governs the preceding empty nucleus between the last two consonants, while the FEN in MA does not. In other words, MA does not necessarily epenthesize a vowel in between consonants, but it requires the vowel to be phonetically realised, whereas MSA does not, as illustrated by (11) below.

(11) FEN governs preceding nucleus in MSA



In (11) the FEN licenses the word-final consonant as indicated by the dashed line, while the governmental capability of the FEN extends to the preceding syllable, as V2 is an empty nucleus that is governed by the FEN.

The licensing and government relationship in MA word-finally is slightly different. It is similar to MSA in that it has a FEN that governs its consonant as the final constituent in the word with a phonetic representation. However, it is different in that the FEN's governmental capability is restricted to its own CV unit. The FEN in MA, contrary to MSA, does not govern the vowel of the preceding CV unit, which has a phonetic representation as licensed by the FEN, as illustrated in figure (12).

(12) FEN does not govern preceding nucleus in MA

ʔ	a	k	ə	l	
|	|	|	|	|	|
C	V1	C	V2	C	V3

$\longrightarrow$ Government

$\dashrightarrow$ Licensing

3.3 Triconsonantal sequences

Consonantal sequences of more than two consonants is not tolerated in MA. For instance, the triconsonantal root of the verb [lʕb] ‘to play’ will always have a vowel breaking up the sequence of consonants in all of its forms. The location of the vowel depends on the government and licensing of other nuclei from the right edge, as no two consecutive nuclei may be empty. For example the masculine form of the past tense of [lʕɪb] ‘(he) played’ has a vowel between the second and third consonant, whereas the feminine form of the past tense [lɪʕbɪt] has a vowel between the first and second consonant. The figures of (13) illustrate how the position of the vowel depends on the prohibition of have two empty nuclei after one another in both forms.

(13) Prohibition of triconsonantal clusters in MA

(a) Masculine form of past tense

l		ʕ	ɪ	b	
C	V1	C	V2	C	V3

(b) Feminine form of past tense

l	ɪ	ʕ		b	ɪ	t	
|	|	|	|	|	|	|	|
C	V1	C	V2	C	V3	C	V4

⟶ Government

------▸ Licensing

An attentive reader would have noticed that the figures of (13), which we have used to illustrate the prohibition on triconsonantal sequences, are very similar to the figures of (8) which were used to illustrate the permissibility of word-initial sequences. This is because both conditions stem from the prohibition on allowing to empty vowel positions after one another. One can go as far as to claim that both of them are consequences of the government and licensing abilities of the FEN of MA, which will be discussed further in the following subsection.

3.4 The FEN of MA

A lot of the observations made in this paper about the consonantal sequences of MA allow us to make certain statements about the governmental properties of the FEN of the variety, as compared to those of MSA and other varieties of Arabic. Which, in turn, would shed light on what makes the behavior of the consonant clusters of this variety so different from that of most other varieties of Arabic.

First of all, we have seen earlier in example (9) among others, that the FEN of MA does not have the ability to govern the preceding nucleus as the FEN of MSA does. Instead, the FEN of MA requires the preceding nucleus to have a phonetic representation. While full nuclei of MA have the ability to govern the preceding nucleus allowing it to not have a phonetic representation, the empty nuclei, including FEN's, of MA do not possess that property.

Secondly, and in light of the first point on the FEN of MA, we can deduct an answer to the question of why the variety does not allow tautosyllabic triconsonantal clusters. As we have touched on in earlier sections, forms such as [ʃbkat] are forbidden in MA. Instead, we get forms such as that in example (9b) [ʃabkat]. This can be explained by the fact that the language does not allow two empty nuclei after one another. This is prevalent in that the FEN does not govern the preceding nucleus, because that would result in two consecutive empty nuclei. The prohibition of triconsonantal sequences is another clear example of the language not allowing two adjacent nuclei to not have a phonetic representation.

4. Conclusion and discussion

This study was conducted with the objective of providing a better understanding and a clear image of what sets Mukallaw Arabic (MA) variety apart from most other varieties by answering some of the theoretical questions surrounding the permissibility of consonant clusters. Viewing the phenomena from the point of view of Strict CV made it possible to uncover relationships between different syllabic constituents and CV units that are relevant in permitting and prohibiting the clusters of MA. Namely, thanks to Strict CV, it is clear that all the conditions of permitted and prohibited clusters follow naturally from the prohibition on having two empty nuclei consecutively. Such conclusion was not and cannot be reached by analysis using branching onsets without the aid of additional theoretical tools, which risks circularity. Furthermore, it was known that Mukallawi Arabic (MA), as well as a few other varieties of Arabic such as; Casablanca Moroccan Arabic, deviate from a widespread pattern among different varieties of Arabic in that they allow word-initial clusters and prohibit word-final clusters. Comparing how the government and licensing relationships fare in MA with how they do in Modern Standard Arabic (MSA) has shed light on what sets MA apart from other varieties of Arabic. This study builds upon the work of Al Tairi (2010) and expands it, by asking and answering questions that were not addressed in the previous study.

