On the treatment of super-heavy syllables in Arabic Dialects: an Optimality Theoretic approach to historical typology

Hassan Bokhari
Indiana University Bloomington

Abstract

This study provides a historical typological Optimality Theoretic analysis of the treatment of potential super-heavy syllables in six Arabic varieties: Hijazi, Egyptian, Emirati, Kuwaiti, Algerian, and Palestinian. The analysis in this study uses the same violable OT constraints for all languages, and the differences between the grammars are represented by the order in which the constraints are ranked relative to one another. The similarities and differences between these varieties are examined from the point of view of one approach to historical OT (Cho 1998), which states that individual pairs of constraints may be ranked or unranked in relation to one another, one operation at a time, meaning that switching the order of two constraints takes two steps historically. According to Cho (1998, 45), "each step of a sound change should be viewed as a change in the ranking of constraints." Cho's approach in detecting the historical typological differences between varieties by counting the steps of constraint reranking is compatible with a common approach to historical linguistics. Specifically, Wichmann et al. (2010) provide a quantitative method for determining the geographic homeland of a group of related languages, which takes into account a simple linguistic-difference metric and the geographic distance between the languages. Using constraint reranking in place of Wichmann et al.'s linguistic-difference metric to calculate the homeland of Arabic dialects results in an area around Hijaz as the homeland of Arabic dialects, since Hijazi, Egyptian and Emirati dialects form a cluster of geographically close, but linguistically diverse dialects.

1   Introduction

This paper considers the typology of how Arabic dialects deal with the potential generation of super-heavy syllables word-internally. Super-heavy syllables are syllables with at least three timing slots in the rhyme,

including CVVC, CVCC, and CVC:.[1] Syllables like these are problematic for a language such as Arabic, since normally three moras would be associated with them, one for each slot in the rhyme, while a syllable is in principle restricted to having two moras. To avoid making word-internal potential super-heavy syllables tri-moraic, some dialects of Arabic epenthesize a vowel allowing a syllable-final consonant to be resyllabified (creating only two timing slots in the rhyme), while other dialects shorten the vowel (again, creating only two timing slots in the rhyme), and yet other dialects allow super-heavy syllables to surface faithfully, but they fail to associate a mora with the final slot (i.e., they allow three timing slots in the rhyme, but still only allow two moras per syllable). However, potential super-heavy syllables are allowed word-finally in almost all Arabic dialects, since the last timing slot in a word-final rhyme is extra-metrical — i.e., it can never bear moraic weight. A potential super-heavy syllable is any syllable that would normally surface with three timing slots in the coda; however, due to phono-logical processes like those mentioned above, some or all potential super-heavy syllables in a language may never surface this way. It should be noted that there are no syllables underlyingly (Davis 2011), so the term "potential super-heavy syllable" only considers what the default syllable structure would be like, before any other phonological processes.

This paper provides an Optimality Theoretic (OT) analysis of four Arabic dialects based on a representative data set. These analyses will use the same violable constraints, and the differences between the grammars of the languages will be represented by the order in which the constraints are ranked relative to one another. The constraints used in this study are all motivated empirically in other studies, and are normally posited as being universal to all languages. The relative ranking of the constraints to one another will account for why the different dialects treat super-heavy syllables in different ways.

The similarities and differences between these analyses are examined from the point of view of one classical approach to historical OT (Cho 1998). This additional level of analysis helps explain how these different grammars came about and how they are related to one another.

Section 2 reviews the OT literature on potential super-heavy syllables in Arabic dialects, as well as the literature on historical approaches to OT. Section 3 presents the methodology used in this study, and background information about the participants. The data that was gathered is presented in section 4, and analyzed in section 5, which also outlines the analytical methods used, including the constraints and historical methods. Section 6 discusses the analysis and concludes.

---

[1] C = Consonant, V = Vowel, C: = Geminate.

2   Literature Review

For this research, several different types of literature on phonology have been consulted. General literature on phenomena like epenthesis, moras, and super-heavy syllables includes Kager (1999) and Davis (2011). Previous sources which provide analyses of overlapping phenomena in various Arabic dialects include Abu-Mansour (1991) on Meccan Arabic syllable structure and epenthesis; Bamakhramah (2009), an OT analysis of weight and epenthesis in Hadhrami Arabic and Meccan Arabic in comparison to Classical Arabic; Kabrah (2014) on the stress system of Meccan Arabic, which is partially dependent on the weight system; Kiparsky (2003), a typology and high-level analysis of the syllabification patterns of a number of Arabic dialects; Alqahtani (2014) on non-final super-heavy syllables in Najdi Arabic, and Al-Mohanna (2008) on syllabification in Hijazi Arabic. Sources consulted on historical OT include Cho (1998), Bermúdez-Otero (2006), Holt (2015), and Rees (2008).

According to, among others, Holes (2004, 18–19), from the second half of the 7th century to the beginning of the 8th century, Arabic spread from the Arabian Peninsula to an area that stretched from Spain to Persia. This quick expansion of the language into the geographical areas where it is spoken in modern times is not when the modern dialects began to differentiate from one another. Instead, Arabic as spoken in each area gradually changed to its modern form. However, as Holes (2004, 3–4) points out, modern Arabic dialects form a complicated sort of dialect continuum, where speakers from nearby areas speak similar varieties, but speakers from distant areas speak very different varieties. According to Bassiouney (2009, 20), there are five main dialect groups in modern Arabic: Arabian dialects, Mesopotamian dialects, Levantine dialects, Egyptian dialects, and Maghreb (or Western) dialects — as shown in Figure 1.

Figure 1: A map representing the five main dialect groups of modern Arabic

While the work consulted regarding the phenomena that this paper examines offer somewhat different approaches, the general approach

that I take to the question of how syllable weight can be dealt with in Arabic involves the interaction between markedness and faithfulness constraints. For example, the word /bint-na/ surfaces with epenthesis as ['bintana] in Hijazi Arabic because $*\sigma_{\mu\mu\mu}$ (a markedness constraint which prohibits syllables from surfacing with three moras) outranks DEP-V (a faithfulness constraint that prohibits forms from surfacing with a vowel that was not present in the underlying representation). If these two constraints were ranked in the opposite order, then /bint-na/ would surface faithfully as ['bintna]. One tenet of classical Optimality Theory is that all languages use the same set of universal constraints, and grammars may only differ in terms of their ranking of the constraints. While a constraint may not be deleted, it may be so low-ranked that it is not active in a language — i.e., it has no real impact on the grammar; conversely, a constraint that is not active in one language may be active in another language simply due to the fact that it is highly ranked, or even undominated, and not because that language's grammar has an additional constraint.

Work in historical OT assumes that when a language's phonology changes, what is happening is the constraints are being reranked, since according to OT, constraints cannot be deleted or added. There are different theories as to how, why, and in what way this happens. For example, one approach assumes that constraints are normally 'demoted,' meaning they lose their domination relationship over another constraint. Another approach (Cho 1998) assumes that individual pairs of constraints may be ranked or unranked in relation to one another, one operation at a time, which means that switching the order of two constraints takes two steps historically. Parallel OT contrasts with rule-based historical phonology, where a list of sound changes is simply described.[2] One advantage of the OT approach is that it can explain the relationship between different sound changes in a language that would normally not be relatable. Other issues in historical OT relate to lexicon-optimization, i.e., the fact that the lexicon is restructured in OT as needed to produce input forms that most efficiently derive the current output form. Some previous work also argues that reranking happens when the evidence for a particular ranking is not very robust in a language. This may be the case for super-heavy syllables in modern Arabic dialects, since Classical Arabic did not have good evidence for how to treat word-internal super-heavy syllables. This problem is described in more depth later in this paper.

---

[2] Harmonic Serialism and Stratal OT can be utilized to replicate rule-based historical phonology or ordered diachronic phonological change.

Bermúdez-Otero (2006) raises some important questions related to historical approaches in OT. The main issue that Bermúdez-Otero (2006, 498) raises is that markedness may simply be a side effect of a series of sound changes, which are caused by simple restrictions on phonetic mechanisms. The sound changes, in turn, are caused by speakers "misparsing … phonological structures that pose articulatory, acoustic, or auditory difficulties", reminiscent of Haspelmath's (2006) claim that phonological markedness is in fact usually just a way of talking about phonetic difficulty.[3] An example of this is contextual voicing neutralization, where many languages devoice codas, but preserve a voicing contrast in onsets. The explanation provided for languages like this relies on the fact that voicing is harder to perceive in coda position than in onset position, so that if voicing is neutralized in one context, it will be in the coda, and not the onset. The normal expression of this as a markedness constraint is said to be problematic because the 'markedness' of voicing in the coda is simply due to perceptual difficulty.

