A Phonological Typology of Modern Arabic Varieties

Emily Lindsay-Smith

St. Catherine's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2021

This thesis is dedicated to
my wife, Tris,
and all my family
— be they Lindsays, Smiths or Joneses —
for I could not have done any of this without you.

Acknowledgements

This thesis would not have been possible without the support and encouragement of many people. I am profoundly grateful to all those who have helped me to this point.

First and foremost, I wish to express my heartfelt thanks to my supervisor, Professor Aditi Lahiri. Over the past four and a half years, it has been a privilege and honour to receive such steadfast support and guidance. There are no words sincere and deep enough to express my gratitude for all you have done for me.

I must also thank those who have played a role in developing the ideas of this thesis. Particular thanks go to Wolfgang de Melo and Holly Kennard for their useful feedback at my Confirmation of Status; subsequent thanks also to audiences at the Manchester Phonology Meeting and Phonetik und Phonologie Tagung who heard earlier versions of this work; anonymous reviewers at the Texas Linguistics Society; not to mention Ricardo Bermudez-Otero for his useful insights.

I am also very grateful to my examiners, Elan Dresher and Wolfgang de Melo, for an enjoyable viva experience and valuable comments on my thesis that have improved this final version.

I am deeply thankful to my linguistics colleagues and friends at the Language and Brain Laboratory at Oxford, who have provided me with so much support, cups of tea, and fascinating chats. I have missed our lunches so much during the interruptions of lockdown and COVID, and I am so grateful for the opportunity to know you all. Thank you to Yoolim, Kim, Nina, Swetlana, Sandra, Hilary, Emily D., Colin, Henning, Yaxuan, Holly, Josh, Zoe, Anna, and to my non-LABLAB friends: Maren, Hannah, Amanda, Edwin, Raquel, Fin. Particular thanks must go to Colin, Amanda, Emily D. and Anna for proofreading, and to Leah for her help compiling the data appendix. Your support has been invaluable, and all remaining errors are my own.

Without the generous support of the Economic and Social Research Council, this work would not have been possible, and I would not have been able to spend the past few years immersing myself in a topic that has brought me such joy, and so bored my (non-linguist) friends and family.

I never would have made it this far without the ongoing support of my family. Thank you to my parents, Yvonne and Andrew Lindsay, who have always believed that I could do anything and so completely supported me. Thank you to my sisters, Joey and Georgie, for picking me up when things are hard, and putting up with nonsensical linguistics chat. Thank you to my grandmothers, Pat Jones and Elisabeth Lindsay, your support and love is deeply felt, and I wish my grandfathers were here to see this. Thank you to my in-laws, Diana and Kevin Smith, as well as Ellie, Joe and Sam Williams. Your support has been invaluable and I feel so welcomed into your family. I cannot list the myriad ways in which each of you have been there, but you have made all the difference.

Finally, and most importantly, I give my deepest thanks to my wife, Tris. Thank you for being there for me throughout, for loving me when I am stressed and when I am excited and all the states in between. You are the best woman I have ever met. You may not believe in syllables but I love you anyway.

Abstract

Modern Arabic colloquial varieties exhibit significant phonological variation that has only partially been incorporated into previous typological efforts. This thesis motivates a phonological typology of sixteen modern varieties, exploring not just the crucial ingredients of the phonological grammar including syllable and foot structure, but also the lexical and postlexical nature of prosodic structure.

I argue that there are two axes for phonological variation in Arabic: TOLERANCE and REPAIR, that is, varieties differ in terms of which types of syllables they tolerate and in how they repair violations of syllable structure, and these axes do not covary. In terms of TOLERANCE, varieties differ in acceptance of quantity. Varieties that permit long segments have restrictions on the *segmental* level, whereas varieties that do not have restrictions on the *moraic* level, and varieties that allow long consonants or clusters but not long vowels have restrictions on the *X-slot* tier. In addition to the canonical 2 or 3 elements, individual varieties may also permit an extra element in their medial syllable rime according to variety-specific restrictions including coronality, fast speech, and sonority. As such, this typology provides mechanisms to account for the data in each and every variety, rather than merely providing a canonical set of features that not all varieties fit.

REPAIR is concerned with how violations of syllable structure are repaired in terms of epenthesis direction and the licensing of word edge extrasyllabic segments. I propose that catalectic moras are required to incorporate extrasyllabic material across the word. The ONSET/CODA distinction can therefore be expressed as whether the extrasyllabic material is incorporated as the onset or coda of a syllable containing a catalectic mora. Epenthesis occurs where the catalectic mora is not permitted to surface without overt segmental material. The difference between domain edge behaviour therefore follows from this — where the catalectic mora is the peripheral element, the surface cluster is permitted; but where it is non-peripheral, epenthesis occurs as the catalectic mora is filled by an epenthetic vowel. 'Extreme' ONSET/CODA varieties are those that permit catalectic moras word-internally in domain edge syllables.

Furthermore, I explore the prosodic word structure of Arabic, finding that there is a preference for affixes to attach after a stressed syllable that accounts for lengthening phenomena and apparent stress exceptions.

This thesis demonstrates that there are constraints on both the input and output of the phonological grammar, and that the variation rests on the acceptability or violation of the constraints. Thus, the Arabic typology emphasizes two aspects: first, variation in phonology is not just on the surface, and second, abstract quantitative analyses including catalexis are fundamental to our understanding of how prosodic structure can survive across a range of dialects.

Contents

List of Figures

List of Tables

Glossary

1	first person
2	second person
3	third person
APPL	applicative
COMP	complementizer
CONJ	conjunction
DAT	dative
DEF	definite
DET	determiner
DU	dual
F	feminine
FUT	future
HAB	habitual
IMP	imperative
IMPF	imperfective
IND	indicative
LOC	locative
M	masculine
NEG	negative
NON-FIN	Non-Finality
OBJ	object
OT	Optimality Theory
PASS	passive
PFV	perfective
PL	plural
POSS	possessive
PRF	perfect
PRS	present
PST	past
PTCP	participle
SBJ	subject
SBJV	subjunctive
SG	singular

List of Abbreviations

C . . . . Coda, elsewhere Consonant

CSS-Morph . Morphologically-triggered vowel shortening and raising

CSS-Phon . . Phonologically-triggered vowel shortening without raising

ERR . . . . . . End Rule Right

H . . . . Heavy (syllable)

L . . . . Light (syllable

L–R . . . . . . Left-to-Right Parsing Directionality

MSA . . . . . . Modern Standard Arabic

MST . . . . . . Metrical Stress Theory

MT . . . . Moraic Trochee

N . . . . Nucleus

Non-Fin . . . . Non-Finality

O . . . . Onset

OT . . . . Optimality Theory

R . . . . Rime

R-L . . . . . . Right-to-Left Parsing Directionality

S . . . . Strong (branch of a foot)

SSC . . . . . . Stress Subordination Convention

V . . . . Vowel

W . . . . Weak (branch of a foot)

X . . . . X-slot

σ . . . . Syllable

μ . . . . Mora

1 Introduction

Contents

Modern Arabic vernacular varieties display drastic differences in their phonology, syntax and lexicon despite their close historical origins. This thesis deals with the phonological typology of sixteen modern varieties, exploring not just the crucial ingredients of the phonological grammar, including syllable and foot structure, but also the lexical and postlexical nature of prosodic structure.

1.1 Arabic Context

Arabic is a non-concatenative Central Semitic language with around 335,176,770 native speakers (Eberhard et al. 2020) across North Africa, the Middle East, and expatriate communities around the world from Australia to Venezuela. ‘Arabic’ is an umbrella term that covers Modern Standard Arabic, the formal standard variety, a wide range of vernacular varieties, and Classical Arabic as a liturgical language.

Arabic originated in the Arabian Peninsula and Syrian steppe, before being spread through invasion from the second half of the seventh century. The conquests of the Rashidun Caliphs (632-661) after the death of Muhammad spread into the Levant, Egypt, Mesopotamia and Persia; the conquests under the Ummayads (661-750) spread into Sindh, North Africa, Hispania, Transoxiana and modern-day Afghanistan. Arabic declined in importance and geographical spread after the Mongol sacking of Baghdad in 1258, withdrawal from Andalusia from 1492, and rise of the Ottoman Empire. However, Arabic continued to be spoken throughout the region as a mother tongue. The predominantly Arabic speaking areas of North Africa and the Middle East today are depicted in the map in Figure 1.1. Here we see marked in brown that Arabic is spoken across North Africa — Morocco, Algeria, Tunisia, Libya — and into Egypt and the Sudan. In the Middle East, it is spoken in the Levantine areas of Palestine, Lebanon, Syria and Jordan, as well as Mesopotamian Iraq. It is also spoken in the Arabian Peninsula, in Saudi Arabia, Kuwait, Bahrain, Qatar, the United Arab Emirates (UAE), Oman and Yemen in the south of the Peninsula.

Figure 1.1: Map of the Arabic Speaking Areas of the Middle East and North Africa from Farrag et al. (2017)

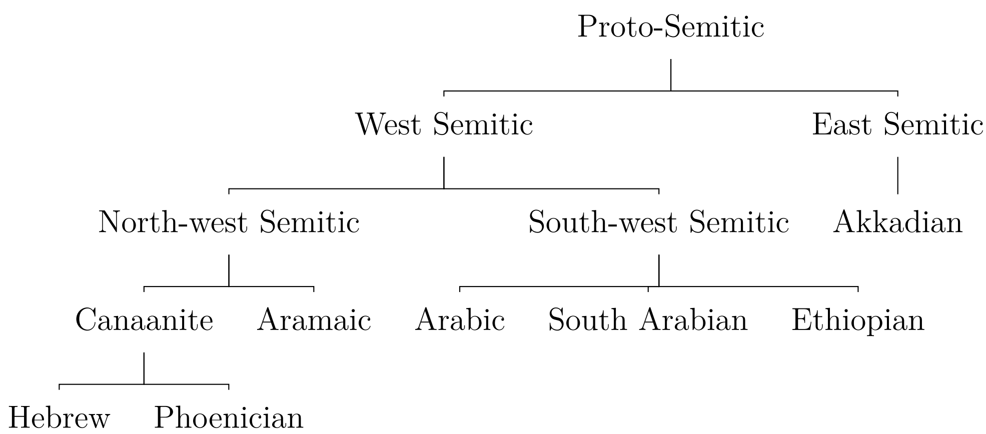
In what follows I will briefly trace the relationship between Arabic and other Afroasiatic languages in Section 1.1.1, the different types of Arabic in Section 1.1.2, the historical origins of the modern varieties in Section 1.1.3, the ways Arabic varieties have been classified in Section 1.1.4, the diglossia throughout the region in Section 1.1.5 and the research on Arabic dialectology from the 19th century onwards in Section 1.1.6 before introducing the data used in this thesis in Section 1.2 and the theoretical background in Sections 1.3 and 1.4. Finally, I give an overview of the contributions of the thesis and its structure in Section 1.5.

1.1.1 Afroasiatic Genealogy

Semitic is part of the Afroasiatic language family, though there is controversy over whether this family is a true family or prehistoric Sprachbund. If it is assumed to be a true language family, the internal classification of this family is a controversial matter.

The widespread approach for most of the twentieth century (Nöldeke 1899; Nöldeke 1911; Brockelmann 1908) based classification on geographical location, as shown in Figure 1.2 below.

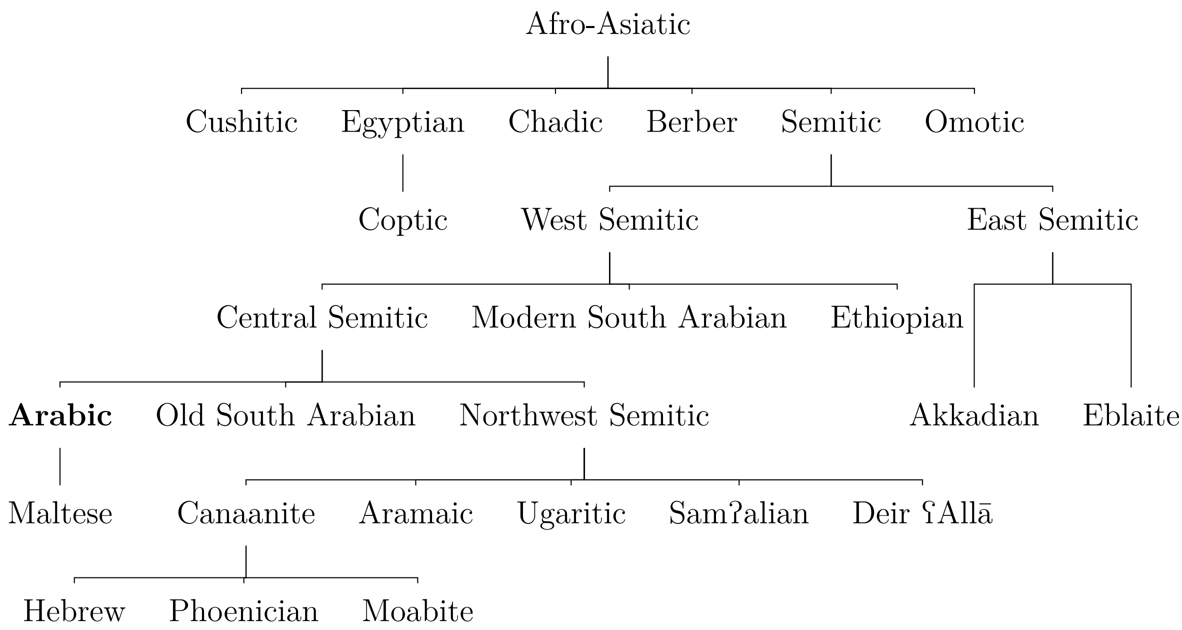
Figure 1.2: Traditional Classification of Semitic Languages



From Versteegh (2014, p. 11)

Hetzron (1974; 1975; 1976) challenged this, arguing for the importance of morphological innovations within Semitic for the grouping of Semitic languages. Arabic seems to share the conservative phonological system of southern semitic languages but shares morphological innovations with the northern parts of Central Semitic (Al Sharkawi 2017). Across Semitic more broadly, the placement of Ethiopic and South Arabian are the key differences between Hetzron's morphological innovation driven model and the traditional geographic model. While Hetzron's

Figure 1.3: Afroasiatic Language Family focusing on Semitic



approach has been disputed (Nebes 1994; Porkhomovsky 1997; Stempel 1999; Lipiński 2001; Haelewyck 2007), it remains the foundation for subsequent work (Rubin 2011, p. 261).

The family tree in Figure 1.3 below is a modification of his model (Rubin 2008, p. 62) with the main subgroupings of the Afroasiatic family added (van der Hulst and Hellmuth 2010, p. 617). Of course, the genealogical model of the family tree cannot account for the effects of language contact explained by a wave model, and indeed as Labov (2007, p. 345) argues 'any general view of language descent must be prepared to integrate the two models'. Furthermore, Ullendorff (1970) , Garbini (1984) and Al Sharkawi (2017) question the legitimacy of attempting to propose a genetic relationship between these languages at all, given quite the degree of contact between them. However, this thesis is not concerned with the fine details of Semitic language family development, so the tree in Figure 1.3 is given to contextualise Arabic amongst related languages[1].

[1]For more information, see Faber (1997), Bennett (1998), Appleyard (2003), Rubin (2008), Weninger et al. (2011), Versteegh (2014) and references within.

1.1.2 Types of Arabic

'Arabic' covers a multitude of varieties distributed across time and space. For the varieties distributed over time, the following terms are used:[2] *Old Arabic* refers to pre-diasporic Arabic though the exact chronological bounds vary between scholars. *Classical Arabic* is the language of the Qur'ān, used from pre-Islamic Arabia[3] to the Abbasid Caliphate[4]. This is the variety discussed by classical grammarians such as Siːbawayhi (d.793) and lexicographers such as Ibn Khaliːl (d.791) . It is still used today across the world as a liturgical language. *Middle Arabic* refers to texts deviating from Classical Arabic Grammar. This is not a chronological stage per se, but a style.[5]

More important for this thesis are the varieties used today. Colloquial varieties are learned as mother tongues, and used in all non-formal domains (Jarrar et al. 2017). While primarily used orally, they are found in literature, lyrics, films, soap operas, social media, cartoons and adverts. There is no standardised spelling in either the Arabic or Latin script, but conventions particularly on social media are emerging. For vernacular variety use in literature, see Aguadé (2006) for Morocco, and see Rosenbaum (2004) and Woidich (2010) for Egypt. A standard form known as Modern Standard Arabic (MSA) is used across the Arabic speaking world, but is not learnt as a mother tongue — rather it is acquired at school. It is used in literature, official documents, news bulletins, and sermons.