This study has found that there is absolutely no hierarchical difference between the first and second consonant of the cluster, as evidenced by vowel-zero alternations and the distributional capabilities of C1 and C2. These alternations are attributed to the fact that the language does not allow two consecutive nuclei to be empty. In addition, a full nucleus in MA has the ability to govern the preceding nucleus allowing it to be empty, even when that nucleus is the first one after a word boundary, resulting in a word-initial consonantal sequence. Whereas in MSA, the first nucleus after a word boundary is obligatorily full, as the following (full) nucleus cannot govern it, posing a prohibition on consonantal sequences word-initially. It is noteworthy to emphasize that an analysis using branching onsets would have trouble explaining how triconsonantal clusters are prohibited while word-initial clusters are not. For instance, it would seem contradictory to say that a structure such as CVC.CCVC is ill-formed when both CVC and CCVC are accepted structures in the language. Such studies would need to invoke external specialized tools in order to account for that, which is often rather suspicious. Whereas, using Strict CV, one can show that the location of the CC cluster is gauged from the right edge; the fact that word-initial clusters are permitted is epiphenomenal.

Furthermore, the study shows that the reason why MA does not permit word-final consonantal sequences is because its final empty nucleus (FEN) does not possess the ability to govern the preceding nucleus, which separates the final two consonants before the word boundary. Instead, the FEN of MA licenses the preceding nucleus, thereby requiring it to be full. On the other hand, the FEN of MSA, as well as many other varieties, have the ability to govern the preceding nucleus, allowing it to be empty, which results in the final two consonants colliding, forming a word-final consonantal sequence.

In conclusion, the paper provides new insights into the permissibility of consonant sequences in MA, which highlight their dependency on government and licensing relationships between nuclei and onsets across and within CV units. The findings of this paper would be supported if similar government and licensing pattern are found to be displayed by the nuclei and onsets of the few varieties that are similar to MA in their tolerance of word-initial sequences and intolerance of word-final sequences, such as Casablanca Moroccan Arabic. Finally, it would be even more interesting to see if Strict CV can shed light on whether there are nuanced differences between how these ‘similar’ varieties permit their consonantal sequences.

5. Appendix

Appendix 1: Consonantal inventory of MA

Consonants of MA (Al Tairi, 2010: p.8).

	Bilabial	Labio-dental	Dental	Alveolar	Emphatic Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	b			t d	tˤ dˤ			k g			ʔ
Nasal	m			n							
Trill				r							
Tap				(ɾ)							
Fricative		f	(θ) (ð)	s z	sˤ	ʃ			χ ʁ	ʕ ħ	h
Affricate							(d͡ʒ)				
Approximant	w										
Lateral approximant				l			j				

6. Bibliography

- Al-Saqqaf, Abdullah. 1999. *A descriptive linguistic study of the spoken Arabic of Wadi Hadramout, Yemen*. Doctoral dissertation, University of Exeter.

- Al Tairi, H. 2010. *Consonant Clusters in Mukallaene Arabic Dialect: Observations and an OT Account* (Doctoral dissertation, Oakland University).

- Boudlal, Abdelaziz. 2001. *Constraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. Doctoral dissertation, Mohammed V University, Rabat, Morocco.

- Clements, G. N. & S. J. Keyser (1983). *CV phonology: a generative theory of the syllable*. Cambridge, Mass.: MIT Press

- Dell, François, and Elmedlaoui, Mohamed, 2002. Syllables in Tashlhiyt Berber and in Moroccan Arabic (Kluwer International Handbooks in Linguistics 2). Dordrecht & Boston: Kluwer Academic Publishers, 2002. Pp. xvi+ 384. ISBN: 1-4020-1076-1. *Journal of the International Phonetic Association*, *34*(2), 215-217.

- Faust, Noam, & Ulfsbjorninn, Shanti. 2018. Arabic stress in strict CV, with no moras, no syllables, no feet and no extrametricality. *The Linguistic Review*, 35(4), 561-600.

- Haddad, Ghassan. 1984. Epenthesis and sonority in Lebanese Arabic. *Studies in the linguistic sciences* 14(1): 57-88.

- Haddad, Yousef. 2006. Dialect and standard in second language phonology: The case of Arabic. *SKY Journal of linguistics* 19:147-171.

- Harris, John. 1997. Licensing Inheritance: an integrated theory of neutralisation. *Phonology, 14*(3), 315-370.

- Harris, John & Edmund Gussmann. 2002. Word-final onsets. *UCL Working Papers in Linguistics 14*. 1–14.

- Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: The university of Chicago press, Ltd.

- Ito, Junko. 1989. A prosodic theory of epenthesis. *Natural language & linguistic theory 7*(2):217-259.

- Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud. 1990. Constituent structure and government in Phonology. *Phonology 7*. 193–231.

- Lowenstamm, Jean. 1996. CV as the only syllable type. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology models and methods*, 419–442. European Studies Research Institute, Manshester: University of Salford.

- Prince, Alan & Smolensky, Paul. 1993. *Optimality Theory: Constraint interaction in generative grammar*. (Technical report no.Ruccs-TR-2) New Brunswick, NJ: Rutgers University Center for Cognitive Science.

- Prince, Alan & Smolensky, Paul. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

- Scheer, Tobias. 2004. *A lateral theory of phonology. 1: What is CVCV, and why should it be?* (Studies in generative grammar ; 68.1. 790318873). Berlin [etc.]: Mouton de Gruyter.

- Watson, Janet. 2002. *The phonology and morphology of Arabic*. New York: Oxford University press.