In Cho's (1998, 45) OT approach, "each step of a sound change should be viewed as a change in the ranking of constraints." This approach means that "both synchronic variation and change are based on the same phonological principles" (Cho 1998, 45). According to this approach, there are only two types of change that can occur: creating a ranking between two unranked constraints, or undoing a ranking between two ranked constraints. Hence, a reversal in ranking between two constraints has to happen in two steps, not one.

As an example of this, Cho (1998) examines *r*-deletion in English, including linking-*r* and intrusive-*r*. Linking-*r* is the preservation of /r/ between vowels in certain dialects of English where coda /r/ is normally deleted; for example, the /r/ in "Homer arrived" is preserved in dialects with linking /r/ even though the same /r/ is deleted in "Homer left". Intrusive-*r*. is the insertion of a non-underlying /r/ between a short low vowel and a following vowel in a subset of the dialects where linking-*r* occurs; for example, an /r/ is inserted between "Wanda" and "arrived" in "Wanda arrived" in dialects with intrusive-*r*.. For Cho's (1998) analysis, there are three constraints, as defined in (1).

(1)   CODA CONSTRAINT – prohibits the occurrence of /r/ in a syllable coda
       FINAL-C – prohibits a short vowel at the end of a prosodic word
       FAITHFULNESS – the output should match the input

Cho (1998) tracks five stages in the development of different dialects of English. The first stage is the most conservative, and corresponds to

---

[3] Also see John Ohala's listener-based theory of sound change (Ohala 1981, 2012).

most North American dialects, with no deletion or insertion of /r/. The second stage is a transition between the first and third stage, with optional deletion of coda /r/, and hence consistent use of linking-*r*. The third stage is attested in formal Received Pronunciation (RP), with consistent deletion of coda /r/ and use of linking-*r*. The fourth stage is a transitional stage between the third and fifth stages, which is like stage 3, but also has optional use of intrusive-*r*, and is attested in parts of England. Stage 5, which is also attested in many parts of England, has consistent /r/ deletion and use of linking-*r*, as well as consistent insertion and use of intrusive-*r*.

Stage 1 is characterized by FAITHFULNESS outranking the CODA CONSTRAINT, which in turn outranks FINAL-C. In this stage, since FAITHFULNESS is ranked most highly, any /r/ in the input is preserved in the output, and no insertion or deletion can occur. The second stage is characterized by the 'undoing' of the ranking between FAITHFULNESS and the CODA CONSTRAINT. This means that the speaker has the option of pronouncing the /r/ or deleting the /r/ in the coda of forms like "Homer left". Stage 3 makes deletion mandatory in this environment by creating a new ranking in which CODA CONSTRAINT now outranks FAITHFULNESS (which in turn outranks FINAL-C). In stage 4 the ranking between FAITHFULNESS and FINAL-C is undone. Because of this, there is no clear winner between "Wanda arrived" without *r*-insertion and with *r*-insertion, leading to optional insertion. Stage 5 makes insertion mandatory by creating a new ranking in which FINAL-C outranks FAITHFULNESS. Throughout these stages, the FAITHFULNESS constraint 'works its way down' the constraint ranking. Stages 2 and 4 demonstrate variation with constraints unranked relative to each other, and also serve as intermediate stages historically. Each subsequent stage corresponds to the level of innovation in a given dialect.

Bermúdez-Otero's (2006) treatment of intrusive-*r* is that it is a case of reanalysis. He argues that the status of /r/ as an epenthetic consonant is not predicated by markedness, since /r/ is extremely marked and epenthetic consonants should be unmarked. Instead, the phonological rule that resulted in linking-*r* was inverted to create the intrusive-*r* rule. Rule inversion is argued to be a type of reanalysis, and reanalysis is not subject to markedness laws "because it is not driven by phonetics, but by cognitive principles governing the relationship between phonology, morphology, and the lexicon" (Bermúdez-Otero 2006, 499; also see Vennemann 1972).

The present study will analyze how word-internal super-heavy syllables are dealt with in Arabic dialects using an approach similar to that of Cho (1998).

3   Methodology

This study makes use of data elicited from six speakers, each speaking a variety of one of the following six Arabic dialects: Saudi,[4] Egyptian, Emirati, Kuwaiti, Algerian, and Palestinian. Data were presented to the participants in the form of written English words, and each participant was asked to pronounce each word in their native dialect of Arabic. The forms were presented in English and not in Arabic orthography, in order to avoid an orthographic effect, where the speakers might use a more Modern Standard Arabic-like pronunciation. The reason for such an effect is that all the participants are highly educated in Arabic, and so know Modern Standard Arabic (MSA) well; also, since modern Arabic dialects are rarely written, written Arabic is assumed to be a representation of MSA. Lastly, the academic and formal nature of the elicitation procedure could induce the participants to automatically produce MSA.

The pronunciations were recorded using a digital audio recorder, and transcribed carefully. These transcriptions constitute the primary body of data used in this study. Before starting each recording session, a test session was conducted with each participant to make sure that they understood the task. In the case of the Algerian speaker, several stems that were problematic in her dialect were identified during this session, and replaced with more appropriate stems with the same target syllable structures.

The stimuli were composed to target specific phonological structures. Three noun stems representing different types of potentially super-heavy syllable were chosen (CVVC, CVCC, and CVCː), as well as two stems which end with regular heavy syllables, which can never be super-heavy (CVCVC and CVCVV). The words used for all of the dialects except Algerian were /naːr/ 'fire', /bint/ 'daughter, girl', /ʔumː/ 'mother', /dʒamal/ 'camel', /samaː/ 'sky'. These are all very well entrenched parts of the basic lexicon of each dialect, and so should be resistant to borrowing, and hence have cognates in each dialect (Zenner et al. 2014). All of these words had cognates in all of the dialects, except for /ʔumː/ in Algerian Arabic, which was replaced with /dʒɨdː/ 'grandfather', which has the same prosodic structure and can also be considered to be well entrenched.

Three forms of each word were elicited: the base form of the noun, with no additional morphology; the first person plural possessive form, with the consonant-initial suffix /-na/ 'our'; and the first person singular

---

[4] The specific variety of Saudi Arabic that was investigated is Hijazi Arabic. This is the term used throughout the rest of this paper.

possessive form, with the vowel-initial suffix /-i/ 'my'. In Algerian Arabic, there is a different set of possessive suffixes for inanimate nouns, so the inanimate noun /naːr/ 'fire' was replaced with /χɛːl/ 'maternal uncle'. No animate nouns were found at the time of original elicitation to replace the stems with non-potentially super-heavy syllables, so the possessive forms of these stems were elicited with the knowledge that they may end up being problematic.

3.1   Participants

All participants are at least bilingual — i.e., they know their native dialect of Arabic and are L2 speakers of English. They are also all between the ages of 20 and 35. Below, each participant is referred to by the name of their dialect, and some further information is given about the specific dialect they speak and their educational background.

- Hijazi   speaker of the Meccan dialect, male linguistics
               graduate student

- Egyptian   speaker of the Cairene dialect, female linguistics
               graduate student

- Emirati   speaker of the standard Emirati dialect from Abu Dhabi,
               male undergraduate in a non-linguistics field

- Kuwaiti   speaker of the Bedouin Kuwaiti dialect, male graduate
               student in non-linguistics field

- Algerian   speaker of the Western Algerian dialect, female
               graduate student in linguistics

- Palestinian speaker of the Central rural Palestinian dialect, male
               graduate student in non-linguistics field

While only one speaker was consulted for each dialect, each dialect is assumed to be uniform in regard to how it deals with potential super-heavy syllables. Despite the gender distribution, the range of educational backgrounds, differences in experience with linguistics, level of knowledge of MSA, and knowledge of other languages, only the variable of dialect is relevant to this study. These other variables may influence the speakers' production of potential super-heavy syllables, but since they are not the focus of this study, they will not be addressed further.

The Hijazi, Emirati, and Kuwaiti speakers all represent Arabian dialects of modern Arabic, the Egyptian speaker is the only represent-tative of Egyptian Arabic, the Rural Palestinian speaker is the only representative of Levantine Arabic, and the Algerian speaker is the only

representative of Maghrebi Arabic. This covers four of the five modern dialect areas described by Bassiouney (2009).

4   Data

The data for Hijazi Arabic, along with a basic moraic analysis, is presented in (2).