1.1.3 Ancestor of the Modern Varieties

Whether the modern colloquial varieties shared the same common ancestor, or a range of closely related ancestors, remains unresolved and controversial. This is not an unusual problem for Historical Linguistics — later varieties tend to continue a sister variety of a better attested earlier variety. For example, Anglo-Saxon is best

[2]These terms do not correspond to coherent historical stages, or even necessarily linguistic commonalities. See Owens (2006) for a criticism.

[3]According the traditional narrative, Muhammad was born c.570, began preaching from c.609, and died 632. By his death, most of the Arabian peninsula had been converted to Islam.

[4]Abbasid Caliphate ruled 750-1258, and then 1261-1417 under the Mamluk Sultanate of Cairo.

[5]See Larcher (2001).

attested in the West Saxon form, but Middle English is descended from Anglian, Pali continues a sister variety of classical Sanskrit, and Middle Chinese is best attested in the Wu group rather than Mandarin.

Different approaches have been taken to this question for Arabic. Where monogenesis is assumed, there are disagreements with regard to this origin — are the modern varieties descendants of the Classical Arabic described by the Arab grammarians, or are they descendants of a colloquial language spoken in Mekka and Medina before the advent of Islam (Vollers 1906; Versteegh 1984; Holes 2004), or are they descendants of a post-Islam military koiné (Fück 1950; Ferguson 1959a)? However, others argue that the modern varieties are descended from many different ancient varieties (Edzard 1998; Cohen 1970). While many features distinguishing modern varieties from Classical Arabic have been identified by Ferguson (1959a), Cohen (1970) and Versteegh (1984), Behnstedt and Woidich (2005) demonstrated that some dialects fail to exhibit many of these features, whereas some of these supposedly modern innovations may have existed in ancient varieties of Arabic.

Furthermore, it is clear that the Arabic spoken in the 6th to 10th centuries was not a single variety but rather a multiple distinct varieties. These varieties were not neatly contained, but rather there is evidence of extensive intertribal contact (Al Sharkawi 2017). Arab grammarians noted and named the allophonic variation present in dialectal phenomena, such as *ʕanʕanah*,[6] *kaʃkaʃah*,[7] *taltalah*,[8] and *ʕaǰʕajah*,[9] although these colloquial varieties were not their primary focus (Sibawayhi and Hārūn 1982/AH 1402; Rabin 1951; Cadora 1992; Al Sharkawi 2017).

Investigating the historical origins of the modern varieties is complicated by the limited material, but remains a fruitful and important avenue of research.

Regardless of the exact origin of these varieties, it is clear that they exhibit both close similarity and drastic differences. How to classify these axes of variation has

[6] *ʕanʕanah*: the substitution of the pharyngeal fricative ʕ for a glottal stop, as in ʔan:aka > ʕan:aka, *that you*.

[7] *kaʃkaʃah*: substitution of post-alveolar fricative ʃ for velar plosive k in feminine singular suffix, as in biki > biʃi, *in you (f.s.)*.

[8] *taltala*: use of /i/ as the theme vowel in the imperfective subject marker, rather than /a/.

[9] *ʕaǰʕajah*: substitution of /ig/ for /iː/, as in tamimiː > tamimig *from Tamimi*.

been a key focus of comparative dialectology in the 20þ century.

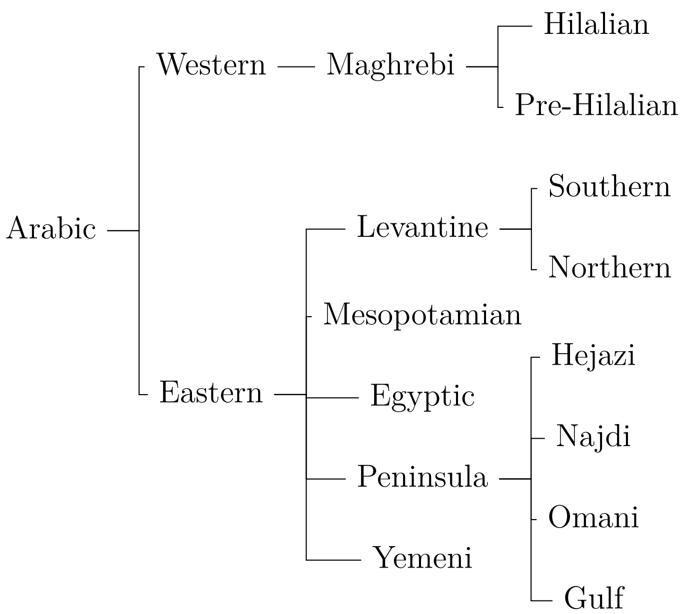
1.1.4 Classifying Modern Arabic Varieties

In comparative dialectology work, Arabic varieties have been classified not based directly on linguistic features alone, but rather on extralinguistic features that are then, post hoc, confirmed through linguistic features. The main axes of classification are geographical, historical, and social.

In terms of geographical divisions, Arabic varieties are divided into Western and Eastern, or Maghrebi and Mashriqi. The geographical boundary between the Eastern and Western dialects runs southward from the Mediterranean coast along the western borders of Egypt to Lake Chad (Palva 2006, pp. 605-6).[10] A key linguistic feature correlated with this boundary is the morphological exponents for subject. In the Western varieties we find [ni-ktib] ‘I write’ and [ni-ktib-u] ‘we write’; whereas in the Eastern varieties we find [ʔa-ktib] ‘I write’ and [ni-ktib] ‘we write’. In addition, there is often a lack of phonemic distinction between short vowels in the Western varieties, often only recognised as a schwa. For a review of the phonetic differences on geographical lines, see Al-Wer and de Jong (2018). This boundary between Eastern and Western varieties has become blurred through population movement (Woidich 1993; Behnstedt and Woidich 2005). Further geographical divisions are recognised though the exact grouping varies between scholars, with Jarrar et al. (2017) not including Yemeni as a separate group but subsuming it under Peninsular Arabic, and not all agree with the more fine grained final level of differentiation displayed in Figure 1.4 below.

[10] But see Owens (2003) for a criticism.

Figure 1.4: Geographical Classification of Modern Arabic Varieties



In terms of historical divisions, these are done broadly based on the spread of Arabic. Arabic originated in the Arabian Peninsula and Syrian steppe, before being spread through invasion from the second half of the seventh century. As described above, the conquests of 632-661 under the Rashidun Caliphs after the death of Muhammad spread into the Levant, Egypt, Mesopotamia and Persia; the conquests under the Ummayads (661-750) spread into Sindh, North Africa, Hispania, Transoxiana and modern-day Afghanistan. Accordingly, Zone I is the Arabian Peninsula, Zone II is the Levant, Iraq, Egypt, North Africa and Iran, and Zone III is the Sprachinseln found further afield, in Anatolia, Iran, Afghanistan, Uzbekistan, Cyprus, Malta and sub-Saharan Africa (Behnstedt and Woidich 2013). However, this approach has its limitations. Despite all forms of 'colonial' Arabic being in Zone II, striking variation is found across these forms of Arabic. Whether it is due to population movement and contact in these areas (Watson 2011b) or a matter of pre-diaspora variation in Arabic in these regions (Retsö 2000; Owens 2006) is unclear. Furthermore, the Zone III areas have little in common with each other beyond being Sprachinseln, and exhibit the results of both different origins in time in addition to varying degrees of contact with other neighbouring languages.

In terms of social divisions, classifications occur on grounds of Sedentary versus Nomadic, Urban versus Rural, and Sectarian differences. Marçais (1938) first introduced the Sedentary versus Nomadic or Bedouin distinction to account for the patterns of arabization of North Africa. This distinguishes between the varieties related to those of the nomadic Bedouin and those from settlements. Bedouin varieties have distinct features, such as the voiced [g] reflex of the Classical Arabic *q, feminine plural forms, and interdental consonants. Bedouin varieties retain [θ] and [ð] for the Classical Arabic *θ and *ð, whereas sedentary varieties have [t] and [d] respectively. The Urban versus Rural distinction is similar, with the glottal stop or voiceless reflex of the Classical Arabic *q noted as a particular urban feature. Sectarian divides are also noted, such as in Baghdad where traditionally a distinction has been made between Muslim, Jewish and Christian varieties. However, these sectarian lines really reflect population movement, where the Baghdad Muslims came from Bedouin communities in southern Iraq but Christians came from towns in Northern Iraq (Blanc 1964; Abu-Haidar 1991; Mansour 1991). However, these classifications are increasingly unreliable, due to dialect levelling, spread of urban speech around urban centres and along trade routes, as well as population movement.

1.1.5 Diglossia

Western Arabicists have explored the complex relationship between MSA and colloquial varieties since the early 20th century, and often describe it as *diglossia* — though exactly what this term means has changed over time.

Marçais (1930) claimed 'la diglossie arabe' in North Africa involved two completely separate entities: 'l'arabe écrit' and 'l'arabe parlé'. Lecerf (1932) makes similar claims about the Levant, but suggested there were signs of a shift with more mixing between the two following the advent of the printing press.

Diglossia in the sense of a hierarchical contrast between two varieties used by the same language community was formalised by Ferguson (1959b, pp. 328-335) who made the following distinction between the standardised or prestigious form, H, and the vernacular form, L. He argues there are specialised functions for each variety with limited overlap. Ferguson's claims as to the specialised functions and distributions are as follows. The H variety is considered superior to L by native speakers, and there is usually a large body of esteemed literature written in H. L is acquired as a mother tongue while H is acquired through formal education. As a result, speakers have a better command of L. There is a long tradition of grammatical study of H while limited or descriptive studies of L either do not exist or were written very recently. This diglossic situation continues for centuries, though uncodified intermediate forms can arise with vocabulary borrowings from H but use of L morphology and syntax. There is often much difference between H and L grammatical structures and those of L are often (viewed as) simpler. The majority of vocabulary is shared between H and L with variations in form and semantics, but often there are pairs of lexemes where usage is in solely H or L contexts. Finally, the phonological system can be similar or can diverge between H and L. Where it diverges, L is the basic system, and H diverges as a subsystem.

Subsequent approaches have described diglossia as a series of 4 or 5 levels (Meiseles 1980; Badawi 1973)[11] or a gradual continuum (Blanc 1960) between colloquial and literary forms with potentially infinite numbers of levels (Bassiouney 2009). Albirini and Powesland (2011) argue that the distinction between MSA and colloquial dialects is determined by function not by context. Eid (1982; 1988) proposed applying the code-switching framework to analyse 'diglossic' switching between H and L. This was taken up by Walters (1996a; 1996b), Boussofara-Omar (2003) and Bassiouney (2006) among others. Earlier works imply a stricter division between H and L than this, perhaps reflecting a change over time following increased education and literacy, and therefore wider fluency in MSA. This trend is likely to continue following the advent of social media, with more writing in colloquial Arabic and more mixing of varieties (Mejdell 2018).

[11] Badawi and Jankowsky (1985, p. 16) distinguishes between *fusˤ*ħ*a al-turaːtˤ*, 'language of heritage', that is, Classical Arabic; *fusˤ*ħ*a al-*ʕ*asˤr*, 'contemporary language', that is MSA; ʕaː*mːiyat al-mutˤaqːafiːn* 'colloquial of the educated'; ʕaː*mːiyat al-mutanawːariːn* 'colloquial of the semi-educated'; ʕaː*mːiyat al-*ʔ*umːiyːiːn* 'colloquial of the uneducated'.

1.1.6 Research on Arabic Dialectology from the 19th Century Onwards

Arabic dialectology began at the end of 19th century, with initial focus on grammars and dictionaries such as Spitta (1880). Dialect atlases began to be produced in the early 20th century, such as Bergsträßer (1915) for Palestinian, Cantineau (1940) for Hōrān, Syria and Cantineau (1946) for Palmyra, Syria. Abul-Fadl (1961) pioneered research on the geographical distribution of phonological and morphological features in the ʃarqiyːa province in Nile Delta. Subsequent dialect atlases have been produced for Egyptian varieties (Behnstedt and Woidich 1985; Behnstedt 1987; Behnstedt 1988), North Yemeni varieties (Behnstedt 1987; Behnstedt 1992) and Syrian varieties (Behnstedt 1997). Today, many works can be found describing individual varieties, ranging in detail and length. However, coverage of the whole Arabic speaking world is still patchy (Versteegh 2014).

Studies of Arabic phonology have primarily focused on changes to phoneme inventories and segmental processes of syncope and epenthesis (Garbell 1958; Birkeland 1952; Jakobson 1957; Mustawafi 2018; Holes 2018). There are a multitude of studies on contemporary stress (see Watson (2011a) for a summary), and some work on the correlation of syllable structure and other phonological processes across groupings of varieties (Broselow 1992; Kiparsky 2003; Watson 2007; Farwaneh 2009).

However, the existing work on the phonological typology of modern Arabic varieties is incomplete, unable to account for non-canonical behaviours, and uses a disparate range of strategies to solve apparent exceptions. In this thesis, I contribute a new bipartite typology of Arabic stress and syllable structure, with better empirical coverage of quantity tolerance, superheavy syllables, stress, and the morphological interface with prosodic structures. This research has broader consequences for linguistic theory, in particular with regards to theories of syllable structure restrictions, the reanalysis of catalexis as a tool to repair degenerate syllables as well as degenerate feet, and the clitic versus affix distinction.

1.2 Investigating Sixteen Modern Arabic Varieties

In this thesis I explore sixteen modern Arabic varieties: Moroccan Casablanca, Tunisian, Algerian, Libyan Tripoli, Cairene, Lebanese, Palestinian, Wadi Ramː, Iraqi, Makːan, Rufaidah, Haːyil, Rwaili, Muscat, Qatari, and Sˤanʕaːniː. These were chosen to express the range of variation found across the subgroups depicted in Figure 1.4 on page 8. Where these varieties are spoken, the primary designations of the variety and the main data source is detailed in Table 1.1 and discussed below. The location of these varieties can be seen in the map in Figure 1.5.

Sources differ in the detail they provide as to the origin of their data and specificity of social details around speakers of the variety in question. While the origin of much of this data is not ideal, in that there is extensive introspection of native speakers and where elicitation occurs details are lacking as to how robust this procedure has been, this is typical of much existing work on Arabic. I chose data from these sources in accordance with my own knowledge of Arabic developed through extensive university-level study. While in some cases I have used loanwords to illustrate phonological patterns, I have ensured that the key points of my analysis do not depend solely on loanword adaptation.

Table 1.1: Data Sources Used in the Thesis

	Location	Characteristics	Source
Moroccan Casablanca	Casablanca, Morocco	Maghrebi	Boudlal (2001)
Algerian	Oran, Algeria	Maghrebi	Bouhadiba (1988)
Tunisian	Nabeul, North-East Tunisia	Maghrebi	Maamouri (1967)
Libyan Tripoli	Tripoli, Libya	Maghrebi, Jewish	Yoda (2005)
Cairene	Cairo, Egypt	Egyptic	Watson (2002)
Palestinian	Palestine; Jordan; diaspora communities	Levantine, Southern	Abu-Salim (1982a)
Lebanese	Beirut, Lebanon	Levantine, Northern	Haddad (1984)
Wadi Ramː	Wadi Ramm, Jordan	Levantine, Bedouin	Al Mashaqba (2015)
Iraqi	Bahgdad, Iraq	Mesopotamian	Majdi (1988)
Muscat	Muscat, Oman	Peninsula, Omani	Glover (1988)
Qatari	Qatar	Peninsula, Gulf	Al-Sulaiti (1993)
Rufaidah	Rufaidah and Dammam, Southern Hijaz, Saudi Arabia	Peninsula, Hejazi	Prochazka Jr. (1988)
Haːyil	Riyadh, Saudi Arabia	Peninsula, Shamːari Najdi, Bedouin	Prochazka Jr. (1988)
Rwaili	Riyadh, Saudi Arabia	Peninsula, ʕAnazi Najdi, Bedouin	Prochazka Jr. (1988)
Makːan	Makka, Saudi Arabia	Peninsula	Kabrah (2004)
Sˤanʕaːniː	Sanʕa, Yemen	Yemeni	Watson (2002)

Figure 1.5: Map of the Locations of the Varieties from the thesis

Moroccan Casablanca is a Maghrebi variety spoken in Casablanca, a coastal city in Western Morocco. As the biggest economic and industrial city in Morocco, Casablanca has a heterogenous population due to movement particularly from the local rural areas into the city, and there has been historical and modern contact with Berber[12], French, Spanish and English among other languages. Despite the heterogeneity of the population, a recognisable homogenous variety of 'Casablanca Moroccan Arabic' exists, and exhibits Bedouin features due to the historical Bedouin settlements in the area. The data source for this variety is Boudlal's 2001 doctoral thesis completed at the Université Mohammed V in Rabat, Morocco. Boudlal, a native speaker of Moroccan Casablanca, elicited data from family members and friends born in Casablanca whose parents do not speak Berber and have lived in Casablnca for a long time. In addition to the elicited data, his work includes data from published work (especially Harrell (1962) and Abdelmassih (1973)) that is confirmed by native speakers. Beyond this, Boudlal is not explicit about the nature of data collection or origins of individual items, except for the stress data. For stress placement, he carried out a native speaker judgement test on a corpus of 39 items with 60 subjects as well as measuring the fundamental frequency, intensity and duration of syllables in recordings from 5 subjects.

Algerian is a Maghrebi variety spoken in Oran, a city on the North-West coast of Algeria. There has been extensive contact with other languages in Oran, notably Berber, Spanish, Turkish, French and other local varieties. The data source for this variety is Bouhadiba's 1988 doctoral thesis completed at the University of Reading. Bouhadiba is a native speaker, and as such the data he cites does in part come from his own introspection; however, he states some data is collected from 20 hours of tape recordings of apparently representative speakers of Oran Arabic collected in Oran and surrounding suburbs. Informants are chosen based on phonetic/semantic considerations of their speech. Which data comes from his own introspection and which from elicitation is not specified.

[12]This is disputed; see Diem (1979, pp.52–55), El Aissati (2006), Taine-Cheikh (2007), Chtatou (2009), Bensoukas and Boudlal (2012) and Aguadé (2018, p. 35) inter alia.

Tunisian is a Maghrebi variety spoken in Nabeul, a seaside resort on the North-East coast of Tunisia. As with the other Maghrebi varieties, there has been extensive dialect contact with Spanish, Turkish, Italian and French. The data source for this variety is Maamouri's 1967 doctoral thesis completed at Cornell. As a native speaker born in Nabeul, part of the data comes from his own introspection. However, some is also specifically elicited from 9 male Tunisian students aged 20-30 studying at Cornell. These informants come from across Tunisia, all speak fluent French and at least some English.

Libyan Tripoli is a Maghrebi variety that was spoken in by the Jewish population of Tripoli, on the Mediterranean coast of Libya about 100km from the Tunisian border. In 1948, there were around 20,000 Jews in Tripoli, presumably speaking this variety, but most have left through subsequent migrations. The data source for this variety is Yoda's 2005 research monograph describing the phonology and morphology of this variety. The data comes from 35 hours of recordings with 9 informants (5 men, 4 women) who previously lived in Tripoli city but have subsequently migrated to Israel (7 in 1948-1951, 2 in 1967).

Cairene is an Egyptic variety spoken in Cairo, the capital of Egypt. Cairo is a large city, and has faced extensive language contact including French, English, Turkish and other Arabic varieties. Cairene is a well-studied Arabic variety. The main data source for this variety is Watson's 2002 research monograph discussing the phonology and morphology of Cairene and Sˤanʕaːniː. Watson does not specify the source for her data, but the data she includes does not contradict the expected patterns from other work on Cairene. All Cairene data comes from Watson unless otherwise specified. The second source for Cairene data is Broselow's 1976 doctoral thesis completed at the University of Massachusetts, Amherst. Broselow uses predominately already published sources of data as well as a native informant.

Palestinian is a Levantine variety spoken in Palestine, Jordan, and diaspora communities around the world. The main data source for this variety is Abu Salim's 1982 doctoral thesis completed at the University of Illinois at Urbana-Champaign. Abu Salim is disappointingly vague as to the origin of his data and indeed specificity about where this variety of Palestinian Arabic is spoken. He is a native speaker, born in Jordan, so data is likely predominately through introspection. His data is described as urban Palestinian by Younes (1995) who discusses a differing variety described as 'Rural Palestinian'. While this lack of specificity is frustrating, it is not uncommon in many works on Arabic. However, the data he provides is coherent with regards to other published work on Palestinian Arabic.

Lebanese is a Levantine variety spoken in Beirut, the capital of Lebanon. Lebanon has experienced extensive language contact, particularly with French and English. The data source for this variety is Haddad's 1984 doctoral thesis completed at the University of Illinois at Urbana-Champaign. Haddad does not specify the origin of the data included, but does thank Lebanese informants in his acknowledgements.

Wadi Ram: is a Levantine variety spoken in Wadi Ramm, the largest valley in Jordan, in the south near the Saudi Arabian border. Most speakers are from the Zalabian and Zawaidih subtribes, with around 6250 speakers. The bedouin population maintains a nomadic pastoral lifestyle, but there is language contact through the tourist industry in the area, so some are fluent in English, as well as contact with neighbouring Arabic varieties. The data source for this variety is Al Mashaqba's 2014 doctoral thesis completed at the University of Salford. AlMashaqba is the most explicit of all the sources used in this thesis, specifying that he collected data through fieldwork with snowball sampling of male informants aged 26-72 who have lived most of their lives in Wadi Ramm. AlMashaqba collected audio recordings of monologues, direct elicitation sessions and conversations between native speakers, and confirmed his data afterwards with native consultants.

Iraqi is a Mesopotamian variety spoken in Baghdad, the capital of Iraq. Alongside Arabic, Kurdish, Turkish and Syriac are widely spoken in Iraq. The data source for this variety is Majdi's 1988 doctoral thesis completed at the University of Connecticut. Majdi is not explicit as the origin of his data. Some is previously published in Erwin (1963) and Blanc (1964) and Odden (1978), but some of this data he refutes presumably based on his own introspection.

Muscat is a Omani Peninsular variety spoken primarily by ’Ibaːdˤi Muslims in Muscat, the capital of Oman. There is extensive language contact in this port city. Glover (1988, pp.12–13) notes that "Among the languages spoken are Egyptian, Sudanese, and Levantine Arabic; English, Dutch, French, German, Greek and other European languages; Hindi, Urdu, Balu:chi, Bengali, Gujarati, Sindhi and Singhalese among Indo-Iranian languages; Malayalam and Tamil among Dravidian languages; and Philippine languages among others from Southeast Asia". The data source for this variety is Glover’s 1988 doctoral thesis at the University of California. In light of the extensive language contact, Glover primarily collected her data from women living in Muscat as their speech is more conservative. She elicited data through monologues and dialogues from predominately monolingual speakers.

Qatari is an Gulf Peninsular variety spoken in Qatar. The data sourse is Al-Sulaiti’s 1993 doctoral thesis completed at the University of Lancaster. Al-Sulaiti is a native speaker, so part of her data comes from introspection, but includes recorded speech of around twenty speakers in conversation as well as monologues from five speakers.

Rufaidah, Rwaili, Haːyil are Peninsular varieties spoken in Saudi Arabia. Rufaidah is a sedentary Hijazi variety, whereas Rwaili and Haːyil are Bedouin Najdi varieties. With the spread of education in Saudi Arabia bringing more knowledge of Modern Standard Arabic, and urban expansion facilitating more language contact, there gradual dialect levelling is occurring in this region (Prochazka Jr. 1988, p. 9). The source for all three varieties is Prochazka Jr.’s 1988 research monograph, with data elicited through questionnaires with native speakers.

Makːan is a Hejazi Peninsular variety spoken in Makka, Saudi Arabia. The data source is Kabrah’s 2004 doctoral thesis completed at Boston University. Kabrah is vague as to the source of her data, but presumably it is through introspection as she is a native speaker born in Makka.

Sˤanʕaːniː is a Yemeni variety spoken in the old city of Sanʕa, in the capital of Yemen. It is very conservative variety, with phonology, morphological and syntactic features typically considered rural or Bedouin (cf. Fischer and Jastrow

(1980, p.24)). The main data source for this variety is Watson's 2002 research monograph discussing the phonology and morphology of Cairene and Sˤanʕaːniː. Watson is not explicit as to the sources of her data, but she did conduct fieldwork in Sanʕa so this is the likely source.

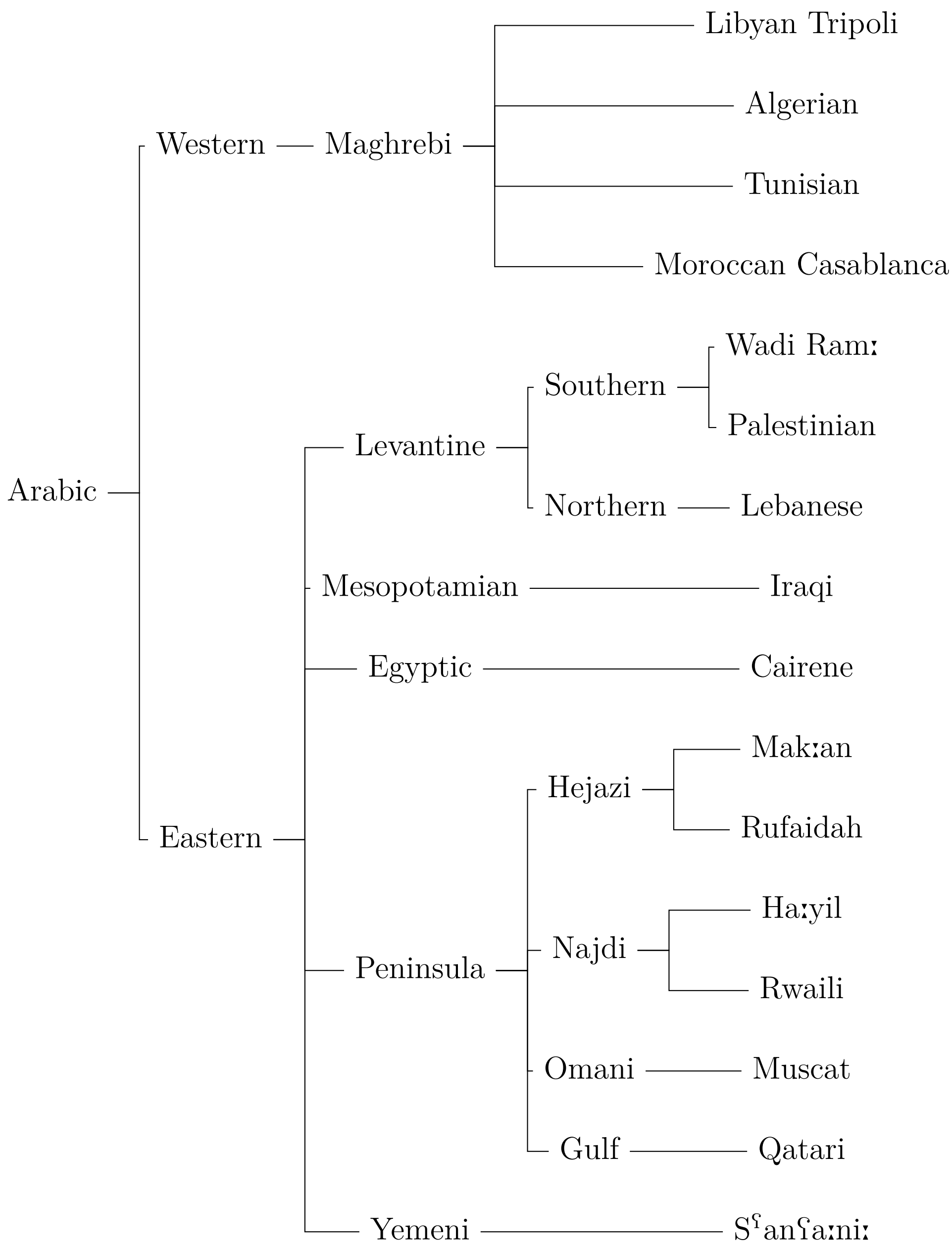
1.2.1 Justification of Choice of Varieties

These varieties were chosen as they represent a wide range of types of Arabic varieties, there was sufficient data to explore all the questions raised, and the data available was internally coherent in terms of the wider literature, expected phonological patterns, and my own knowledge of Arabic. The four Maghrebi varieties cover the different countries in North Africa. Cairene is perhaps the most studied Arabic variety, as well as the variety best understood across the region as it is spoken on TV. The three Levantine varieties provide a range of the phenomena found in this region, and include a Bedouin variety. The six Peninsula varieties include varieties from Oman and the Arabian Gulf coast, as well as four varieties including two Bedouin varieties from Saudi Arabia that represent the main dialect groups in this region. Sˤanʕaːniː is included as a Yemeni variety, that also sometimes is included as its own group distinct from the Peninsular varieties. This distribution across the geographical classifications is illustrated in Figure 1.6.

I did not include Arabic spoken in central Africa, such as in the Sudan, as the high degree of contact with neighbouring non-Semitic languages has influenced its phonology as noted in Watson (2007). Nor did I include the Arabic varieties spoken in parts of Asia, Mauritania or Malta, for the same reason. Wadi Hadˤrami would have been a good addition to the Yemeni group, but the data was contradictory as to the behaviour of medial geminates (Al-Saqqaf 1999). Other studies of Bedouin varieties lacked the comprehensive array of data needed for this particular study, but the analysis developed here would be enriched by inclusion of Bedouin data in future work.

This array of data is sufficient for the work completed here. Future investigations on this topic would benefit from an even larger pool to confirm the results, and in

Figure 1.6: Geographical Classification of Modern Arabic Varieties in the Thesis



particular more Bedouin and rural varieties would give a more thorough overview of the variation found in Arabic.

1.2.2 A Note on Transcription

The International Phonetic Alphabet (IPA) is used to transcribe the data. This is not standard among the data sources used, nor indeed on much work written by Arabists rather than linguists, who use the America Library Association – Library of Congress system, or variants thereof, thus I have converted the disparate systems of the sources into one consistent transcription for the whole thesis. Note that both geminate consonants and long vowels are marked with ː here following IPA conventions.

Transcriptions throughout the thesis are broad rather than narrow, and much phonetic detail is therefore obscured. The consonant inventory of the modern varieties is broadly as in Figure 1.7.

Figure 1.7: IPA Chart of the Arabic Consonantal Phoneme Inventory

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyn-geal	Glottal
Plosive	b		t tˤ d dˤ				k g	q		ʔ
Nasal	m		n							
Fricative		f	θ ð ðˤ	s sˤ z	ʃ	ʒ		χ ʁ	ħ ʕ	h
Affricate					dʒ					
Trill				r						
Approx-imant	w					j				
Lateral Approx-imant					l					

At least three vowel qualities are distinguished for modern Arabic varieties [a,i,u], however some varieties also display mid vowels [e,o], and some transcriptions include a much broader array of vowels. Vowel and consonant length are contrastive in all modern Arabic varieties.

However, there are a few points on which even broad transcriptions will show major differences between varieties. Geographical variation in the realisation of the phoneme known in Classical Arabic as *jīm* is particularly notable. Classical Arabic realised *jīm* as a voiced palatal stop (Gairdner 1925; Fischer and Jastrow 1980;

Watson 1992) or a voiced palatalised velar stop (Schaade 1911; Cantineau 1960). Table 1.2 depicts its realisation in modern Arabic varieties (Watson 2002, p. 16).

Table 1.2: Realisation of *jīm* across Modern Arabic Dialects

Symbol	Description	Where Used
[ʤ]	voiced palatoalveolar affricate	most Bedouin dialects, rural Syrian, Jordanian, Palestinian, Mesopotamian, northern Yemen
[g]	voiced velar stop	Cairene, Yemeni dialects in Ta'izz and Hugariyyah
[ɟ]	voiced palatal stop	parts of Arabian Peninsula, northern Yemeni dialects, Upper Egypt, parts of Sudan
[j]	palatal glide	Syrian desert, Khuzistan, Hadramawt, Dhofar, Gulf Arabic
[ʒ]	voiced palatoalveolar fricative	much of the Levant (major cities of Beirut, Damascus, Jerusalem), most Maghrebi dialects

Note that while the table above broadly accounts for the phonetic patterns by geography, there are in fact complex sociolinguistically-governed factors involved in exact use of which allophone following urbanisation and migration, particularly in the Levant. For example, women tend to keep or adopt urban Palestinian variants, while men tend to use either MSA or Bedouin variants (Miller 2007).