(2)   Hijazi

gloss	**UR**	**SR**	**/-naː/ 'our'**	**/-i/ 'my'**

'fire'   /naːr/   [naːr]   ['naː.ra.na]   ['naː.ri]

'daughter' /bint/   [bɪnt]   ['bin.ta.na]   ['bin.ti]

'mother'   /ʔumː/   [ʔumm]   ['ʔum.ma.na]   ['ʔum.mi]

'camel'   /ʤamal/   [ʤa.mal]   [ʤa.'mal.na]   ['ʤa.ma.li]

'sky'   /samaː/   ['sa.ma]   [sa.'maː.na]   [sa.'maːj]

All three stems with potentially super-heavy syllables (when a consonant-initial suffix is added), undergo epenthesis. This allows the final consonant in each stem to resyllabify as the onset to the syllable whose nucleus is the epenthetic vowel, avoiding any potentially super-heavy syllables. In addition, the long vowel in /samaː/ shortens word-finally.

The data for Egyptian Arabic, along with a basic moraic analysis, is presented in (3).

(3)   Egyptian

gloss	UR	SR	/-naː/ 'our'	/-i/ 'my'

'fire'   /naːr/

```
           μ μ        μ μ   μ        μ μ μ
           |/         | |   |        |/  |
'fire'   /naːr/  [naːr]  ['nar.na]   ['naː.ri]
```

```
           μμ         μμ  μ   μ      μμ  μ
           | |        | |  |   |     | |  |
'daughter'/bint/  [bint]  [bin.'ti.na] ['bin.ti]
```

```
           μ μ        μμ    μ   μ     μμ    μ
           | |        | |   |   |     | |   |
'mother'  /ʔumː/ [ʔumm] [ʔum.'mi.na] ['ʔum.mi]
```

```
           μ   μ      μ   μμ  μ     μ    μ  μ
           |   |      |   | |  |    |    |  |
'camel'   /gamal/['ga.mal] [ga.'mal.na]['ga.ma.li]
```

```
           μ   μ   μ      μμ  μ     μ    μμ μ
           |   |   |      |/  |     |    |/  |
'sky'    /samaː/['sa.ma] [sa.'mæː.na] [sa.'mæː.ja]
```

The three stems with potentially super-heavy syllables exhibit differing behaviors. The vowel in the stem /naːr/ shortens when a consonant-initial suffix is added, i.e., in the form ['nar.na], avoiding a potential super-heavy syllable. In forms of the stems /bint/ and /ʔumː/ with a consonant initial suffix, a vowel is epenthesized after the last consonant of the stem, allowing it to resyllabify, avoiding any potential super-heavy syllables. It is noteworthy that the epenthetic vowel is stressed in the Egyptian data, unlike in other dialects where epenthesis occurs, though this study will not explore the differing stress systems of Arabic dialects. Like in most of the other dialects, the long vowel in /sama:/ is shortened word-finally.

The data for Emirati Arabic, along with a basic moraic analysis, is presented in (4).

(4)   Emirati

gloss	**UR**	**SR**	**/-naː/ 'our'**	**/-i/ 'my'**

```
                        μ μ         μ μ   μ        μ μ  μ
                        |/          |/    |        |/   |
'fire'      /naːr/     [nɒːr]      ['nɒːr.na]     ['nɒː.ri]

                        μ μ         μ μ  μ  μ       μ μ  μ
                        | |         | |  |  |       | |  |
'daughter' /bint/      [bɯnt]      ['bɯn.tɨ.nə]   ['bɯn.ti]

                        μ μ         μ μ   μ   μ      μ μ    μ
                        | |         | |   |   |      | |    |
'mother'   /ʔumː/      [ʔʊmm]      ['ʔʊm.mɨ.nə]   ['ʔʊm.mi]

                        μ    μ       μ    μμ  μ      μ    μ   μ
                        |    |       |    ||  |      |    |   |
'camel'  /ʤamal/      ['ʤɛ.mɛl]    [ʤɛ.'mɛl.nə]  ['ʤɛ.mɛ.li]

                        μ   μ    μ    μμ  μ        μ    μμ
                        |   |    |    |/  |        |    |/
'sky'    /samaː/     ['sɪ.mɒ]   [sɪ.'mɒː.nə]   [sɪ'mɒːj]
```

Like the Egyptian data, the treatment of potential super-heavy syllables differs depending on the prosodic shape. The form of /naːr/ with a consonant-initial suffix does not undergo epenthesis, but the forms of /bint/ and /ʔumː/ do undergo epenthesis with a consonant-initial suffix.

The data for Kuwaiti Arabic, along with a basic moraic analysis, is presented in (5).

(5) Kuwaiti

gloss	UR	SR	/-naː/ 'our'	/-i/ 'my'

```
                          μμ          μμ   μ        μμ  μ
                          |/          |/   |        |/  |
'fire'       /naːr/     [nɑːr]    ['nɑːr.na]    ['nɑː.ri]

                          μμ          μμ   μ        μμ  μ
                          ||          ||   |        ||  |
'daughter' /bint/       [bɨnt]    ['bɨnt.na]    ['bɨn.ti]

                          μ μ         μ μ   μ       μ μ    μ
                          | |         | |   |       | |    |
'mother'   /ʔumː/      [ʔʊmm]    ['ʔʊmm.na]   ['ʔʊm.mi]

               μ    μ        μ   μμ  μ         μ    μ μ
               |    |        |   ||  |         |    | |
'camel' /dʒamal/ ['dʒɪ.mɛl] [dʒɪ.'mɛl.na]  ['dʒɪ.mɛ.li]

               μ   μ    μ     μμ  μ       μ   μμ
               |   |    |     |/  |       |   |/
'sky'     /samaː/['sɪ.ma] [sɪ.'mɑː.na] [sɪ'mɑːj]
```

In Kuwaiti Arabic, there is no epenthesis or vowel shortening, even in forms with consonant-initial suffixes, which create potential super-heavy syllables, including situations where there are two consonants in the coda, and in situations where a long vowel precedes a coda consonant. In the form [dʒɪ.'mɛl.na], the coda consonant receives a mora, but in ['nɑːr.na], the coda consonant does not receive a mora, since this would create a trimoraic syllable.

The data for Rural Palestinian Arabic, along with a basic moraic analysis, is given in (6).

(6)   Rural Palestinian

gloss	**UR**	**SR**	**/-naː/ 'our'**	**-/i/ 'my'**
		μμ |/	μ μ μ |/ |	μμ μ |/ |
'fire'	/naːr/	[nɔːr]	[ˈnɔːr.næ]	[ˈnɔː.ri]
		μμ ||	μμ μ || |	μμ μ || |
'daughter'	/bint/	[bɪnt]	[ˈbɪnt.næ]	[ˈbɪn.ti]
		μμ ||	μμ μ || |	μμ μ || |
'mother'	/ʔɪmː/	[ʔɪmm]	[ˈʔɪmm.næ]	[ˈʔɪm.mi]
		μ μ | |	μ μμ μ | || |	μ μ μ | | |
'camel'	/dʒamal/	[ˈdʒɛ.mɛl]	[dʒɛ.ˈmɛl.næ]	[ˈdʒɛ.mɛ.li]
		μ μ | |	μ μμ μ | |/ |	μ μμ | |/
'sky'	/samaː/	[ˈsɛ.mæ]	[sɪ.ˈmæː.næ]	[sɛ.ˈmæːj]

This Rural Palestinian dialect of Arabic exhibits the same patterns as Kuwaiti Arabic in terms of faithfulness in the presence of a consonant-initial suffix.

The data for Algerian Arabic, along with a basic moraic analysis, is presented in (7).

(7) Algerian

gloss	UR	SR	/-na:/ 'our'	/-i/ 'my'
		μ μ |/ 	μ μ μ |/ | 	μ μ μ |/ |
'maternal uncle'	/χɐ:l/	[χɐ:l]	['χɐ:l.nɑ]	['χɐ:li]
		μ μ | | 	μ μ μ | | | 	μ μ μ | | |
'daughter'	/bint/	[bɨnt]	['bɨnt.nɑ]	['bɨn.ti]
		μ μ | | 	μ μ μ | | | 	μ μ μ | | |
'grand-father'	/dʒɨd:/	[dʒɨdd]	['dʒɨdd.nɑ]	['dʒɨd.di]
		μ | 	μ μ μ μ μ | | | | | 	μ μ μ μ | | | |
'camel'	/tʃmɨl/	[tʃmɨl]	[tʃmɨl 'tɛʕ.nɑ]	[tʃmɨl 'tɛ.ʕi]
		μ μ |/ 	μ μ μ μ | | | | 	μ μ μ | | |
'sky'	/sma:/	[smɑ:]	[smɛ 'tɛʕ.nɑ]	[smɛ 'tɛ.ʕi]

The Algerian data are similar to the Palestinian and Kuwaiti data in terms of faithfulness in the presence of consonant-initial suffixes. Since the word 'camel' is considered inanimate in Algerian Arabic (and so does not take the same possessive morpheme as the other stems), and is furthermore mono-syllabic (so it doesn't show a stress shift like in other Arabic dialects), the word [mu.'ʕal.lim] 'teacher' was examined in addition. The first-person plural possessive form of this word is [mu.ʕal.'lim.na], showing both that there is no epenthesis, and that stress shifts in order to remain on the penultimate syllable as in other Arabic dialects.