In order to contextualise this work in the broader phonological theory, I will now remind the reader of the nature and importance of structure above the level of the phoneme and above the level of the word.

1.3 Beyond the Phoneme Structure: Why Syllables Matter

Syllables are useful to account for phonotactic constraints, that is, the distribution of different segments. Thus, we can account for why /pf/ is a possible sequence across syllables in English, as in *cup.ful*, but not syllable initially **pful* or syllable finally **cupf*. However, it is not that this is physiologically impossible — /pf/

is a perfectly acceptable sequence in the German *Pfeffer*, ‘pepper’. Furthermore, the syllable is valuable in the structural description of phonological rules, such as the allophones of /t/ in (1).

(1) Allophones of /t/

a. aʔ.læn.tɪc

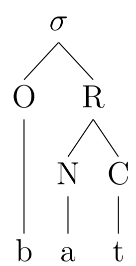
b. ə.t^{h}rəʊ.ʃəs

Although /t/ follows a vowel and precedes a liquid in both (1-a) and (1-b), syllable structure can account for the allophonic distribution, with /t/ glottalised syllable finally in (1-a) and aspirated syllable initially in (1-b). The syllable is also the target in morphological processes such as reduplication and truncation, as well as the domain of stress and tone. Finally, speech errors often obey syllable structures, substituting the same syllable position as in *beg and eacon sandwiches* instead of *egg and bacon sandwiches*, and the syllable functions as a psycholinguistic entity in ambisyllabic phonemes (Gussenhoven 1986a) and monitoring tasks (Cutler et al. 1986; Mehler et al. 1981; Aquil 2011).

Phonotactic constraints make most sense if there is some degree of internal structure to the syllable. Constituency-based models recognise three constituents: Onset (O), Nucleus (N), Coda (C). The most widely accepted model is that shown in Figure 1.8 below.

Figure 1.8: Syllable Constituency

Key: σ = Syllable, O = Onset, R = Rime, N = Nucleus, C = Coda



This model assumes the nucleus and coda form a constituent: the rime (R). This approach is assumed here, given the relative independence of the onset and rime for well-formedness conditions and that the rime is the element affecting rhyming.

Other constituency proposals include that in Figure 1.10 proposed for Japanese by Kubozono (1989), on the grounds that speech errors split Japanese syllables between the nucleus and rest of the syllable, and that in Figure 1.9 proposed by Saporta and Contreras (1962) and supported by Clements and Keyser (1983) due to co-occurence restrictions. Co-occurrence restrictions between the onset and nucleus occur for example in English, where anterior fricatives /f,v,s,z,θ,ð/ cannot occur before /uːr/, and stop plus /w/ clusters cannot occur before back vowels including /uː, ʊ, aw/ (Clements and Keyser 1983, p. 20). Similarly, co-occurence restrictions occur between the nucleus and the coda, as short vowel + CCC sequences can be found with a final coronal member, but with a long vowel only Vː+CC sequences can be found (Selkirk 1978), and as such they argue there are no grounds to favour one grouping of constituents over another.

Figure 1.9: Saporta and Contreras (1962) Syllable Model

Key: σ = Syllable, O = Onset, N = Nucleus, C = Coda

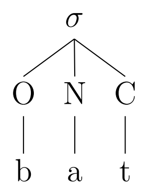
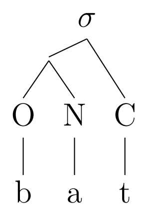


Figure 1.10: Kubozono (1989) Syllable Model

Key: σ = Syllable, O = Onset, N = Nucleus, C = Coda



How the syllable functions in Arabic is not a solved matter. There are ongoing questions about the permitted size of syllables in different varieties and how violations of syllable structure are repaired which I engage with in the course of this thesis. To understand the broader context of this work, one must be familiar with the issues of sonority discussed in Section 1.3.1, syllable weight discussed in Section 1.3.2, and superheavy syllables discussed in Section 1.3.3.

1.3.1 Sonority

The syllable can be analysed in terms of organising segments in waves of sonority (Sievers 1881; Jespersen 1904; Fischer-Jorgensen 1952; Blevins 1995). Sonority can account for restrictions on segment ordering within syllables, with sonority increasing to the nucleus and decreasing to the coda, ending often more sonorously than the onset, as encapsulated by the Sonority Sequencing Principle (Clements 1990; Selkirk 1984; Morelli 2003). There is no single phonetic correlate to sonority, but the hierarchy in (2) is widely accepted (Bell and Hooper 1978; van der Hulst 1984; Clements 1987; Clements 1990; Kenstowicz 1994; Smolensky 1995; Holt 1997).

(2) The Sonority Hierarchy
Plosives < Fricatives < Nasals < Liquids < Glides < Vowels

The vowel is the optimal nucleus, but sonorant[13] nuclei exist, as in the US English [jr̩n] 'yearn'. Some scholars identify obstruent nuclei in Nuxalk Salishan[14], such as in [sx̩s] 'animal fat'. However, this is controversial, with Newman (1947) arguing for no syllabification, Bagemihi (1991) viewing obstruents are non-syllabic, and Shaw (1993) proposing non-nuclear syllables instead.

Not only have patterns of sonority been identified within syllables, but also importantly sonority-based patterns also occur across syllable boundaries. There are cross-linguistic tendencies for sonority patterns between adjacent syllables (Hooper 1976; Murray and Vennemann 1983; Vennemann 1988). This is termed the Sonority Contact Law which states that:

> 'A syllable contact A.B is the more preferred, the greater the sonority of the offset A and the less the sonority of the onset B.

From Davis and Shin (1999, p. 286)

That is, *al.ka* with falling sonority across the syllable boundary is preferred to *ak.la* with rising sonority across the syllable boundary. This accounts for a range of synchronic and diachronic processes. For a review, see Seo (2011). This has been adapted by Alhammad (2018) to account for a preference for falling sonority between a syllable boundary and a semisyllable in Arabic.

[13]Consonants can be divided into sonorants and obstruents. Sonorants are produced with continuous non-turbulent airflow, and include nasals, liquids, and glides; obstruents are produced with obstructions to the airflow, so include plosives and fricatives.

[14]Spoken in the town Bella Coola in British Columbia, Canada.

1.3.2 Weight

Some languages treat syllables differently on account of their internal structure, as if some syllables are heavier than others (Allen 1973). Pre-generative approaches discuss this in terms of syllable structure types. However, the Prague school (Jakobson 1971) proposed the mora as a unit of weight mediating between the CV tier and the syllable (Hayes 1984; Hyman 1985). Language specific mora distribution explains why the same syllable structure has different weights in different languages (Gordon 2006).

The onset never carries a mora as it makes no contribution to stress in any known language. Vocalic nuclei always carry moras, but whether long vowels carry two moras is language dependent. The greatest variation is in whether the coda carries moras through Weight by Position (Hayes 1989). The different options for mora distribution allows for three types of languages (Hayes 1995), as shown in Figure 1.11.[15]

Figure 1.11: Languages by Mora Distribution

Type	Example	Heavy(H)	Light(L)
Rime Weight	English	CVː CVC	CV
Nucleus Weight	Selkup	CVː	CVC CV
Coda Weight	Dutch	CVC	CVː CV

Subsequent research has suggested that stress assignment in quantity sensitive languages can be more complex than a simple heavy versus light syllable distinction. Proposals include weight preference scales for Hindi (Kelkar 1968), Mehri (Watson and Al-Mahri 2018) and Sˤanʕaːniː (Watson 2002); vowel height for Nanti (Crowhurst and Michael 2005); and a sonority threshold for moraicity in Lithuanian (Zec 1995). Context-dependent weight has also been found where CVC syllables are heavy unless word final (Kager 1989; Alber et al. 1997; Rosenthall

[15]While 'H' and 'L' were used in the discussion of diglossia above, their meaning is different in phonology. As such, through the rest of this thesis 'H' will refer to heavy syllables, and 'L' to light syllables.

and van der Hulst 1999; Moren 2000). Context-dependent weight occurs in all modern Arabic varieties, commonly analysed as consonant extrametricality, though I will posit alternative extrametrical units.

In addition to accounting for some aspects of stress, the mora also best explains Compensatory Lengthening. This is the lengthening of a segment following the deletion of a neighbour to preserve weight. It only occurs when the segment has been deleted from the coda, as in (3) from the Ingvaeonic parent language of Friesian and English, which lost nasals before fricatives, triggering vowel lengthening:

(3) munθ → muːθ
'mouth'

However, when Middle English lost velar stops before nasals, changes to the onset did not trigger Compensatory Lengthening in (4).

(4) knɔt → nɔt
'knot'

Evidence for the underlying single mora of geminates can be seen in Koya (Tyler 1969) where long vowels are shortened after concatenation of geminates, but not other consonants, as shown in (5).

(5) Geminate Moracity

a. CVː + Cː → CVCː (shortening)
b. CVː + C → CVːC (no shortening)

This reflects a common cross-linguistic tendency to avoid trimoraic syllables, although Estonian (Lehiste 1966; Prince 1980), Hindi (Kelkar 1968), and Nanti (Crowhurst and Michael 2005) permit them with superheavy syllables taking precedence over heavy syllables for stress assignment.

1.3.3 Superheavy Syllables

There is cross-linguistic variation in the syllable inventories that exist, as depicted in the following Table 1.3 adapted from Zec (2007, p. 165).

Table 1.3: Cross-linguistic Syllable Inventory

(C)	C	V	(C)	(C)	Totonak
(C)	C	V	(C)		Dakota
	C	V	(C)	(C)	Klamath
	C	V	(C)		Temiar
(C)	C	V			Arabela
	C	V			Senufo
(C)	(C)	V	(C)	(C)	English
(C)	(C)	V	(C)		Spanish
	(C)	V	(C)	(C)	Finnish
	(C)	V	(C)		Turkish
(C)	(C)	V			Pirahã
	(C)	V			Fijian

Permitted syllable sizes are rarely symmetrical across the word, but rather depend on position with the word. Of particular importance to this thesis are syllables larger than CVC and where these are permitted. If we assume that every element in a coda carries a mora following Weight by Position, then it would follow that there is a three–way opposition between monomoraic CV, bimoraic CVC and CVː, and trimoraic CVːC and CVCC syllables. However, cross-linguistically three-way oppositions are relatively rare (Trubetzkoy 1939; Hayes 1995; Broselow, Chen, and Huffman 1997) and trimoraic syllables are avoided in languages such as Latin (Martinet 1955), Hausa (Hayes 1986), Japanese (Kubozono 1999), Fula (Sherer 1994), Lithuanian (Zec 1995) and Koya (Tyler 1969), as well as English and other Germanic languages (Árnason 1980). This avoidance can be seen in vowel shortening processes, epenthesis, and stray erasure of unsyllabifiable elements.

In Wolof[16], the affixation to form the inversive from the imperfective creates a closed heavy CVː through gemination which undergoes Closed Syllable Shortening

[16] A Niger-Congo languages spoken in Senegal, Gambia and Mauritania.

in (6) (Kaye et al. 1990). Note that the CVːC is acceptable at the end of the word as the final consonant is extrametrical.

(6) Closed Syllable Shortening in Wolof

a. rɔːf

‘to put in’ (imperfective)

b. rɔpːi

‘to take out’ (inversive)

c. yɛːw

‘to tie’ (imperfective)

d. yɛwːi

‘to untie’ (inversive)

In Koryak[17], the most complex syllable permitted is CVC, so if a larger syllable is created this needs to be repaired through epenthesis, as shown in (7) (Spencer 1996, p. 63–64).

(7) Epenthesis in Koryak

[təp.ŋə.lon]

t -pŋlo -n
SBJ.1SG -ask -OBJ.3SG

‘I asked him’

In Bengali[18], no coda clusters are permitted, so this is repaired in (8) through degemination (Lahiri 2001, p. 1349).

(8) Deletion in Bengali

/boʃ-ʧːi/

boʃ-ʧi
sit -PRS.IND

‘he sits’

[17] A Paleosiberian language spoken in Kamchatka.
[18] An Indo-Aryan language spoken in Bengal

Similarly in Spanish, medial CVCC syllables are not permitted so when created through affixation the /p/ is deleted in (9-b).

(9) Spanish Stray Erasure

a. eskulp-ir
   'to sculpt'
b. eskul-tura
   'sculpture'
c. *eskul**p**tura

However, some consonants do seem to persist outside of canonical CVC syllable templates, in particular at word edges, and it has been questioned whether they are truly part of the syllable, or instead part of an extrasyllabic unit of some sort. Representations of these tend to be on an 'ad hoc, language–specific basis' (Vaux and Wolfe 2009, p. 102). Reflecting the heterogeneity of representations, a wide range of terms are also used to describe them, including semisyllables, reduced syllables, the first third of sesquisyllables, presyllables, half syllables, headless syllables, degenerate syllables, consonantal syllables and minor syllables (Herr 2011, p. 35).

Vaux and Wolfe (2009, p. 104–5) classify these different types of representations as follows. *Complex Margins* involve segments associated with the syllable directly or via an extra onset node to form more complex onset or coda clusters. *Degenerate Syllables* involve a syllable with an empty nucleus and overt coda or just a coda. *Appendices* attach a segment to a prosodic node higher than the syllable whereas *stray segments* do not attach to higher level structures at all.

Different languages have different restrictions on the number of semisyllables — or equivalent — in a row. Cho and King (2003, p. 194) argue there is 'at most one semisyllable per morpheme' but this is called into question by cases such as Dutch *her-f-s-t*, 'autumn', (Trommelen 1983; van der Hulst 1984). They can occur word initially, word medially, or word finally. Word initially, they are found in Acoma (Miller 1965), Ladakhi (Koshal 1979; Bell and Saka 1983), Greek, Sanskrit and Bella

Coola (Steriade 1982). Word medially, they are found in English, with stop-fricative-stop clusters (Vaux and Wolfe 2009, p. 128). Word finally, they are found in French (Rialland 1994; Dell 1995), English and Wolof (Charette 1985; Kenstowicz 1994).

Whether or not semisyllables contain moras seems to vary cross-linguistically. Moraic semisyllables are argued for Estonian and Saami (Bye 1997), Burmese (Green 2005), and Arabic (Kiparsky 2003). Mora-less syllables work for French (Féry and van de Vijver 2003), Mon-Khmer languages (Donegan and Stampe 1983), Lemi Chin (Herr 2011), and Proto-Indo-European (Keydana 2012). A few scholars suggest both are needed, such as Shaw (1992) to account the moraicity of liquid semisyllables but amoraicity of obstruent semisyllables in Kammu.

1.4 Beyond the Word Structure: Prosodic Word Formation

The prosodic word mediates between phonology and morphology, and phonology and syntax, and roughly corresponds to the grammatical word. Evidence for its existence can be found in phenomena that use the prosodic word as their domain, including glide formation, syllabification, and stress.

Hannahs (1995, p. 1131) argues that prosodic words are the domain for glide formation in French. Underlying high vowels /i, y/ can become glides [j, ɥ], but only in stem+suffix strings such as (10-b), not at the end of a prefix as in (10-c) or start of a compound as in (10-d).

(10) French Glide Formation Hannahs (1995, p. 1131)

a. colonie
   [kɔlɔni]
   ‘colony’

b. colonial
   [kɔlɔnjal]
   ‘colonial’

c. anti-alcoolique
[ãtialkɔlik]
*[ãtjalkɔlik]
'anti-alcohol'

d. tissu-éponge
[tisyerɔ̃ʒ]
*[tisɥerɔ̃ʒ]
'terry-cloth'[19]

It is often assumed that in Germanic languages such as Dutch (Booij 1995), English (Raffelsiefen 1999) and German (Hall 1999), as well as languages such as Italian (Peperkamp 1997), and Korean (Kang 1992), that syllabification occurs within the prosodic word domain, whereas in languages such as French it occurs at the phrase level so *petit ami*, 'boyfriend', and *petit tamis*, 'little sieve', are homophonous: [pə.ti.ta.mi]. However, such a statement elides over the complexity of syllabification — aspiration, flapping, and [r] insertion occur within the phrasal domain in English, and all are closely related to syllabification. Liaison also occurs across syllable boundaries in English, with the coda of one word becoming the onset of the next, as in *fi.ve eggs* and *Ol.d English.* Many phonological processes are sensitive to syllable boundaries, and as such resyllabification can occur cyclically and postlexically. Thus, there are limitations to using syllabification and related processes as evidence of the prosodic word domain.

The prosodic word is the domain for primary stress in Arabic, as it is in many languages. In Jarawara,[20] primary stress falls on the penultimate syllable as in (11-a). In reduplicative formations such as (11-b) and in compounds such as (11-c), there are multiple primary stress peaks which correspond to the penultimate syllable of their constituents, rather than the penultimate syllable of the compound or formation itself:

(11) Jarawara Prosodic Word Formation (Dixon 2002, p. 128)

[19]Terry cloth is the material used to make towels and fluffy dressing gowns
[20]A dialect of Madí, an Arawan language spoken in Brazil

a. keˈtebe

‘run, follow’

b. ˈkete-keˈtebe

‘run a lot’

c. to -ˈwa -ka -ˈtima -ˈmaro
away -APPL -in.motion -upstream -FUT.PRF
‘took upstream’

There is a limit on the size of the minimal prosodic word. Assuming the prosodic hierarchy reflects the constituency of prosodic units, the minimal prosodic word is usually the foot (McCarthy 1993; McCarthy and Prince 1998). To maintain the minimal prosodic word under derivation, vowels can be lengthened, or consonants added. In Bengali, ‘nose’ is underlyingly *nak* but undergoes vowel lengthening to meet minimal word requirements as in (12-a), but does not need to undergo lengthening when a suffix is added in (12-b).

(12) Bengali (Lahiri and Fitzpatrick-Cole 1999, p. 128)

a. naːk

‘nose’

b. naki

‘nasal’

Prosodic word formation involves both morphological and phonological processes. Kiparsky (1982) argued that not only do you need rule ordering in general, but that different rules applied at different points in the word formation process. Kiparsky proposed Lexical Phonology with two types of rules: lexical and postlexical. Within lexical rules, there are several levels, within which rules can apply cyclically. Lexical rules apply within the lexicon, interacting with word formation rules, are conditioned by lexical information, typically have exceptions, and are subject to derived environment constraints. Postlexical rules apply within syntax and can act within words or across word boundaries, are only conditioned by phonological not morphological information, do not have exceptions and can apply within roots.

1.5 Overview and Structure of the Thesis

Modern Arabic varieties differ significantly in their phonology, grammar and lexicon. In the phonology alone, the variation includes epenthesis position, syncope application, vowel lengthening processes, vowel shortening processes, permissibility of medial and final consonant clusters, permissibility of larger than CVC medial syllables, foot type, parsing directionality, extrametricality, which elements are incorporated into the prosodic word, and exceptions to stress. How to reconcile this wealth of variation into a typology with predictive power has been attempted as will be discussed in Chapter 2, Section 2.2. These existing attempts, however, are incomplete, focusing on smaller sets of correlations, and face empirical limitations. This thesis explores not just the crucial ingredients of the phonological grammar, including syllable and foot structure, but also the lexical and postlexical nature of prosodic structure across sixteen modern Arabic varieties, ranging from Morocco to Iraq to Yemen.

I explore these issues through the following more specific research questions:

- How can we account for the distribution of syllable types across Arabic?
- What is the structure of final CVXC syllables?
- How are stress parameters distributed across Arabic varieties and how can we account for exceptions to them?
- How are linearly attached formatives incorporated into the prosodic word?

These questions allow me to develop a more comprehensive phonological typology than my predecessors that not only better accounts for the variation found across Arabic, but also has consequences for phonological theory more broadly.

In Chapter 2, I review previous work on Arabic syllable structure typology, which suggests that varieties can be classified in terms of where they epenthesise into a morphologically derived tri-consonantal cluster. I test this approach against the sixteen varieties considered here. I find that there are more medial CVːC and

CVCː syllables than predicted by Kiparsky (2003) and Watson (2007), and propose a new analysis. I argue that there are two axes for phonological variation across Arabic: Tolerance and Repair, that is, what type of syllables they tolerate and how they repair violation to syllable structure. These axes do not covary. I find that Arabic varieties differ in how they tolerate quantity. Some dialects permit medial syllables with long vowels or geminate consonants but not consonant clusters. This can be accounted for with restrictions at different levels of syllable structure. Varieties that permit long segments have restrictions on the *segmental* level, whereas varieties that do not have restrictions on the *moraic* level, and varieties that allow long consonants or clusters but not long vowels have restrictions on the *X-slot* tier. In addition to the canonical two or three segments, moras or X-slots, individual varieties can also permit an extra element in their medial syllable rime according to variety-specific restrictions, thus allowing for some restricted consonantal clusters. These restrictions can include coronality, fast speech, or sonority. As such, this typology provides mechanisms to account for the data in each and every dialect, rather than merely providing a canonical set of features that not all dialects fit.

In Chapter 3, I delve into the issue of the structure of final CVXC syllables using evidence from stress. I revisit the stress analysis of the sixteen varieties in question. All final CVXC syllables are unaffected by extrametricality, which means that the final C must be in a separate syllable that contains a catalectic mora to ensure that the C itself is non-peripheral. I demonstrate that not only does catalexis account for the behaviour of extrametricality, it also accounts for the correlation between medial epenthesis position and whether domain edge clusters are permitted. In this way, I construe catalexis not merely as a device to resolve degenerate feet, but also degenerate syllables. Previous work on Arabic syllable typology found no link with stress position. However, considering the new typology of syllable behaviours I propose in Chapter 2, I revisit this. I find that while the syllable that stress falls on does not covary with the syllable parameters (as found by Kiparsky (2003)), there is a relationship between the extrametrical unit parameter and whether varieties permit long segments (the new axis of variation

I motivate in Chapter 2). Specifically, varieties with syllable restrictions on the segmental level have syllable extrametricality, varieties with syllable restrictions on the moraic level have consonant extrametricality, and varieties with X-slot syllable restrictions do not exhibit any extrametricality.

In Chapter 4, I explore cliticization and affixation in the sixteen varieties, considering their impact on syllable and foot structure. I review previous literature on clitics, including the assumptions in the literature that particular endings are clitics (specifically, attached direct and indirect object pronouns, dative markers, and negation markers). These assumptions include low selectivity; phonological differences; optionality; and stressability. I dispute the claim that these are clitics and show that they are in fact suffixes. These endings do not have low selectivity in the sense used elsewhere in the literature. The apparent differences in phonological behaviours compared to subject markers come from paradigm behaviours in the subject forms combined with a requirement for other endings to attach after a heavy monosyllabic foot. However, I show this preference is also found in the subject markers of biliteral[21] and weak[22] verbs. This preference for attaching to a stressed foot also accounts for well–recognised but ill–understood exceptions to stress rules across Arabic varieties. In many varieties, the third person feminine singular past tense verb plus a vowel initial bound pronominal object receives stress and is never syncopated. Rather than needing exceptional reversals in parsing directionality or other solutions, this can be accounted for as the same requirement to attach after a stressed foot as other affixes. It also accounts for why final long vowels can escape syllable extrametricality: they are heteromorphemic, contain an affix, and so require stress. I show these endings attach at Level 2 in the lexical phonology. This has to still be part of phonological word formation as words with only subject markers also undergo the phonological processes on Level 2. Therefore, it is not that these processes are occurring within a clitic group level instead. Finally I

[21] The canonical Arabic root contains three consonantal segments. However some verb roots only have two, and as such are named biliteral.

[22] The canonical Arabic root contains three consonantal segments. Weak verbs are those where the final root segment is a glide.

show that it is not that there is no need for a clitic/affix distinction in Arabic: there is a clear difference between prefixes and clitics, as only prefixes impact stress assignment. In most of the varieties I explore, there are clitic tense/mood/aspect markers, definite article, and negation; however, in some varieties these are prefixes and can impact stress assignment accordingly.

In Chapter 5, I conclude the thesis, clarify the contributions to phonological theory more widely and discuss potential future historical and experimental research. I explore the limitations of a constraint-based approach to account for the interactions motivated throughout the thesis.

2

The Role of Quantity in Syllable Structure Restrictions

Contents

2.1 Introduction

Arabic varieties have previously been divided into groups based on the position of an epenthetic vowel in a morphologically derived triconsonantal cluster. In (13-a), Cairene epenthesises after the central consonant, and in (13-b) Iraqi epenthesises before the central consonant, whereas in (13-c) Moroccan does not epenthesise at all. Note that here and throughout epenthetic vowels are underlined.

(13) Epenthesis Position in a Triconsonantal Cluster (Kiparsky 2003, p. 3)

a. Cairene CCvC

ʔul -ti̲ -l -u
say.PFV -SBJ.1SG -DAT -OBJ.3SG.M
‘I said to him’

b. Iraqi Cv̲CC
gil -i̲t -l -u
say.PST -SBJ.1SG -DAT -OBJ.3SG.M
‘I said to him’

c. Moroccan CCC
qəl -t -l -u
say.PST -SBJ.1SG -DAT -OBJ.3SG.M
‘I said to him’

How to account for these syllabification patterns, and the extent to which the epenthesis position is correlated with other syllable behaviours, has been a key focus in studies of Arabic phonology to date.

In this chapter, I argue that previous typological accounts of Arabic syllable behaviour are incomplete and that a bifurcated typology, with one axis for medial syllable tolerance and another axis for syllable repair and domain edge behaviour, is required to account for the patterns discussed here. I demonstrate that varieties can be grouped by the level of structure they impose syllable restrictions on — whether segmental, moraic, or X-slot — which accounts for the variation in tolerance of medial CVXC syllables, and that this axis does not covary with the relationship between epenthesis position and whether word edge clusters are permitted. By CVXC I mean syllables where the X can be a vowel or consonant, thus subsuming CVːC, CVCː and CVCC.

In Section 2.2 I discuss previous work on Arabic syllable typology. In Section 2.3 I discuss previous work on quantity which has not been adequately considered in the Arabic phonology literature, before evaluating the extent to which the sixteen varieties I explore here fit the existing typological expectations in Section 2.4. I demonstrate in Section 2.5 that more medial CVXC syllables surface than expected, and that CVːC and CVCː syllables pattern together. Informed by previous work on quantity, I lay out a new proposal for Arabic syllable typology in Section 2.6.

2.2 Previous Work on Arabic Syllable Structure Typology

A range of syllable structure types are found in Arabic. These are summarised in Table 2.1, which includes the traditional Arab grammarian names for these syllable types.

Figure 2.1: Arabic Syllable Types with Traditional Arabic Names

Key: C = consonant, V= vowel

Light	al kasiːr al maftuːħ	CV	kataba	‘he wrote’
Heavy	al tawiːl al maftuːħ	CVː	saːlim	‘intact’
	al muɣlaq	CVC	man	‘who’
Superheavy	al muqfal	CVːC	qaːl	‘he said’
	al madiːd al muqfal bisamilain	CVCC	nasr	‘prose’
		CVCː	madː	‘he extended’

Azzabi (2001, p. 112–113)

The distribution of CV and CVC syllables is unrestricted.[1] Usually, researchers claim that CVXC is only found word or utterance finally — however, as this chapter will demonstrate, there is a much higher prevalence of CVCC, CVCː and CVːC syllables word medially than has previously been recognised.

Templatic morphological gemination is a very common process across Arabic, and there are no restrictions as to which segments can undergo gemination. Gemination

[1]Most works agree there are no onset-less syllables in Arabic. However, Azzabi (2001) claims VC syllables do exist. Such apparently onsetless syllables are often described as beginning with a glottal stop such as those in (i).

(i) a. ʔana
‘I’
b. ʔalbanja
‘Albania’

Not all scholars agree this should be considered an onset consonant (Azzajaj 1979; Ibn Jinnii 1970). The Arab grammarians distinguished between two types of glottal stops: *hamzat al-qatʕ* that is always pronounced and present in the underlying morphophonological structure; and *hamzat al-wasl* that is only pronounced before vowel-initial morphemes without a preceding consonant, with inclusion controlled by the phonological environment.

Contrary to Azzabi’s (2001) claim not to produce any *hamzat al-wasl* glottal stops, Kaye (1991) argues based on anecdotal evidence and orthography that the contextually-produced *hamzat al-wasl* is shifting to become a *hamzat al-qatʕ* — that is, prothetic glottal stops are becoming part of the underlying representation.

occurs productively to mark particular verb forms. The 10 verb templates in Modern Standard Arabic are detailed in Figure 2.2. They have general meanings, such as reflexive or causative, allowing different verbs to be formed producively from the same root with nuances of meaning. Note that not all templates are attested for all root forms or in all modern varieties, whether in Classical/ or in any modern variety. For example, Palestinian doesn't permit *nn* or *nm* clusters. The Form I verb *nisi* 'forget' does not realise the Form VII passive **innasa*, but rather has a related passive in Form VIII: *intasa* 'be forgotten' (Laks 2011).

Figure 2.2: Modern Standard Arabic Verb Forms

Verb Form	Meaning	Pattern		Examples	
		Present Tense	Past Tense		
Form I	Base Form	yaCCuC	CaCaCa	kataba	write
Form II	Causation	yuCaCCiC	CaCCaCa	kattaba	made s.o. write
	Intensity			kassara	smashed
	Estimation			kaððaba	considered s.o. a liar
Form III	Do s.thng to s.o.	yuCaaCiCu	CaaCaCa	kaataba	correspond
	Attempt to do s.thing to s.o.			qaatala	tried to kill, fought
Form IV	Causation	yuCCiCu	ʔiCCaCa	ʔaʕlama	inform
Form V	Intransitive of Form I	yataCaCCaCu	taCaCCaCa	takawwana	be composed of
Form VI	Reciprocal of III	yataCaaCiC	taCaaCaCa	takaataba	write to each other
Form VII	Reflexive of Form I	yanCaCiCu	ʔinCaCaCa	ʔinqalaba	it overturned
	Passive of Form I			ʔinkasara	it broke
Form VIII	Related to Form I	yaCtaCiCu	ʔiCtaCaCa	ʔijtamaʕa	collect (intrans)
Form IX	Change/State of defects	yaCCaCu	ʔiCCaCCa	ʔiʔwajja	became bent
	Change/State of colour			ʔihmarra	blushed
Form X	Reflexive of Form IV	yastaCaCiCu	ʔistaCaCaCa	ʔistaʕadda	prepared (intrans)
	Asking or using			ʔistahsana	approved

Note that vowel and consonant length is important for maintaining morphological distinctions, as shown in (14) and (15).

(14) Vowel Length in Morphological Distinctions

a. sˤadˤara -∅
stem.PFV -SBJ.3SG.M
'he stemmed, originated, appeared' FORM I

b. sˤaːdˤara -∅
seize.PFV -SBJ.3SG.M
'he seized, impounded, confiscated' FORM III

(15) Consonant Length in Morphological Distinctions in Algerian (Bouhadiba 1988, p. 323)

a. kadːb -u
accuse.of.lying.PFV -SBJ.3PL
'they accused someone of lying'

b. kadb -u
lie.PFV -SBJ.3PL
'they lied'

In addition to the the verb template system, geminate is used (less productively) in the formation of nouns of place and profession as in (16-a–b) as well as the plurals of lexicalised active participles in (16-c–d) (Rosenthall 2018) and in intense/habitual adjective formation in (16-e–f) (El Zarka 2018).

(16) Sources of Gemination in Arabic (Rosenthall 2018; El Zarka 2018)

a. kalːaːf

'stablehand'

b. xabːaːz

'baker'

c. baːhil (sg), buhːal (pl.)

'free'

d. ʔaːbiq (sg.), ʔabːaːq

'fugitive'

e. sikːiːr
   ‘drunken, drunkard’

f. kaðːaːb
   ‘liar’

Geminates can also occur through assimilation of the definite article /l/ to a following apical consonant (traditionally called ‘sun letters’, compared to non-assimilatory ‘moon letters’) as shown in (17-a-b). In additional to the apical consonants, in some modern colloquial varieties there is also assimilation to [k,g,b,f,m,n,j,ʤ] as in (17-c), and in West Yemenite varieties such as Jiblah the article assimilates to all consonants (Zaborski 2018) as in (17-d). Some gemination can also occur through sandhi phenomena, haplology and both regressive and progressive assimilation (Danecki 2018).

(17) Definite Article Assimilation/Gemination (Zaborski 2018)

a. /al-dars/ → [ad-dars]
   ‘the lesson’

b. /al-rajul/ → [ar-rajul]
   ‘the man’

c. /il-kitab/ → [ik-kitab]
   ‘the book’

d. /al-beːt/ → [ab-beːt]
   ‘the house’

How syllabification works in Arabic, and indeed how to account for differing patterns of epenthesis, has been an ongoing controversy. Selkirk (1981) argued that syllables are built around vowels to form CVC syllables, with unassigned consonants triggering vowel epenthesis. The position of the epenthetic vowel is determined by whether the unsyllabified consonant is assigned directly to a degenerate syllable, or to a nuclear mora of a degenerate syllable — and as such, forms the onset or rime respectively of the syllable. This can be seen in (18), where Cairene and

Iraqi adapt the loanword name ‘Fred’ according to whether the initial consonant /f/ forms an onset or coda.