The transcriptions made for this research include close transcriptions of the vowels. It is worth noting that the quality of epenthetic vowels differs in the three dialects with epenthesis: [a] in Hijazi, [i] in Egyptian, and [ɨ] in Emirati. The question of what vowel is epenthesized is not the focus of this study, and so will remain unaddressed. This topic is an interesting one to consider in future research, but will not be discussed further in this paper (see Broselow 1976, Abu-Mansour 1991, Farwaneh 2009).

As this paper is about historical phonology, use of Classical Arabic or MSA data might be considered ideal for determining what the original treatment of potential super-heavy syllables may have been. However, underlying word-internal super-heavy syllables are at best rare in these

two varieties. Forms like /naːr-na/ did not exist, and instead a case morpheme always needed to intervene between the noun and the possessive suffix, e.g. /naːr-u-na/ 'fire-NOM-our'. This allowed a potentially super-heavy syllable to resyllabify, surfacing as e.g. [ˈnaː.ru.na].[5] Because no faithful forms of this sort would ever produce a super-heavy syllable in Classical Arabic, there are no data on which to argue for the relative rankings of various constraints. Basically, the classical data are not comparable to the modern dialectal data. However, this fact suggests that the loss of case morphology historically may have given rise to the very set of problems that this paper investigates in the modern Arabic dialects. Gruber-Miller (1990) presents several potential reasons why essentially all modern dialects of Arabic have lost case morphology, including final short vowel loss, and stress shift. However, these phonological reasons do not explain why the case suffixes were not preserved before possessive suffixes. Either way, their loss may have motivated the need in some dialects for an epenthetic vowel before consonant-initial suffixes, at least in situations following different types of potential super-heavy syllables.

5   Analysis

This study utilizes the optimality-theoretic framework in order to explain the change from a certain Arabic dialect to another using constraints that will be defined in the following subsection.

5.1   The Constraints

The analysis of the various dialects to come below makes use of the set of universal constraints given in (8).

---

[5] Similar to [ˈnaː.ru.na], super-heavy syllables do not occur in Classical Arabic/MSA in the verb paradigm. For example, the word /baːʕ-a-ha/ becomes [ˈbaː.ʕa.ha] 'he sold it (fem.).' The perfect tense marker /a/ added to the verb in the surface form forces resyllabification of the last consonant of the verb root, forming a new light syllable. This example is cognate with Hijazi [ˈbaː.ʕa.ha] in which [a] is not a tense marker, but rather it is an epenthetic vowel to prevent a word-internal super-heavy syllable from surfacing.

(8) OT Constraints

   a) Max-μ_V – Do not delete a vocalic mora. Assign a constraint violation for any vocalic mora in the input that is not realized in the output (Davis 2017, 247).

   b) Max-C_{slot} – Do not delete a consonant slot. Assign a constraint violation for any consonant slot in the input that does not have a correspondent in the output.

   c) Weight-By-Position (WbP) – A coda consonant must be moraic. Assign a constraint violation for any coda consonant that is not moraic (Kager 1999, 147).

   d) *CCC – No three consonants in a row. Assign a violation when three consonants appear in a row. This phonological condition was originally proposed by Kisseberth (1970).

   e) Dep-μ_C – Do not insert a consonantal mora.[6] Assign a constraint violation for any consonantal mora in the output that is not present in the input.

   f) *$\sigma_{\mu\mu\mu}$ – No trimoraic syllables.[7] Assign a violation for any syllable that has three moras.

   g) Dep-V – Do not insert a vowel. Assign a constraint violation for any vowel in the output that does not have a correspondent in the input.

5.2 The individual grammars

The ranking of the constraints determines the grammar of each dialect. The grammar of each dialect is presented below as a Hasse diagram, where each line represents the relative ranking of the constraints it connects. Each constraint that is connected by a line to a constraint below it is understood to outrank (or dominate) that constraint. Because of the property of strict dominance, it also outranks the constraints that are dominated by the constraints that it outranks. For example, if constraint A outranks constraint B, and constraint B outranks constraint C, then constraint A outranks constraint C by transitivity. In order to highlight the differences between dialects, I begin the discussion of each dialect with the corresponding Hasse diagram. For each individual ranking presented in the Hasse diagrams, a corresponding ranking argument, in the form of an argument tableau, can be constructed. Because of the large number of constraint dominations in this study, only

---

[6] Bamakhramah (2009) calls this constraint Dep-μ-C.
[7] Bamakhramah (2009) calls this constraint *3μ.

the most crucial comparative tableaux are presented. The Hasse diagram for Hijazi Arabic is presented in (9).

(9)   Hijazi

$$\text{*CCC} \qquad \text{MAX-}\mu_V \qquad \text{MAX-C}_{slot} \qquad \text{WBP} \qquad \text{*}\sigma_{\mu\mu\mu}$$
$$\text{DEP-V}$$
$$\text{DEP-}\mu_C$$
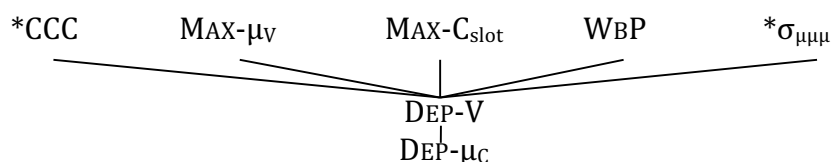
In Hijazi Arabic, to avoid a potential super-heavy syllable word-internally, DEP-V is violated, allowing a vowel to be epenthesized, instead of shortening a long vowel (i.e., a violation of MAX-µV, as happens in Egyptian Arabic). Furthermore, WBP is satisfied in Hijazi Arabic, meaning that ['naː.ra.na] surfaces for /naːr-na/ instead of the faithful form ['naːr.na] (the Emirati, Kuwaiti, Algerian, and Palestinian form), as shown in tableau 1.

µ µ µ |/ | /naːr-na/	WBP	MAX-µ$_V$	*σ$_{\mu\mu\mu}$	DEP-V
µ µ µ |/ | a. ['naːr.na]	* !			
µ µ µ µ |/ | | ☞ b. ['naː.ra.na]				*
µµ µ || | c. ['nar.na]		* !		
µµµ µ |/| | d. ['naːr.na]			*!	

Tableau 1: A vowel is epenthesized to avoid a violation of WBP and MAX-µ$_V$.

If DEP-V were to dominate either WBP or MAX-µ$_V$ then either the candidate with a deleted vocalic mora or the fully faithful candidate would wrongly surface, as in Egyptian or Emirati and the conservative dialects, respectively. In addition, if DEP-V were to dominate *σ$_{\mu\mu\mu}$ then ['naːr.na] with a trimoraic initial syllable would wrongly surface.

In Hijazi Arabic, as in Egyptian and Emirati Arabic, it is preferable to epenthesize than to allow a tri-moraic syllable to surface. This is due to the fact that $*\sigma_{\mu\mu\mu}$ and WBP dominate DEP-V, which results in ['bin.ta.na] in Hijazi for /bint-na/ instead of *['bint.na], as shown in tableau 2.

μ μ | | /bint-na/	$*\sigma_{\mu\mu\mu}$	WBP	DEP-V
μμ μ μ || | | ☞ a. ['bin.ta.na]			*
μμ μ || | b. ['bint.na]		*!	
μμμ μ ||| | c. ['bint.na]	*!		

Tableau 2: Epenthesis occurs to avoid having a trimoraic (super-heavy) syllable.