(18) Onset vs Coda Dialects

a. Cairene: An Onset Dialect

fi̲.red

‘Fred’

b. Iraqi: A Coda Dialect

ʔi̲f.red

‘Fred’

Itô (1986), Itô (1989) and Farwaneh (1995) argued instead that the distinction is due to the direction of syllabification. Itô (1989) assumes the onset and nucleus share a mora, as in the model of syllable constituency proposed by Kubozono (1989). Ito assigns moras before building syllables from either edge, adding vowels where needed to form legal syllables. This is shown in Example (19), where in the second line of (19-a) and (19-b) the syllabification without epenthetic vowels is shown, before the epenthetic vowels are inserted in the third line.

(19) a. Cairene Left to Right

/ka.tab.t.lu/

katab	-ti̲	-l	-u
write.PFV	-SBJ.1SG	-DAT	-OBJ.3SG.M

‘I wrote to him’

b. Iraqi Right to Left

/ki.ta.bt.lu/

kitab	-i̲t	-l	-u
write.PFV	-SBJ.1SG	-DAT	-OBJ.3SG.M

‘I wrote to him’

An alignment constraints account of this analysis was proposed by Mester and Padgett (1994).

However, Broselow (1992, p. 25) showed that this cannot account for cases in right-to-left syllabification where an unsyllabified consonant has no consonant to serve as its onset. Directionality would predict that /klaːb/, ‘dogs’, would be syllabified as [*kilaːb] rather than the attested [ʔiklaːb]. This correct form is predicted by the onset/rime approach in that the unsyllabified consonant will always be assigned to the rime of a new syllable.

Futhermore, Broselow (1992) showed that the epenthesis patterns are correlated with other syllable-based behaviours, specifically syncope and acceptance of medial CVːC syllables, in her exploration of Cairene, Iraqi, Makkan and Sudanese. She argues that onset dialects have only ‘structure-preserving’ syncope that removes monomoraic syllables to maximise bimoraic syllables,[2] whereas rime dialects have ‘promiscuous’ syncope that removes monomoraic syllables regardless of the resulting syllabification. She also finds an asymmetry in the permissibility of CVːC versus CVCC syllables. Rime dialects permit CVːC word medially and word finally but CVCC syllables word finally only. Onset dialects do not permit either CVːC or CVCC word medially, but both word finally. These findings are summarised in Table 2.1 below.

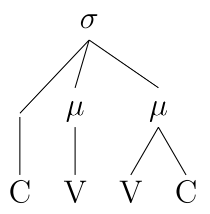
Table 2.1: Broselow (1992) Syllable Behaviours across Varieties

	Rime Dialects	Onset Dialects
Syncope Type	Promiscuous	Structure Preserving
Medial CVːC	Yes	No
Final CVːC	Yes	Yes
Medial CVCC	No	No
Final CVCC	Yes	Yes

To account for these patterns, Broselow proposes adjunction to mora, where the final C of a CVːC syllable shares the mora (μ) of the final V, as shown in Figure 2.3 below.

[2]Broselow (1976) originally argued that the maximal and optimal syllable in Arabic is bimoraic and the phonological processes conform to this constraint. This is supported by McCarthy and Prince (1990a) for Classical Arabic.

Figure 2.3: Broselow Adjunction to Mora



Broselow claims this process only supports CV:C syllables, on the grounds that a large sonority distance between the segments sharing a mora is needed, and a sufficiently large distance cannot be found in consonant clusters compared to vowel plus consonant combinations. She suggests that for rime dialects, adjunction can apply at the word and the phrase level, whereas for onset dialects it can only occur at the phrase level, with different dialects having preferences for epenthesis or delinking (and subsequent vowel shortening). Thus, Broselow extends Selkirk's analysis that rime varieties tolerate syllables with a nuclear mora dominating a consonant whereas onset varieties do not with the idea that adjunction to mora can capture the asymmetry in CV:C distribution.

Farwaneh (1995) explored 11 varieties including Levantine, Bedouin, Iraqi, Libyan, Egyptian, Sudanese and Saudi varieties, and similarly to Broselow (1992) found that other features are correlated with epenthesis, albeit Farwaneh expressed epenthesis direction as a matter of direction of syllabification. For Farwaneh, there is a continuum between Onset and Coda varieties, with Onset varieties using left-to-right syllabification, permitting no medial CVCC but some CV:C, allowing final clusters but limited removal of open syllables; whereas Coda varieties use right-to-left syllabification, permitting CV:C in all positions, some CVCC, allowing initial clusters and having open removal of open syllables.

Kiparsky (2003) extended this correlation of syllable-based phenomena with epenthesis position to cover a broader range of behaviours, and introduced a third group to cover dialects that do not break up a word-internal CCC cluster. He names these groups based on the pattern surrounding the central element of the cluster.

Onset varieties are named ‘CV’ (see (20-a)), rime varieties ‘VC’ (see (20-b)), and those that do not epenthesise are ‘C’ (see (20-c)):

(20) Three Epenthesis Types (Kiparsky 2003, p. 3)

a. Cairene: CV
ʔul -ti̠ -l -u
say.PFV -SBJ.1SG -DAT -OBJ.3SG.M
‘I said to him’

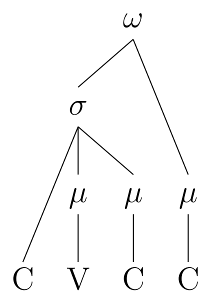
b. Iraqi: VC
gil -i̠t -l -u
say.PFV -SBJ.1SG -DAT -OBJ.3SG.M
‘I said to him’

c. Moroccan: C
qəl -t -l -u
say.PFV -SBJ.1SG -DAT -OBJ.3SG.M
‘I said to him’

The features he correlates with these groups are shown in Table 2.2.[3]

Kiparsky does consider whether stress might pattern with the groups as well, but finds that there is no correlation between epenthetic position and stressed syllable. He argues that mora licensing can account for this pattern, where unsyllabified consonants (those outside the CV, CVː, or CVC syllable template) are dominated by a mora that is not affiliated with a syllable but instead attached to the word node directly, as shown in Figure 2.4 below.

Figure 2.4: Kiparsky (2003) Adjunction to Word



[3] I have not included pausal glottalisation or desonorisation here, viewing them as areal features as argued by Watson (2007).

Table 2.2: Kiparsky (2003) Dialect Group Criteria

	CV	VC	C
-CC# CLUSTERS	permitted	not permitted /sonority restrictions	permitted
#CC- CLUSTERS	no high vowel deletion in initial position	initial vowel prosthesis	permitted
INITIAL GEMINATE CLUSTERS	vowel prosthesis	permitted	permitted
STEM METATHESIS	no	yes	drop vowel
POSTGEMINATE SYNCOPE	no	yes	yes
CLOSED SYLLABLE SHORTENING	yes	no	no
OPAQUE STRESS/EPENTHESIS INTERACTION	no	yes	no

The difference between these dialect groups is in whether they permit semi-syllables lexically or postlexically. He argues that CV dialects do not permit semisyllables; VC dialects only permit semisyllables at the lexical level; and C dialects permit semisyllables at both the lexical and postlexical levels. This is summarised in Table 2.3 below.

Table 2.3: Kiparsky (2003) Analysis

Dialect Group	Lexical	Postlexical
CV	No	No
VC	Yes	No
C	Yes	Yes

Watson (2007) noted that a set of varieties from Saudi Arabia and Yemen that appear to be CV varieties based on epenthesis after the second element in a triconsonantal cluster also display features associated with VC varieties. In particular, they permit medial CVːC syllables, permit final CVCCC and CVːCC, and delete high vowels after geminates. In fast speech, they share some features of C varieties, such as permitting initial consonant clusters and medial -CCC- clusters. Watson argues that these varieties, which include Makːan and Sˤanʕaːniː, form a fourth dialect group that she labels 'Cv'. This analysis is summarised below in Table 2.4.

Table 2.4: Watson (2007) Analysis

Dialect Group	Lexical	Postlexical	Mora Sharing
CV	No	No	Word-final CVːC syllables
VC	Yes	No	CVːC syllables
Cv	No	No	CVːC or CVCC syllable
C	Yes	Yes	none

Farwaneh (2009) links the VC/CV/C analysis with the earlier Onset/Coda distinction, arguing that VC and C varieties are Coda dialects that differ postlexically, with CV as Onset dialects. She argues that these distinctions can be accounted for within Optimality Theory as the constraint `Final-C` applying at different levels of structure. I summarise this in Table 2.5, where CODA (extreme) refers to the North African C dialects.

Table 2.5: Farwaneh (2009) Analysis: Level of `Final-C` Application

Dialect Type	Stems	Words	Syllables	Moras
ONSET	Yes	Yes		
CODA	Yes	Yes	Yes	
CODA(extreme)	Yes	Yes	Yes	Yes

There are several unresolved issues with these proposals. Mostly crucially, as I will demonstrate in the subsequent parts of this chapter, more medial CVXC syllables surface than predicted. Whilst there is some recognition in the literature of medial CVːC syllables surfacing, more CVCC syllables occur than expected, and different types of CVCC syllables behave differently.

Secondly, the formalisation of final consonant clusters in CV varieties has limitations. Both Kiparsky (2003) and Watson (2007) claim that semisyllables are impermissible in CV dialects but also recognise that final clusters do surface, such as [katabt], 'I wrote', in Cairene. Whilst domain edges do tend to exhibit different behaviours to domain internal syllables, these theories that aim to account for larger than CVC syllables should be able to account for them in all word positions. In diagrams, Kiparsky (2003) appears to assume an alternative type of non-full syllable, where a consonant is attached directly to the word node, designated by Vaux and Wolfe (2009) as an 'appendix', but no claims are made as to the nature or distribution of this appendix. Farwaneh (2009) attempts to solve this through use of constraints, but her use of `Final-C` cannot account for either the variable epenthesis in final CC clusters in coda dialects, or the distribution of initial clusters that appears to be tightly correlated with the final edge behaviours.

Thirdly, the analysis of Cv varieties from Watson (2007) misses some generalisations. Recall that Watson demonstrated that there are a set of dialects around Yemen and Saudi Arabia that appear to be CV dialects based on the epenthesis position, but share features of VC dialects (and even C in fast speech). The features in question are the medial CVːC syllables, final CVCCC and CVːCC, high vowel deletion after geminates, and, in fast speech, initial consonant clusters and medial -CCC- clusters. Watson accounts for this by describing them as a fourth dialect group, 'Cv', in which semisyllables are not permitted lexically or postlexically, and mora sharing is permitted iff the syllable rhyme contains a long segment. This analysis is limited in that it cannot account for why initial consonant clusters are permitted in fast speech. Furthermore, the behaviour described is very similar to that of C dialects, yet requires completely different analyses — C varieties permit semisyllables at both lexical and postlexical levels, whereas Cv dialects do not permit semisyllables at either level, and mora sharing is permitted iff the syllable rhyme is CVːC or CVCC. This formalisation fails to express the similarities seen between Cv and C varieties.

While it has been recognised that CVːC syllables are distributed differently to CVCC, there has been no exploration of the behaviour of CVCː syllables in these typologies. In the next section, I will review previous work on quantity, which will be integrated into my proposal for an Arabic syllable structure typology.

2.3 Previous Work on Quantity

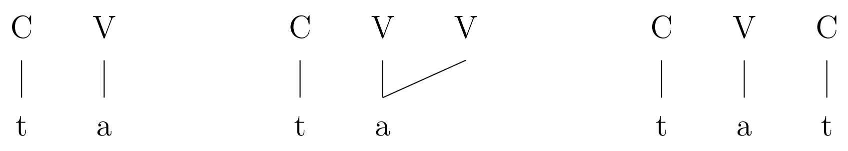
The behaviour of long vowels and geminates has been a problem in linguistics since Pāṇini. Pre-generative work recognised that long vowels and geminates vary in whether they behave as a single segment or as a sequence of segments. Trubetzkoy (1938) viewed geminates as single 'long' phonemes that were ambisyllabic intervocalically. Similarly, Swadesh (1937) viewed them as behaving like single consonants rather than clusters. However, Bithell (1952), Hockett (1955) and Brinkmann (1966) advocated for treating geminates as a sequence of segments, with possible double articulation. Chomsky and Halle (1968) formalised the featural approach as one in which geminates could be $\left[\ +\text{Long}\ \right]$ with singleton consonants as $\left[\ -\text{Long}\ \right]$, but didn't rule out a bisegmental approach for assorted phenomena. Kenstowicz (1970) found that Lithuanian treats long vowels and short vowels the same way for vowel raising, lowering and rounding processes, suggesting a featural $\left[\ +\text{Long}\ \right]$ representation, while accentual processes treat long vowels as a sequence of two vowels suggesting a bisegmental (aa) approach. Similarly, Pyle (1971) found that West Greenlandic high vowel lowering before uvulars affected long and short vowels equally, suggesting a featural $\left[\ +\text{Long}\ \right]$ representation; however, metathesis of final consonant clusters is blocked after long vowels or diphthongs, suggesting a bisegmental (aa) approach. Futhermore, Kenstowicz and Pyle (1973) found that geminate consonants behave like single consonants for some rules and clusters for others. They rejected a featural representation for length as it was not independently supported for the languages they explored. Thus, to account for the fact that geminate consonants are not split up through metathesis they propose the geminate integrity hypothesis:

> ‘All other things being equal, a rule which splits up a geminate cluster is less highly valued than a rule which must be constrained from doing so’
>
> *From Kenstowicz and Pyle (1973, p. 27)*

They recognised a broad typological split, with segmental rules (such as vowel harmony and other vowel quality alternations) treating long vowels as a single segment, but prosodic rules (such as accentual alternation, metathesis, epenthesis and syncope) treating them as a sequence of segments. However, the best way to represent these different phenomena and the relationship between the different levels of representation needed was still elusive. The development of autosegmental representations for tone (Leben 1980) and non-concatenative morphology (McCarthy 1981) formalised the relationship between these different representations. However, disagreement persists as to the nature of these representations, in whether a skeletal tier where two timing units dominate a single segment, or moraic representations where a mora dominates the segment best accounts for the data. In both approaches, there is a separation between a melodic or segmental tier containing all feature information and a more abstract tier that encodes length through timing or weight units.

Skeletal Models began with a CV tier, where timing units also encode syllable structure positions (McCarthy 1979; Leben 1980; Clements and Keyser 1983; Hayes 1986), as shown in Figure 2.5.

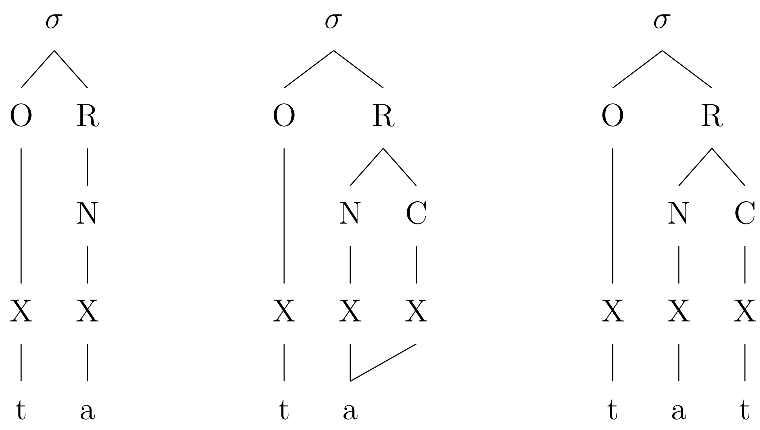
Figure 2.5: CV Theory Approach to [ta], [taː] and [tat]



Given that the consonantal or vocalic status of segments is usually predictable from syllable structure, Levin (1985) developed a model of X-Slots, as shown in Figure 2.6.