If DEP-V were to dominate $*\sigma_{\mu\mu\mu}$ or WBP, then the candidate with a tri-moraic syllable or the fully faithful candidate with a rhyme slot that is not associated with a mora would wrongly surface, respectively. While the form ['bint.na] surfaces in the conservative varieties, the /t/ is not associated with an epenthetic mora in these varieties either. This will be discussed in more detail later. Also, *CCC prevents faithful forms in Hijazi Arabic, but as the tableau shows, these forms would be eliminated even if *CCC were not ranked highly.

In Hijazi Arabic, as in all the other dialects examined in this study, potential heavy syllables, such as in the second syllable of /ʤamal-na/ are also affected by these constraints. Since potential heavy syllables only have two timing slots in the rhyme (at least one of which will always be a vowel slot), unlike potential super-heavy syllables, they will never violate constraints like $*\sigma_{\mu\mu\mu}$ or *CCC. In turn, a consonantal mora can be epenthesized at the expense of satisfying WBP without violating other high-ranked constraints.[8] In Arabic, generally speaking, the presence of

---

[8] Satisfying WBP implies violating DEP-μ_C. This is because a coda consonant is underlyingly not moraic and WBP motivates mora assignment to the syllable coda in the output form. This makes mora association to the syllable coda different than the long vowel, which is inherently associated with two moras.

two moras in a penultimate syllable attracts stress; otherwise, stress would fall on the antepenult.[9] Because the antepenult is not stressed in forms like this in any of the dialects examined, it is clear that a mora is being epenthesized in all of the dialects. Another possible repair strategy to satisfy WBP would be vowel epenthesis, but this is not encountered in any of the dialects examined. Because consonantal mora epenthesis is the winning strategy, DEP-μ_C must be low-ranked in all dialects. The result is that [dʒa.'mal.na] surfaces with the epenthesis of a mora, as opposed to *[dʒa.'ma.la.na] with epenthesis of a vowel, or *['dʒa.mal.na] with no epenthesis, as shown in tableau 3.

If DEP-μ_C were to dominate DEP-V, then the candidate with an epenthesized vowel would wrongly surface. Likewise, if WBP were to be dominated by DEP-μ_C, then the fully faithful form, with no mora epenthesis, would wrongly surface.[10]

μ μ μ | | | /dʒamal-na/	WBP	DEP-V	DEP-μ_C
μ μμ μ | || | ☞ a. [dʒa.'mal.na]			*
μ μ μ μ | | | | b. [dʒa.'ma.la.na]		* !	
μ μ μ | | | c. ['dʒa.mal.na]	* !		

Tableau 3: Vowel epenthesis does not occur between two consonants which can be syllabified as a singleton onset and a singleton coda, and instead a consonantal mora is epenthesized.

Now that we have considered Hijazi Arabic we turn next to the second dialect, Egyptian Arabic. The Hasse diagram for Egyptian Arabic is presented in (10).

---

[9] In some southern Arabian dialects, stress falls on a syllable that contains a geminate coda even when the following syllable contains two moras as in ['dar.ra.sat.ni] 'she taught me' (Alahmari 2018).

[10] A potential candidate such as ['dʒaᵘ.maᵘˡᵘ.naᵘ] would be eliminated by the markedness constraint WSP (Kager 1999, 155), which is assumed to be highly ranked in these dialects.

(10) Egyptian

*CCC   MAX-C$_{slot}$ WBP   *σ$_{\mu\mu\mu}$

DEP-V

DEP-μ$_C$ MAX-μ$_V$
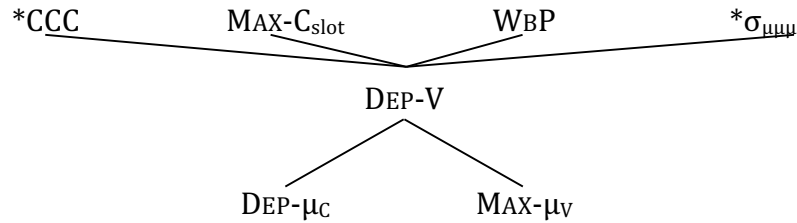
The ranking of the constraints for Egyptian Arabic is pretty similar to Hijazi Arabic, except for the relation between DEP-V and MAX-μ$_V$. The main difference between the two dialects is that in Hijazi Arabic, MAX-μ$_V$ dominates DEP-V, but in Egyptian Arabic, DEP-V dominates MAX-μ$_V$. This difference results in /naːr-na/ surfacing as ['nar.na] with vowel shortening and no epenthesis, as opposed to ['naː.ra.na] with epenthesis and no vowel shortening. The winning candidate in Egyptian Arabic satisfies WBP at the expense of violating the low-ranked DEP-μ$_c$. This is shown in tableau 4.

μ μ μ |/ | /naːr-na/	WBP	DEP-V	MAX-μ$_V$	DEP-μ$_c$
μ μ μ |/ | a. ['naːr.na]	*!			
μμ μ || | ☞ b. ['nar.na]			*	*
μμ μ μ |/ | | c. ['naː.ra.na]		*!		
μ μ | | d. ['nar.na]	*!		*	

Tableau 4: A vocalic mora is deleted to avoid a violation of WBP and DEP-V.

If MAX-μ$_V$ were to dominate either DEP-V or WBP, then either the candidate with epenthesis or the fully faithful candidate would wrongly surface. If DEP-μ$_c$ outranked WBP then the winning candidate ['nar.na] with no mora in the coda [r] would wrongly surface. Other forms are the

same between Egyptian and Hijazi. For example, both display vowel insertion in /bint-na/, indicating a low-ranked DEP-V.

Next, we consider Emirati Arabic. The Hasse diagram for Emirati Arabic is presented in (11).

(11) Emirati
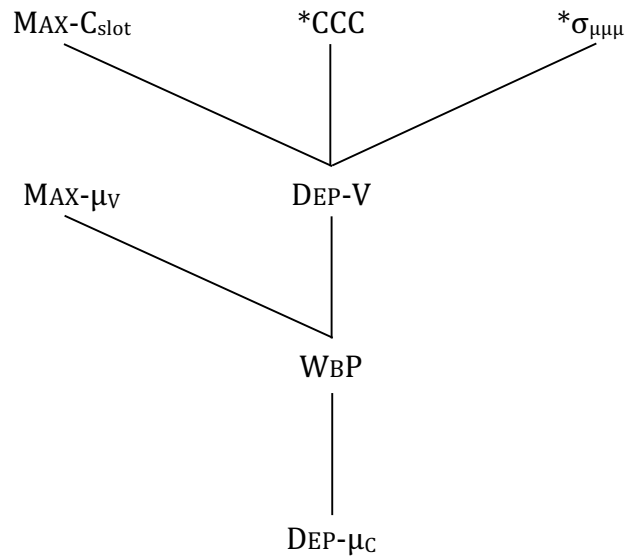
The ranking of the constraints for Emirati Arabic is somewhat similar to Hijazi Arabic, except for the relations between DEP-V and WBP, MAX-μ_V and WBP, and DEP-μ_C and WBP. In Hijazi Arabic, MAX-μ_V and WBP dominate DEP-V, but in Emirati Arabic, DEP-V and MAX-μ_V dominate WBP. This difference results in /naːr-na/ surfacing fully faithfully as ['naːr.na], and not with vowel shortening as in Egyptian Arabic, or with epenthesis as in Hijazi Arabic.

μ μ μ |/ | /naːr-na/	MAX-μv	DEP-V	WBP
μ μ μ |/ | ☞ a. [ˈnaːr.na]			*
μμ μ || | b. [ˈnar.na]	* !		
μμ μ μ |/ | | c. [ˈnaː.ra.na]		* !	

Tableau 5: Since WBP is ranked low, the fully faithful candidate is allowed to surface, and neither epenthesis nor vocalic mora deletion occur.

If WBP were to dominate either DEP-V or MAX-μv, then either the candidate with epenthesis or vocalic mora deletion would wrongly surface, as in Hijazi or Egyptian Arabic, respectively.

In Emirati, like in Hijazi and Egyptian Arabic, /ʤamal-na/ surfaces with an epenthesized consonantal mora, as [ʤaˈmal.na], even though DEP-V and WBP are ranked differently with respect to one another. The important factor is that they both dominate DEP-μC.

The main difference between Emirati and all other dialects is the role of *CCC.[11] Because WBP is ranked below DEP-V in Emirati, it cannot be the trigger for epenthesis in [ˈbin.ta.na]. Instead, *CCC prevents the fully faithful form *[ˈbint.na] from surfacing, as shown in tableau 6.

---

[11] In Emirati, the constraint *CCC is crucial in the analysis as it motivates vowel epenthesis in a sequence of 3 consecutive medial consonants. The constraint *μμμ is also undominated in this dialect as well as other Arabic dialects.

μ μ | | /bint-na/	*CCC	DEP-V	WBP
μμ μ ‖ | a. [bint.na]	*!		*
μμ μ μ ‖ | | ☞ b. [bin.ta.na]		*	

Tableau 6: *CCC prevents /bint-na/ from surfacing faithfully.

Having presented the ranking for the Emirati dialect, we now consider the more conservative dialects. The Hasse diagram for the conservative dialects (Algerian, Kuwaiti, Palestinian) is presented in (12).

(12) Algerian, Kuwaiti, Palestinian

MAX-C$_{slot}$ DEP-V   *σ$_{μμμ}$ MAX-μ$_V$

*CCC   WBP

DEP-μ$_C$
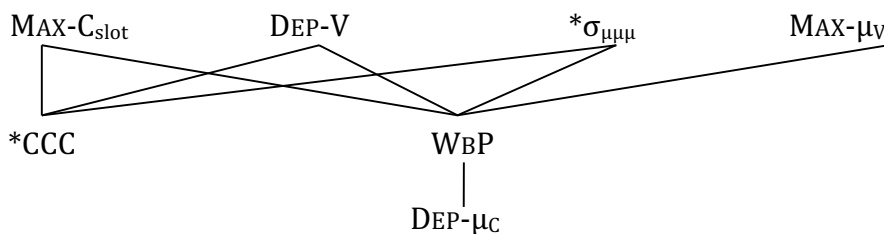
In the conservative dialects, to avoid a potential super-heavy syllable word-internally, no mora is epenthesized, as in Emirati Arabic. Tableau 7 shows that *σ$_{μμμ}$, MAX-μ$_V$, and DEP-V outrank WBP.

μ μ μ |/ | /naːr-na/	*σ_{μμμ}	MAX-μ_V	DEP-V	WBP
μ μ μ |/ | ☞ a. [ˈnaːr.na]				*
μμμ μ |/| | b. [ˈnaːr.na]	*!			
μ μ μ | | | c. [ˈnar.na]		*!		
μ μ μ μ |/ | | d. [ˈnaː.ra.na]			*!	

Tableau 7: Since WBP is ranked low, the fully faithful candidate is allowed to surface, and epenthesis of a vowel, epenthesis of a mora, and vocalic mora deletion are not allowed to occur.

If WBP were to dominate *σ_{μμμ}, MAX-μ_V, or DEP-V, then /naːr-na/ would wrongly surface with an epenthetic consonantal mora, a deleted vocalic mora, or an epenthesized vowel, respectively. Tableau 8 also shows that *σ_{μμμ} and DEP-V outrank WBP as related to /bint-na/ as well as showing that *CCC and WBP outrank DEP-μ_C, resulting in the winning candidate in which two consonants serve in the coda with only one epenthesized mora.

In the following tableau, if *CCC were to dominate DEP-V, then /bint-na/ would wrongly surface with an epenthetic vowel. If DEP-μ_C outranked WBP then the faithful candidate, which does not assign moras to coda consonants would wrongly surface. If WBP outranked *σ_{μμμ} then the candidate that assigns two moras to the coda consonants would wrongly surface.

μ μ | | /bint-na/	*σ$_{μμμ}$	Dᴇᴘ-V	WʙP	Dᴇᴘ-μ$_C$
μ μ | | a. [bint.na]			**!	
μμμ μ ||| | b. [bint.na]	* !			**
μμ μ μ || | | c. [bin.ta.na]		*!		*
μμ μ || | ☞ d. [bint.na]			*	*

Tableau 8: Since Dᴇᴘ-μ$_C$ is the lowest ranked and *σ$_{μμμ}$ is a highly ranked constraint, the winning candidate surfaces with one epenthesized mora, the faithful candidate loses by two violations of WʙP

The rankings above show the OT accounts of the differences and similarities between these dialects. The following section considers the historical steps by which these dialects could have diverged.

5.3   Historical analysis

The process of changing one ranking to another is used to analyze the change from one grammar to another. According to Cho (1998), the only operations that may change a grammar are the unranking of two constraints that were previously ranked relative to one another, or the ranking of two constraints that were not previously ranked relative to one another in an immediate domination relationship. One assumption that is made is that the dialects with no epenthesis or shortening (Algerian, Kuwaiti, and Palestinian) are more conservative than the dialects with these processes. It is also assumed that the fewer the number of operations to get from one grammar to another, the more closely related the grammars are. The number of operations from each grammar to each of the others is presented in the table in (13).

(13) The number of steps between the grammars of each dialect[12]

	Hijazi	Egyptian	Emirati	Others
Hijazi	—	2	4	8
Egyptian	—	—	6	10
Emirati	—	—	—	4
Others	—	—	—	—

The number of operations from one grammar to another that is presented in the table in (13) simply represents a hypothetical shortest path from each grammar to another, and not (yet) the proposed set of historical changes.

Wichmann et al. (2010) provide a quantifiable method for determining the geographic homeland of a group of related languages. They take into account a simple linguistic difference metric and the geographical distance between the languages. Based on the assumption that the area with the most linguistic difference in the smallest area is the homeland, they then compute which of the languages corresponds most closely to the location of the homeland. The linguistic difference metric of Wichmann et al. (2010) could be replaced with the OT reranking metric presented here, and originating in Cho (1998). Using this to calculate the homeland of Arabic dialects would likely result in an area around Hijaz as the homeland of Arabic dialects, since Hijazi, Egyptian, and Emirati dialects form a cluster of geographically close, but linguistically diverse dialects. This result is compatible with historical facts regarding the spread of Arabic. This is also compatible with Holes' (2004) description of modern Arabic dialects as forming a dialect continuum, with Hijazi being similar to its direct neighbors, Egyptian and Emirati. This would be an interesting study to pursue in future research.
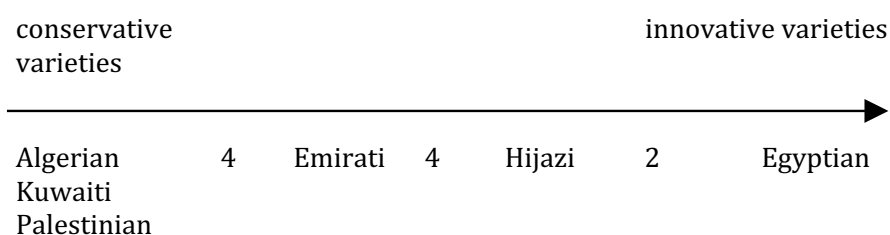
The dialect with the least number of changes from the conservative group of dialects is Emirati. The grammar represented by Emirati can thus be considered to represent an intermediate stage between the conservative varieties and the other varieties. Emirati, Hijazi, and Egyptian all went through a stage similar to the modern grammar of Emirati, and Hijazi and Egyptian diverged from it, while Emirati maintained its constraint ranking. Egyptian and Hijazi later diverged when Egyptian underwent a new innovation. Egyptian, then, represents

---

[12] Here, a 'step' is an operation according to Cho (1998), so that a change from two unranked constraints to one of ranking would be one step and a change from two ranked constraints to one of unranking would also constitute a step.

the most innovative dialect, and is the least conservative of all of these varieties. Egyptian can be assumed to have gone through all the previous stages, represented by the grammars of Emirati and Hijazi. This may be compatible with the life cycle of phonological processes as suggested in Bermúdez-Otero (2020).

This progression is represented in the diagram in (14), which includes the number of steps required to advance to the next grammar.

(14) The dialects based on degree of innovation

conservative   innovative varieties
varieties

$\longrightarrow$

Algerian   4   Emirati   4   Hijazi   2   Egyptian
Kuwaiti
Palestinian

The steps between each of the grammars are listed below. The symbols >> and << indicate that a domination relationship was established between the two constraints (the constraint on the open side came to dominate the constraint on the pointed side), and the symbol X indicates that the ranking relationship was destroyed.

(15) Steps from conservative grammar (12) to Emirati grammar (11)

First step   MAX-C$_{slot}$ >> DEP-V
Second step   *σ$_{\mu\mu\mu}$ >> DEP-V
Third step   *CCC X DEP-V
Fourth step   *CCC >> DEP-V

In the first step in (15), MAX-C$_{slot}$ dominates DEP-V. In the second step, the constraint against trimoraic syllables (*σ$_{\mu\mu\mu}$) dominates DEP-V. In the third step, the domination relationship between *CCC and DEP-V is destroyed (i.e. there is no domination relationship between these two constraints in this step). In the fourth step, *CCC dominates DEP-V.