Figure 2.6: X Theory Approach to [ta], [taː] and [tat]

Key: σ = Syllable, O = Onset, R = Rime, N = Nucleus, C = Coda, X = X-slot



This model was better than the CV tier representation at accounting for compensatory lengthening, as the loss of a C in the coda led to lengthening of the preceding vowel, but could not account for the lack of lengthening in the case of onset loss. In (21-a), we see that when the Ingvaeonic parent language of Friesian and English lost nasals before fricatives, the vowel lengthened to compensate for the loss of the segment by associating to the mora of the nasal. However, in (21-b), when Middle English lost velar stops before nasals, the change to the onset did not trigger Compensatory Lengthening as there is no mora to associate to.

(21) Compensatory Lengthening

a. munθ → muːθ

'Mouth'

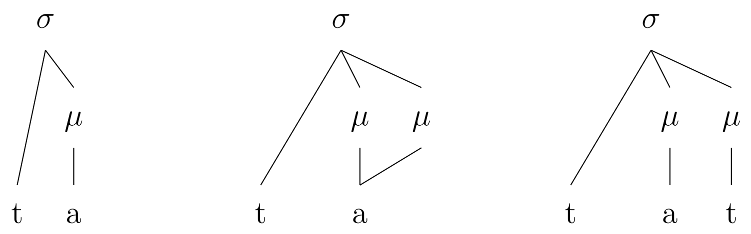
b. knɔt → nɔt

'knot'

This asymmetry is better accounted for under a moraic model, where segmental root nodes are associated to syllable nodes via moras, as shown in 2.7.

Figure 2.7: Moraic Theory Approach to [ta], [taː] and [tat] with Weight By Position

Key: σ = Syllable, μ = Mora



Here, long vowels are bimoraic, short vowels monomoraic, geminates inherently moraic, and singletons amoraic though the latter can be moraic in closed syllables via weight by position (Hayes 1989, p. 258). With moras, the lack of onset compensatory lengthening follows from the fact that the onset does not carry a mora, so there is no floating mora to associate to another segment. Thus, this approach predicts that prosodic processes affect moras, not segments, so segments that are amoraic will not be affected.

However, whilst a moraic model does work better for some phenomena, it is not sufficient for all languages. It would predict that initial geminates should be weightless as they are in the syllable onset — but in some languages they do contribute to weight, including Trukese where [tto] ‘clam (sp.)’ is an acceptable minimal word, but *[tə] ‘islet’ is not so surfaces as [təə] (Davis 1999, p.95). Rubach (1993) found that compensatory rules in Slovak need skeletal representation rather than a weight representation. Furthermore, in some languages there is no isomorphy between rhyme structure and weight (Lahiri and Koreman 1988) — for example, in Seneca Vː is light whilst VC is heavy, and in Malayalam medial geminates do not affect syllable weight while long vowels do (Mohanan and Mohanan 1984). Thus, for some researchers, a mixed model is necessary. Kraehenmann (2003) opts for a model where geminates are a single root node linked to two timing slots on an X-slot based skeletal tier, and weight is independently represented through projection of rhyme structure rather than as a function of length. This is necessary where length and weight are not interdependent — however, outside of these cases, the extra representational machinery of the X-slot is not needed.

It is clear from this that geminates and long vowels share properties and can function differently from consonantal clusters. Different rules can apply to different levels of representation, and long segments can be treated as singletons or as clusters.

Whilst Kenstowicz and Pyle (1973) predict that prosodic rules including syllabification would treat long segments as clusters, I will argue there is variation across Arabic varieties as to at what level syllable based processes apply. This variation in level of representation can account for the distribution of CVXC syllables discussed below.

I will now assess the extent to which the sixteen varieties I explore in this thesis fit the typological expectations laid out in Kiparsky's (2003) criteria, demonstrating that more CVXC syllables surface than expected.

2.4 Analysis of Sixteen Varieties

The sixteen varieties I explore can be classified as CV/VC/C based on their epenthesis position, as demonstrated in Table 2.6, where the epenthetic vowel is underlined. Note that for some varieties the epenthesis evidence has come from phrasal epenthesis due to the limitations of particular data sources.

However, many of these varieties do not exhibit the canonical behaviours expected by Kiparsky (2003) based on epenthesis position. I will briefly summarise the typological expectations discussed above before exploring the extent to which each of these varieties exhibit expected behaviours. Note that several of these terms are potentially misleading, but I have not changed because they are used in this way by Kiparsky (2003), and I do not wish to mislead the reader into finding a distinction I do not intend. The term 'stem metathesis' is potential misleading but is used here because it is used to describe the phenomenon of the -CiCC- stem in VC varieties corresponding to the -CCiC- stem found in CV varieties elsewhere in the literature. This is metathesis in the sense of medial syncope followed by epenthesis (Woidich 1980, p.212). Furthermore, note that while Kiparsky (2003) specifically refers to

Table 2.6: Epenthesis Position and Evidence

	Variety	Type	Epenthesis Evidence	Translation
1	Cairene	CV	ʔul-ti̲-l-ak	‘I told to you’
2	Sˤanʕaːniː	CV	gul-ta̲-l-ih	‘I told to him’
3	Makːan	CV	kalba̲-kum	‘your dog’
4	Iraqi	VC	kitab-i̲t-l-a	‘I wrote to him’
5	Qatari	VC	baʕad i̲ʃkibra	‘and also it is very big’
6	Muscat	VC	ʕaʃu̲b-hum	‘their grass’
7	Palestinian	VC	katab-∅-i̲l-ha	‘he wrote for her’
8	Lebanese	VC	darab-t-i̲l-na	‘I hit for her’
9	Wadi Ramː	C	ʤaːb-∅-l-na	‘he brought for us’
10	Tunisian	C	ktib-t-ha	‘I/ you (m.s.) wrote it’
11	Moroccan Casablanca	VC	faxə̲r-na	‘our coal’
12	Algerian	C	ħkəm-t-l-i	‘I caught for me’
13	Libyan Tripoli	VC	far ə̲zɣir	‘a little mouse’
14	Rufaidah	CV	gul-ti̲-l-i	‘you told me’
15	Rwaili	CV	simiʕ-ti̲-kam	‘I heard you (pl.)’
16	Haːyil	CV	ʃif-ti̲-hum	‘you saw them’

high vowel deletion after geminates, some varieties including Sˤanʕaːniː do not differentiate based on vowel quality and syncopate all vowels after geminates.

The canonical expectation of a CV variety is:

1. to permit final CC clusters;
2. to not permit initial CC clusters and as such not syncopate high vowels in domain initial syllables;
3. to prothesise a vowel before initial geminate clusters;
4. to not undergo stem metathesis;
5. to not undergo postgeminate high vowel deletion;
6. to not permit word medial CVːC syllables;
7. to not undergo opaque interactions between stress and epenthesis and thus to have stressable epenthetic vowels.

Where CV varieties are more tolerant of medial CVːC than expected, they can be classified as Cv under Watson (2007). The canonical expectation of a VC variety is:

1. to not permit final CC clusters or to do so only with sonority restrictions;
2. to permit initial CC clusters or prothesise a vowel before an initial CC cluster;
3. to permit initial geminate clusters;
4. to undergo stem metathesis from CCVC to CVCC;
5. to undergo postgeminate high vowel deletion;
6. to permit internal CVːC syllables;
7. to exhibit opaque interactions between stress and epenthesis.

The canonical expectation of a C variety is:

1. to permit final CC clusters;
2. to permit initial CC clusters;
3. to permit initial geminate clusters;
4. to syncopate the stem vowel from CCVC to CCC;
5. to undergo postgeminate high vowel deletion;
6. to permit medial CVːC syllables;
7. to not undergo opaque interactions between stress and epenthesis.

Note with the licensing of medial CVːC syllables there is a complexity sometimes overlooked in the literature. Arabic contains two separate processes both often referred to a Closed Syllable Shortening (CSS). One type, referred to hereafter as CSS-Morph, is a morphophonological alternation where in one type of Arabic verbs the central long vowel is shortened and often raised in the non-3$^{\text{rd}}$ person forms, as shown in Table 2.7 below.

Table 2.7: Arabic CSS-Morph

Verb	Translation	Verb	Translation
ʃaːf	‘he saw’	ʃuf-t	‘I saw’
qaːl	‘he said’	qul-ti	‘you (sg) said’
naːm-at	‘she slept’	nim-tum	‘you (pl) slept’
saːr-u	‘they walked’	sir-na	‘we walked’

CSS-Morph occurs in all Arabic varieties. The second type, referred to hereafter as CSS-Phon, is the vowel shortening but not raising that occurs when a CVːC syllable is too large for syllable structure restrictions. This is commonly seen in CV varieties. The exact behaviours and distributions of CSS-Morph and CSS-Phon will be discussed in Section 4.3.3 on page 286.

2.4.1 Cairene

From the epenthesis position shown in (22), we would expect Cairene to behave as a CV variety.

(22) CV Epenthesis (Kiparsky 2003, p. 3)

ʔul -ti̠ -l -ak
say.PFV -SBJ.1SG -DAT -OBJ.2SG.M

‘I told you (m.)’

Final Clusters

Cairene permits phrase final CC clusters regardless of the sonority profile of the cluster. In the following (23-a) has a sonority plateau, (23-b) has a sonority fall, and (23-c) has a sonority rise, but none undergo epenthesis.

(23) Cairene Final CC Clusters Permitted (Watson 2002, p.70, 71, 190)

a. katab -t
write.PFV -SBJ.1SG
‘I wrote’

b. bint

‘girl’

c. masˤr
‘Egypt’

This fits what we would expect from a CV variety.

Initial Clusters

Cairene does not permit phrase initial CC clusters. These are never created by syncope, and loanwords undergo epenthesis to avoid initial clusters, as shown in (24). Similarly there are no phrase initial geminate clusters.

(24) Cairene Epenthesis in Initial CC Clusters (Watson 2002, p.135)

a. bi̱.las.tik
‘plastic’

b. ku̱.lubː
‘globe’

This fits what we would expect from a CV variety.

Epenthesis in Triconsonantal Clusters

All medial triconsonantal clusters are broken up by epenthesis to the right of the central consonant, whether these clusters are word internal as in (25-a)–(25-b), or across words as in (25-c)–(25-d). Note that epenthesis occurs across words as the domain of syllabification is the phrase in Cairene.

(25) Cairene Triconsonantal Cluster Epenthesis (Watson 2002, p.64)

a. ʔul -ti̱ -l -ak
say.PFV -SBJ.1SG -DAT -OBJ.2SG.M
‘I told you’

b. kulːu̱ -hum
all -OBJ.PL.M
‘all of them’

c. kunti̱ hina
be.SBJ.1SG here
‘I were here’

d. binti̲ latˤiːfa
daughter Latifa
'Latifa's daughter'

We know that these are epenthetic vowels because they only appear in cases of triconsonantal clusters. Elsewhere, we find for example (26-a) without a following vowel, and (26-b) without such a vowel before the dative as well.

(26) Evidence of Epenthesis (Watson 2002, p. 177), (Broselow 1976, p. 16)

a. ʔul -t
say-PFV -OBJ.1SG
'I told'

b. katab -∅ -l -u
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M
'he wrote to him'

This fits what we would expect from a CV variety.

Stem Metathesis

Cairene doesn't undergo verb stem metathesis, which fits what we would expect from a CV variety.

Postgeminate High Vowel Deletion

Cairene doesn't undergo postgeminate high vowel deletion, as shown in (27).

(27) No Postgeminate High Vowel Deletion (Watson 2002, p.64)

a. kulːu̲ -hum

all -OBJ.PL.M

'all of them'

b. *kulːhum

This fits what we would expect from a CV variety.

Closed Syllable Shortening

Cairene does undergo CSS-PHON to avoid medial CVːC syllables. In (28-a) and (28-c) the bare forms of the nouns are shown with word final CVːC syllables.

However, after affixation of the possessive marker, these syllables are word medial and therefore undergo shortening as shown in (28-b) and (28-d).

(28) Closed Syllable Shortening (Watson 2002, p.66)

a. baːb
‘door’

b. bab -kum
door -POSS.2PL.M
‘your (m.pl.) door’

c. joːm
‘day’

d. jom -ha
day -POSS.3SG.F
‘her day’

This fits what we would expect from a CV variety.

Opaque Stress and Epenthesis

Cairene does not have opaque interactions between stress and epenthesis, as all epenthetic vowels are visible to stress, as shown in (29).

(29) ʔaˈbil̠ -∅ -ak
meet.PFV -SBJ.3SG.M -OBJ.2SG.M
‘he met you (m.s.)’ (Watson 2002, p.95)

This fits what we would expect from a CV variety.

Evaluation

The data above show that Cairene represents a canonical CV variety.

2.4.2 Sˤanʕaːniː

Based on the epenthesis position shown in (30), we would expect Sˤanʕaːniː to behave as a CV variety.

(30) CV Epenthesis (Watson 2002, p.65, 106)

gul -ta -l -ih
say.PFV -SBJ.1SG -DAT -OBJ.3SG.M

'I said to him'

Final Clusters

Sˤanʕaːniː does permit phrase final CC clusters regardless of sonority. In (31-a) there is a sonority plateau and in (31-b)–(31-c) there is a sonority fall.

(31) Final CC Clusters (Watson 2002, p.73)

a. ktab -t
write.PFV -SBJ.1SG
'I wrote'

b. ɟlis -t
stay.PFV -SBJ.1SG
'I stayed'

c. gambar -t
sit.PFV -SBJ.1SG
'I sat'

This fits what we would expect from a CV variety.

Initial Clusters

Sˤanʕaːniː permits phrase initial CC clusters regardless of sonority. In (32-a) there is a sonority plateau; in (32-b) there is a sonority fall; and in (32-c) there is a sonority rise.

(32) Initial CC Clusters (Watson 2002, p.73)

a. fhim -tiː
understand.PFV -SBJ.3SG.F
'you (f.s.) understood'

b. hrib -at
flee.PFV -SBJ.3SG.F
'she fled'

c. ħtˤab

'wood'

This does **not** fit what we would expect from a CV variety, but is more consistent with a Cv variety (as in Watson's (2007) analysis).

Epenthesis in Triconsonantal Clusters

In normal speech all medial triconsonantal clusters are broken up by epenthesis to the right of the central consonant, as shown in (33).

(33) Triconsonantal Cluster Epenthesis (Watson 2002, p.64)

a. baʕdama:
   'after'

b. gablama:
   'before'

c. gul -ta -l -ih
   say.PFV -SBJ.1SG -DAT -OBJ.3SG.M
   'I told him'

We know that these are epenthetic vowels because they only appear in cases of triconsonantal clusters. Elsewhere, we find for example (34-a) without a following vowel, and (34-b) without such a vowel before the dative as well.

(34) Evidence of Epenthetic Vowels (Watson 2002, p. 67)

a. gul -t
   say. -SBJ.1SG
   'I told'

b. yi -biːʕ -l -ak
   SBJ.3 -sell. -DAT -OBJ.2SG.M
   'he sells to you (m.s.)'

However, in fast speech CCC can occur, as shown in (35), where the optionally syncopated vowel is marked with brackets.

(35) Fast Speech Triconsonantal Cluster Variation (Watson 2002, p.115)

a. ji -kt(a)sib
   SBJ.3 -earn.IMPF
   'he earns'

b. ji -ft(a)hin
SBJ.3 -rest.IMPF
'he rests, feels better'

c. ift(a)han -∅
rest.PFV -SBJ.3SG.M
'he rested, felt better'

This does **not** fit what we would expect from a CV variety, but does fit a Cv variety.

Stem Metathesis

Sˤanʕaːniː does not undergo metathesis in verb stems. This fits what we would expect from a CV or Cv variety.

Postgeminate High Vowel Deletion

Sˤanʕaːniː undergoes postgeminate high vowel deletion and subsequent degemination, as shown in (36) where the first line of each example gives the underlying form of the verb. Note in (36-b) and (36-c), this deletion is postgeminate vowel deletion, rather than specifically *high* vowel deletion.

(36) Postgeminate High Vowel Deletion and Degemination (Watson 2002, p.73)

a. /jilabːisuː/
ji -labs -uː
SBJ.3 -dress.IMPF -SBJ.PL.M
'they (m.) dress (someone)'

b. /ragːaʕuː/
ragʕ-uː
sew.PFV -SBJ.3PL.M
'they (m.) sewed'

c. /jiɟamːaʕayn/
ji -ɟamʕ -ayn
SBJ.3 -collect.IMPF -SBJ.PL.F
'they (f.) collect'

This is **not** what we would expect from a CV variety, but does fit a Cv variety.

Closed Syllable Shortening

CSS-Morph occurs when triggered by the subject marker as shown in (37-a), and elsewhere CSS-Phon is optional, as shown in the variation in vowel length in (37-b) marked by brackets.