(16) Steps from Emirati grammar (11) to Hijazi grammar (9)

First step   MAX-μ_V >> DEP-V
Second step   WBP X DEP-V
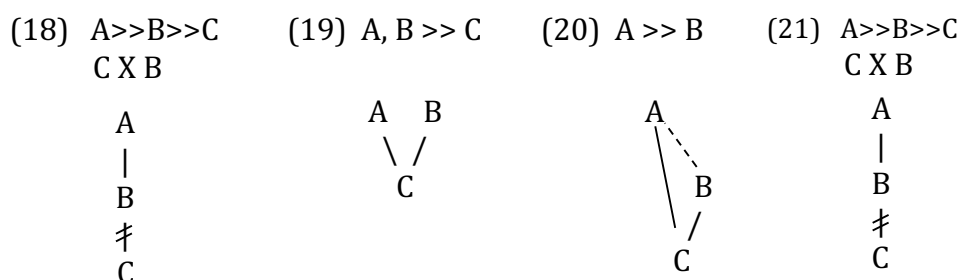Third step   DEP-V >> DEP- μ_c
Fourth step   WBP >> DEP-V

In the first step in (16), Max-μ_V dominates Dep-V. In the second step, the relationship between WbP and Dep-V is destroyed. In the third step, Dep-V dominates Dep-μ_c, and in the final step, WbP dominates Dep-V.

(17) Steps from Hijazi grammar (9) to Egyptian grammar (10)

      First step   Max-μ_V X Dep-V
      Second step   Max-μ_V << Dep-V

In the first step in (17), the domination relationship between the faithfulness constraints Max-μ_V and Dep-V is destroyed. In the second step, Max-μ_V is dominated by Dep-V.

It is assumed that when the ranking relationship between two constraints is destroyed, all ranking relationships that exist due to transitivity through that original relationship are also destroyed. For example, if A >> B >> C, then A is assumed to outrank C by transitivity; then, when the relationship between B and C is destroyed, as indicated by the 'X' symbol in (18), B no longer dominates C. Moreover, the constraint A, which originally dominated B, no longer dominates C through transitivity.

(18) A>>B>>C   (19) A, B >> C   (20) A >> B   (21) A>>B>>C
     C X B   C X B

      A   A   B   A   A
      | \ / |
      B   C   B   B
      ‡   /   ‡
      C   C   C

Another assumption relates to a starting configuration of A and B both outranking C, as in (19) If a single operation then makes A outrank B, as in (20), then the new transitivity relationship between A and C, through B, is assumed to replace the original direct relationship between A and C. The new Hasse diagram would be just like the starting configuration of (18), repeated in (21). Because of this, if C is then unranked with regard to B, as in (21), then it is also no longer ranked relative to A. In other words, if the domination relationship between the last two points in the diagram is destroyed, it implies that the relationship between the first and the last points in the diagram is also destroyed.

An interesting point is that consideration of just these forms does not shed light on the progression of changes that most likely took place. For example, there is only one difference in the forms that separate Hijazi, Emirati, and Egyptian dialects from one another: /naːr-na/ surfaces

faithfully as ['nɒːr.na] in Emirati, with epenthesis as ['naː.ra.na] in Hijazi, and with vowel shortening as ['nar.na] in Egyptian. Based on these data, Emirati appears to be the most conservative variety, and Hijazi and Egyptian both appear to be more innovative than Emirati, though it is unclear which would be more innovative than the other. In this sense, the OT analysis, using Cho's (1998) distance metric, provides both more and different insight into diachronic evolution than simply comparing the number of forms which are different.

The two main ways to represent shared innovation in related languages or dialects are tree diagrams and wave diagrams (Millar 2007, §7.4).

The tree diagram in (22) groups together Emirati, Hijazi, and Egyptian as a distinct group from the conservative dialects, and further groups Hijazi and Egyptian as more similar.

(22) A branching tree diagram based on the analysis of constraint reranking that shows which branch is most innovative
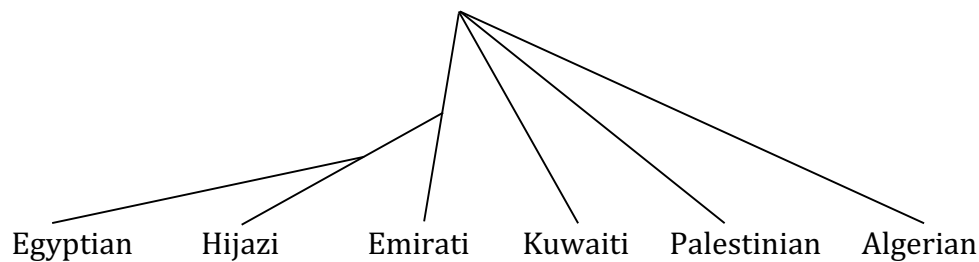

In (22), new innovation is represented by branching, while a continuing straight line shows that there was no new innovation in regards to potential super-heavy syllables. For example, Emirati, Hijazi, and Egyptian Arabic are all grouped together, but because the line they all diverge from continues straight to Emirati, it is understood that Emirati only has innovation that occurred before the three dialects first diverged. Likewise, Hijazi only has innovation that occurred before Hijazi and Egyptian diverged.

The main problem with this tree diagram is that it implies that there was some point when Egyptian, Hijazi, and Emirati Arabic were all the same dialect, to the exclusion of Kuwaiti. As Bassiouney (2009) points out, Emirati, Hijazi, and Kuwaiti Arabic are all Arabian dialects, while Egyptian constitutes its own separate group of Arabic dialects. Because of this, Egyptian should not be grouped with Emirati and Hijazi, and Emirati and Hijazi should be grouped with Kuwaiti Arabic. However, remember that Holes (2004, 3–4) points out that modern Arabic dialects form a complicated sort of dialect continuum. According to Millar (2007, §7.4), tree diagrams are notoriously inadequate at representing the relation-

ships between dialects, because they imply a homogenous ancestral language that underwent a sudden split. What is more likely to have happened is that each of these phonological innovations spread to a set of nearby dialects, but did not affect all closely related dialects. As Millar (2007, 231) points out, innovations can spread between related varieties after they have "begun to diverge quite strongly". This can be represented clearly in a wave diagram, such as the one below in (23), with the numbered circles indicating sets of innovation. The first set of innovations affected Emirati, Hijazi, and Egyptian Arabic (circle 1); the second set affected only Hijazi and Egyptian Arabic (circle 2); and the last set of innovations affected only Egyptian Arabic (circle 3).

(23) A wave diagram based on the analysis of constraint reranking that groups dialects based on shared innovation
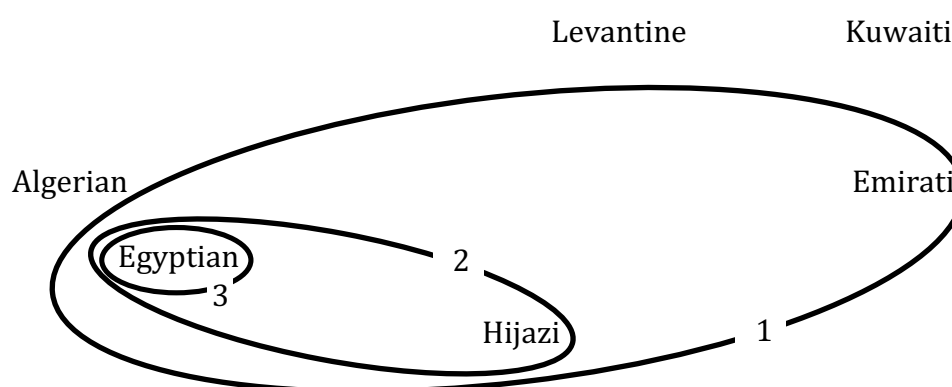
The line labelled 1 in (23) represents dialects that underwent the changes (OT operations) listed in (15) that resulted in epenthesis in forms like /bint-na/; the line labelled 2 represents dialects that underwent the changes listed in (16) that resulted in epenthesis in forms like /naːr-na/; and the line labelled 3 represents dialects that underwent the changes listed in (17) which resulted in vocalic mora deletion (i.e., vowel shortening) instead of epenthesis in forms like /naːr-na/.

Another advantage of this wave diagram is that it also shows that the conservative varieties do not share innovations with each other, and that they are not related in any way phonologically (in terms of potential super-heavy syllables), except that they happen to have mostly fully faithful forms and have not undergone any further innovations in this aspect of their phonologies.

6   Discussion and Conclusion

This study provides an OT analysis of potential super-heavy syllables in six Arabic dialects: Hijazi, Egyptian, Emirati, Kuwaiti, Algerian, and Palestinian. These represent 4 of the 5 main dialect areas of Arabic (Hijazi, Emirati, and Kuwaiti all represent Arabian dialects).

Kuwaiti, Algerian, and Palestinian Arabic are all considered conservative with regard to how they deal with potential super-heavy syllables. In these dialects, all forms with potential super-heavy syllables surface faithfully (i.e., the output form is the same as the input form, with no changes). Epenthesis and mora deletion are blocked because DEP-V and MAX-μ_V, respectively, are undominated, and a mora is not epenthesized to a third rhyme slot because *$\sigma_{\mu\mu\mu}$ dominates WBP.

Emirati is the closest of the other dialects to these conservative dialects. In Emirati Arabic, a vowel is epenthesized before a third consonant when there are three consonants in a row (e.g., in forms like /bint-na/). This epenthesis allows the second consonant to syllabify as the onset of the syllable containing the epenthetic vowel. This epenthesis and resyllabification process can be understood to be due to the fact that the *CCC constraint dominates DEP-V in Emirati, preventing three consonants from ever occurring in a row, and allowing epenthesis, which is used to resolve this problem.

Hijazi Arabic, like Emirati, does not allow three consonants to surface in a row, and repairs such structures with epenthesis. However, unlike in Emirati, /naːr-na/ is not allowed to surface faithfully either. This is because WBP and *$\sigma_{\mu\mu\mu}$ dominate DEP-V. Importantly, *CCC could be demoted, or even removed from the ranking altogether, and the same candidates would win. This demonstrates that in Emirati, *CCC — i.e., a constraint against three consonants in a row — is the main concern of the grammar, while in Hijazi, WBP and *$\sigma_{\mu\mu\mu}$ being undominated together makes it impossible for structures containing three rhyme timing slots to surface at all.

Egyptian Arabic's grammar is very similar to Hijazi Arabic. What makes the grammar of Egyptian different from the grammar of Hijazi is only the relative ranking of MAX-μ_V and DEP-V. Since DEP-V dominates MAX-μ_V in Egyptian Arabic, deletion of a vocalic mora (i.e. vowel shortening) is preferred to vocalic epenthesis as a repair strategy in forms where two of the rhyme timing slots in a potential super-heavy syllable are vocalic. Because of this, forms like /naːr-na/ surface as [nar.na] in Egyptian Arabic.

This study shows that an approach like Cho's (1998) quantification of steps between the grammars can shed light on how grammars change. This topic could be expanded in future studies by examining more Arabic dialects and the sonority effect of the coda cluster in these dialects. In

many Arabic dialects, sonority plays a role in vowel epenthesis between coda consonants — many dialects do not allow rising sonority, such as Hijazi, therefore vowel epenthesis would be the solution to break up a dispreferred coda cluster of rising sonority.

There are still some outstanding issues that are yet to be answered. These include how the loss of case could have sparked the changes that ended up giving rise to epenthesis in some of the dialects, the quality of the epenthetic vowel, stress of the epenthetic vowel and its interaction with vowel and mora epenthesis and deletion, and how the number of constraint reranking operations could be used for other purposes, like the determination of homelands, per Wichmann et al. (2010). Each of these topics could be a potential focus of future research. The wave diagram in (23) may not be sufficient to describe all observed changes among Arabic dialects.

Comments invited

PiHPh relies on post-publication review of the papers that it publishes. If you have any comments on this piece, please add them to its comments site. You are encouraged to consult this site after reading the paper, as there may be comments from other readers there, and replies from the author. This paper's site is here:

https://doi.org/10.2218/pihph.6.2021.6668

Author contact details

Hassan Bokhari
Indiana University Bloomington
Bloomington, Indiana
United States

hbokhari@iu.edu

References

Abu-Mansour, Mahasen. 1991. Epenthesis in Makkan Arabic: Unsyllabified consonants vs. degenerate syllables. In Comrie, Bernard and Mushira Eid (eds.), *Perspectives on Arabic linguistics: papers from the Annual Symposium on Arabic Linguistics,* vol. III, 137–154. Philadelphia: John Benjamins.

Al-Mohanna, Faisal. 2008. *Positional syllable maximality: Syllabification in Hejazi.* ROA 993.

Alahmari, Musa. 2018. *An Optimality-Theoretic analysis of some aspects of the phonology and morphology in a southwestern Saudi Arabian Arabic dialect.* Bloomington, IN: Indiana University Bloomington dissertation.

Alqahtani, Mufteh. 2014. *Syllable structure and related processes in Optimality Theory: An examination of Najdi Arabic.* Newcastle: Newcastle University thesis.

Bamakhramah, Majdi A. 2009. *Syllable structure in Arabic varieties with a focus on superheavy syllables.* Bloomington, IN: Indiana University Bloomington dissertation.

Bassiouney, Reem. 2009. *Arabic sociolinguistics: Topics in diglossia, gender, identity, and politics.* Washington, DC: Georgetown University Press.

Bermúdez-Otero, Ricardo. 2006. Phonological change in optimality theory. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn., vol. 9, 497–505. Oxford: Elsevier.

Bermúdez-Otero, Ricardo. 2020. The initiation and incrementation of sound change: Community-oriented momentum-sensitive learning. *Glossa* 5(1): article 121. 1–32.

Broselow, Ellen J. 1976. The phonology of Egyptian Arabic. Amherst, MA: University of Massachusetts at Amherst dissertation.

Cho, Young-Mee Yu. 1998. Language change as reranking of constraints. In Hogg, Richard M. and Linda van Bergen (eds), *Germanic linguistics volume 2: Selected papers from the 12th International Conference on Historical Linguistics, Manchester, August 1995.* Amsterdam: John Benjamins.

Davis, Stuart. 2011. Quantity. In Goldsmith, John, Jason Riggle, and Alan Yu (eds.), *The handbook of phonological theory*, 2nd edn., 545–562. Oxford: Wiley-Blackwell.

Davis, Stuart. 2017. Geminates and weight manipulating phonology in Chuukese (Trukese). In Haruo Kubozono (ed.), *The phonetics and phonology of geminate consonants*, vol. 2, 230–259. Oxford: Oxford University Press.

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

Gruber-Miller, Ann M. 1990. Loss of nominal case endings in the modern Arabic dialects: evidence from Palestinian Christian middle Arabic texts. In Mushira Eid (ed.), *Perspectives on Arabic Linguistics: Papers from the Annual Symposium on Arabic Linguistics*, 235–253, vol. I. Amsterdam: John Benjamins.

Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42(1). 25–70.

Holes, Clive. 2004. *Modern Arabic: Structures, functions, and varieties.* Washington, DC: Georgetown University Press.

Holt, D. Eric. 2015. Historical sound change in optimality theory: Achievements and challenges. In Honeybone, Patrick and Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 546–562. Oxford: Oxford University Press.

Kabrah, Rawiah. 2014. Stress assignment in Makkan Arabic: A stratal-OT analysis. In Benmamoun, Elabbas and Enam Al-Wer (eds.) *Perspectives on Arabic linguistics XXIV-XXV: Papers from the Annual Symposia on Arabic Linguistics,* 21–62. Philadelphia: John Benjamins.

Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

Kisseberth, Charles W. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1(3). 291–306.

Kiparsky, Paul. 2003. Syllables and moras in Arabic. In Féry, Caroline and van de Vijver, Ruben (eds.), *The syllable in Optimality Theory*, 147–182. Cambridge: Cambridge University Press.

Millar, Robert McColl. 2007. *Trask's historical linguistics*, 2nd edn. London: Hodder Education.

Ohala, John. 1981. The listener as a source of sound change. *Papers from the Parasession on Language and Behavior*, Chicago Linguistics Society. 178–203.

Ohala, John. 2012. The listener as a source of sound change: An update. In Solé, Maria-Josep and Daniel Recasens (eds.). *The initiation of sound change: perception, production, and social factors*. 21–36. Amsterdam: John Benjamins.

Rees, Daniel. 2008. *Towards Proto-Persian: An optimality theoretic historical reconstruction*. Washington, DC: Georgetown University dissertation.

Wichmann, Søren, André Müller & Viveka Velupillai. 2010. Homelands of the world's languages: A quantitative approach. *Diachronica* 27(2). 247–276.

Zenner, Eline, Dirk Speelman & Dirk Geeraerts. 2014. A usage-based onomasiological approach. *Diachronica* 31(1). 74–105.