(37) Closed Syllable Shortening (Watson 2002, p.67)

a. gaːl-t → gult
‘I/you (m.s.) said’

b. ma diri(ː) -t -ʃ
NEG know.PFV -SBJ.1SG -NEG
‘I didn’t know’

Note that word internal CVːC syllables are elsewhere optionally broken up by epenthesis, as shown in (38), where the optional epenthesis is marked by brackets.

(38) Optional Epenthesis after CVːC syllables (Watson 2002, p.69)

a. kitaːb(a̲) -naː
book -POSS.1PL
‘our book’

b. ji -zuːr(a̲) -haː
SBJ.3.M -visit.IMPF -OBJ.3SG.F
‘he visits her’

c. bayn(u̲) -hum
among -OBJ.3PL
‘among them’

This is **not** what we would expect from a CV variety, but does fit a Cv variety.

Opaque Stress and Epenthesis

Sˤanʕaːniː does not undergo opaque interactions between stress and epenthesis

Evaluation

Unlike Cairene, Sˤanʕaːniː is not a canonical CV variety. While it exhibits the epenthesis behaviour of CV varieties; it also permits CVːC syllables as found in VC varieties. On this basis, Watson (2007) has analysed it as a Cv variety instead.

2.4.3 Makːan

Based on the epenthesis in the example shown in (39), we would expect Makːan to behave as a CV variety.

(39) Epenthesis in a Triconsonantal Cluster (Kabrah 2004, p. 74)
kalba -kum
dog -POSS.2PL
'your dog'

Final Clusters

Makːan permits phrase final consonantal clusters, as shown in (40).

(40) Phrase Final Clusters (Kabrah 2004, pp. 19–20)

a. bint
'girl'

b. katab -t
write.PFV -SBJ.1SG
'I wrote'

This is what we would expect from a CV variety.

Initial Clusters

Makːan does not permit phrase initial consonantal clusters, as shown in (41). Kabrah (2004) does not include data as to the presence or absence of initial geminate clusters, but these are unlikely if initial non-geminate clusters are not permitted.

(41) Phrase Initial Clusters (Kabrah 2004, p. 151)

a. kitaːb
*ktaːb
'book'

b. ti-maːr-i
*tmaːri
'she bets/challenges'

This is what we would expect from a CV variety.

Epenthesis in Triconsonantal Clusters

Makːan epenthesises in medial triconsonantal clusters to the right, as shown in (42).

(42) Epenthesis in Triconsonantal Clusters (Kabrah 2004, p. 69)

a. kalba -kum
dog -POSS.2PL
'your dog'

b. ruħ -t -alː -u
go.PFV -SBJ.1SG -DAT -OBJ.3SG.M
'I went to him'

We know that these are epenthetic vowels because they only appear in cases of triconsonantal clusters. Elsewhere, we find for example (43-a) without a following vowel, and (43-b) without such a vowel before the dative as well.

(43) Evidence of Epenthetic Vowels (Kabrah 2004, p. 23, 113)

a. ruħ -t
go.PFV -SBJ.1SG
' I went

b. katab -∅ -l -u
write. -SBJ.3SG.M -DAT -OBJ.3SG.M
'he wrote to him'

This is what we would expect from a CV variety.

Stem Metathesis

Makːan doesn't undergo stem metathesis, as shown in (44) where the CCVC stem surfaces.

(44) Stem metathesis (Kabrah 2004, p. 41)

ni -ktub
SBJ.1 -write.IMPF -SBJ.PL

'we write'

This is what we would expect from a CV variety.

Postgeminate High Vowel Deletion

Makːan doesn't undergo postgeminate high vowel deletion, as shown in (45).

(45) Postgeminate High Vowel Deletion (Kabrah 2004, p. 141)

a. mudarːis -i
teacher -POSS.1SG
'my teacher'

b. *mudarːs-i

This is what we would expect from a CV variety.

Closed Syllable Shortening

Makːan does shorten internal CVːC syllables through CSS-PHON, though these are often broken by epenthesis instead. In (46-a) we see a closed CVC syllable surface [tub], although the corresponding form in (46-b) has a long vowel [tuː]. However, in (46-c), there is no shortening, but epenthesis has occurred instead.

(46) Closed Syllable Shortening (Kabrah 2004, p. 80, 61)

a. maktub -l -u
be.written.PTCP.M -DAT -OBJ.3SG.M
'it was destined/written for him'

b. maktuːb -at -l -u
be.written.PTCP. -SBJ.F -DAT -OBJ.3SG.M
'she was destined for him'

c. baːba -ha
door -POSS.3SG.F
'her door'

However, Makːan does create some internal CVːC syllables through syncope, as shown in (47).

(47) CVːC Syllables through Syncope (Kabrah 2004, p. 136)

a. ji -saːfr -u
SBJ.3 -travel.IMPF -SBJ.PL
'they travel'

b. sˤaːħb -i
friend.m -POSS.1SG
'my male friend'

We would expect CV varieties to shorten all medial CVːC syllables, so in this regard Makːan is closer to C varieties. However, the variation in whether it is shortened is unusual.

Opaque Stress and Epenthesis

Makːan does not undergo opaque interactions between stress and epenthesis.

Evaluation

Just like Sˤanʕaːniː, Makːan is not a canonical CV variety, in that there are more CVːC syllables than expected — thus is closer to a Cv analysis as in the work of Watson (2007).

2.4.4 Iraqi

Based on the epenthesis in the example shown in (48), we would expect Iraqi to behave as a VC variety.

(48) Epenthesis in a Triconsonantal Cluster (Majdi 1988, p. 201)
raːs i̲kbiːr
head big

'a big head'

Final Clusters

Iraqi rarely permits phrase final CC clusters. In (49-a) there is a sonority fall, in (49-b)–(49-c) a sonority rise; in (50-a–b) they are glide plus consonant clusters; and in (50-c–d) they are a nasal or liquid plus consonant cluster with optional epenthesis, marked by brackets.

(49) Phrase Final CC Clusters (Majdi 1988, p. 193)

a. galu̲b

'heart'

b. baħa̲r

‘sea’

c. sahi̲l

‘easy’

(50) Phrase Final CC Clusters (Erwin 1963, pp. 32-33)

a. kawn

‘universe’

b. qaws

‘curve’

c. ban(i̲)ʤ

‘anaesthetic’

d. sil(i̲)k

‘wire’

This is what we would expect from a VC variety.

Initial Clusters

In Majdi (1988) and Broselow (1992), Iraqi does not permit phrase initial CC clusters as in (51); however, in Erwin (1963) it does as in (52).

(51) Phrase Initial CC Cluster Not Permitted (Majdi 1988, p. 126)

a. i̲qmaːs

‘cloth’

b. i̲claːb

‘dogs’

(52) Phrase Initial CC Cluster Permitted (Erwin 1963, p. 31)

a. wlaːja

‘city’

b. druːs

‘lessons’

c. gbaːl

'facing, in front of'

d. rtikab -∅
commit.PFV -SBJ.3SG.M
'he committed'

e. mfakfuk

'having taken apart'

Either variant — with and without a prothetic vowel — is what we would expect from a VC variety.

Epenthesis in Triconsonantal Clusters

Iraqi epenthesises to the left of the central consonant in a triconsonantal cluster, as shown in (53).

(53) Iraqi Triconsonantal Cluster Epenthesis(Majdi 1988, p. 195)

kitab-t-l-a
kitab -i̲t -l -a
write.PFV -SBJ.1SG -DAT -OBJ.3SG.M

'I wrote to him'

We know this is an epenthetic vowel because it does not receive stress, although in this position we would expect it to if it were an underlying vowel, as shown in (54).

(54) Opaque Stress and Epenthesis (Majdi 1988, p. 71)
kiˈtab -i̲t
write.PFV -SBJ.1SG

'I wrote'

This is what we would expect from a VC variety.

Stem Metathesis

Iraqi does undergo stem metathesis, as shown in (55) where the stem changes from CCVC to CVCC between (55-a) and (55-b).

(55) Stem Metathesis (Majdi 1988, p. 25)

a. ʔa -ktib
SBJ.1SG -write.IMPF
'I write'

b. ti -kitb -uːn
SBJ.2 -write.IMPF -SBJ.PL
'you (pl.) write'

This is what we would expect from a VC variety.

Postgeminate High Vowel Deletion

Iraqi exhibits postgeminate high vowel deletion, as shown in (56) where the vowel [a] in the final syllable of (56-a) is syncopated when no longer in the final syllable in (56-b) and (56-c).

(56) Postgeminate High Vowel Deletion (Majdi 1988, p. 19)

a. badːal -∅
change.PFV -SBJ.3SG.M
'he changed'

b. badːl -at
change.PFV -SBJ.3SG.F
'she changed'

c. t -badːl -uːn
SBJ.2 -change.IMPF -SBJ.2PL
'you (pl.) change'

This is what we would expect from a VC variety.

Closed Syllable Shortening

Iraqi does permit medial CVːC syllables, as shown in (57), so does not undergo CSS-Phon.

(57) Medial CVːC Syllables (Erwin 1963, p. 20)

beːt -kum
house -POSS.2PL

'your house'

This is what we would expect from a VC variety.

Opaque Stress and Epenthesis

Iraqi does have opaque stress and epenthesis interactions, such as in (58) where the epenthetic vowel is invisible to stress. Were it visible, we would expect initial stress given final syllable extrametricality.

(58) Opaque Stress and Epenthesis (Majdi 1988, p. 71)
ki'tab -i̲t
write.PFV -SBJ.1SG

'I wrote'

This is what we would expect from a VC variety.

Evaluation

The data above show that Iraqi is a canonical VC variety.

2.4.5 Qatari

Based on the epenthesis in the example shown in (59), we would expect Qatari to behave as a VC variety.

(59) Epenthesis in a Triconsonantal Cluster (Al-Sulaiti 1993, p. 143)
baʕad i̲ʃ -kibra -∅
and.also be -big -3SG

'and also it is very big'

Final Clusters

Qatari does permit phrase final CC clusters with sonority restrictions. Usually the clusters are sonorant–obstruent clusters or geminate obstruents. Thus in (60-a) there is a sonorant–obstruent cluster so the cluster surfaces; in (60-b)–(60-c), these are rare obstruent–obstruent clusters; but in (60-d)–(60-f) the clusters are broken up; and in (60-g) the final geminate is permitted.

(60) Phrase Final CC Cluster (Al-Sulaiti 1993, p. 107, 100, 144, 196)

a. farg
'difference'

b. is-sabt
   'Saturday'

c. sagf
   'ceiling'

d. dihi̠n
   'fat/grease'

e. rami̠l
   'sand'

f. wazi̠n
   'weight'

g. amarː
   'bitter'

This is what we expect from a VC variety.

Initial Clusters

Qatari permits non-homorganic initial CC clusters as in (61-a-c), as well as phrase initial geminate clusters as in (61-d).

(61) Phrase Initial CC Clusters (Al-Sulaiti 1993, p. 107, 154)

a. ħzaːm
   'belt'

b. sħib -∅ -a
   pull.PFV -SBJ.3SG.M -OBJ.3SG.M
   'he pulled it'

c. ktaːb -eːn
   book -DU
   'two books'

d. s -samːir
   SBJ.2SG -hammer.IMPF
   'you hammer'

This is what we expect from a VC variety.

Epenthesis in Triconsonantal Clusters

Qatari epenthesises to the right in triconsonantal clusters, as seen in (62).

(62) Epenthesis in Triconsonantal Clusters (Al-Sulaiti 1993, p. 143)
baʕad i̲ʃ-kibra
CONJ be-big.F
'and also it is very big'

This evidence comes from clusters that occur across words, as syllabification occurs on the phrasal level. Word-internal triconsonantal clusters are not found because the suffixes starting with /h/ undergo assimilation and degemination, and those with /k/ have an underlying vowel before them. There is no data for the first person suffixes in Al-Sulaiti (1993).

(63) H Assimilation (Al-Sulaiti 1993, p. 155)

a. /niʃad-t-ha/ → [niʃatːa]
'I asked her'

b. /gmaːʃ-hum/ → [gmaːʃːum]
'their pearls'

This epenthesis is what we expect from a VC variety.

Stem Metathesis

Qatari does undergo optional stem metathesis. Note in the examples in (64), V refers to a short vowel with a non-specified quality.

(64) Stem Metathesis (Al-Sulaiti 1993, p. 176)

a. j -Vʕzim ~ jʕVzim
SBJ.3SG.M -invite.IMPF
'he invites'

b. j -Vʕrif ~ jʕarf
SBJ.3SG.M -know.IMPF
'he knows'

c. j -Vħdig ~ jħVdig
SBJ.3SG.M -fish.IMPF
'he fishes'

The optionality of metathesis is unusual, but otherwise this is what we expect from a VC variety.

Postgeminate High Vowel Deletion

Qatari does not undergo postgeminate high vowel deletion, as shown in (65) where the vowel surfaces after the geminates.

(65) No Postgeminate High Vowel Deletion (Al-Sulaiti 1993, p. 241)

a. radːi̲ -∅ -ha
return.PFV -SBJ.3SG.M -OBJ.3SG.F
'he returned it'

b. kilːi̲ -hum
all -OBJ.3PL.M
'all of them'

This is **not** what we expect of a VC variety, and is more consistent with a CV variety.

Closed Syllable Shortening

Qatari does not undergo CSS-Phon. Note in (66-b) the form of the 3rd person plural possessive is underlyingly /-hum/ but has undergone assimilation. This will be discussed in Chapter 4.

(66) Closed Syllable Shortening (Al-Sulaiti 1993, p. 113, 155, 233)

a. xanaːfsiːn

'beetles'

b. gmaːʃ -ʃum
pearl.PL -POSS.3PL
'their pearls'

c. ktaːb -na
book -POSS.1PL
'our book'

This is what we would expect from a VC variety.

Opaque Stress and Epenthesis

Qatari does not undergo opaque interactions between stress and epenthesis. This is **not** what we would expect from a VC variety.

Evaluation

Qatari predominately behaves as a VC variety, sharing features with Iraqi. However, the lack of postgeminate high vowel deletion and opaque interactions between stress and epenthesis make it less canonical than Iraqi as these features are more consistent with CV varieties.

2.4.6 Muscat

Based on the epenthesis in the example shown in (67), we would expect Muscat to behave as a VC variety.

(67) Epenthesis in a Triconsonantal Cluster (Glover 1988, p. 223)
ʕaʃu̲b -hum
grass -POSS.3PL.M
'their grass'

Final Clusters

Muscat permits phrase final CC clusters in verbs (68-a), but not in nouns (68-b) unless the final C is a dental-alveolar fricative (68-c) in which case epenthesis is optional, as marked by the brackets.

(68) Phrase Final CC Cluster (Glover 1988, p. 60, 224)

a. xtar -t
choose.PFV -SBJ.1SG
'I chose'

b. ʕaʃu̲b
'grass'

c. xub(i̲)z
'bread'

This is what we would expect from a VC variety.

Initial Clusters

Muscat permits phrase initial CC clusters in (69-a-c), as well as phrase initial geminate clusters in (69-d-e), though these geminates do optionally degeminate as marked by the brackets. Note that in (69-e) the definite article is [b] due to place assimilation, as this is one of the varieties that undergoes definite article assimilation in a broader set than just the coronals.

(69) Phrase Initial CC Clusters (Glover 1988, p. 60, 68)

a. ntaras -∅
be.filled.PFV -SBJ.3SG.M
'he was filled'

b. tqahwaː -∅
have.coffee.PFV -SBJ.3SG.M
'he had coffee'

c. xtaːr -∅ -hum
choose.PFV -SBJ.3SG.M -OBJ.3PL.M
'he chose them'

d. t(ː)asˤal -∅
get.in.touch.PFV -SBJ.3SG.M
'he got in touch'

e. b -(b)eːt
DEF -house
'the house'

This is what we expect from a VC variety.

Epenthesis in Triconsonantal Clusters

Muscat displays a mix of behaviours with triconsonantal clusters. They epenthesise to the left of the central consonant in nouns as in (70-a) unless C1 and C2 are a sonorant–obstruent pair as in (70-b)–(70-c), or if C2 is a dentalveolar fricative as in (70-d) where epenthesis is optional as marked by the brackets. If the cluster contains a geminate, epenthesis does not occur as in (70-e)–(70-f).

(70) Epenthesis in Nominal Medial CCC Clusters (Glover 1988, p. 223, 224, 60)

a. ʕaʃu̲b -hum
grass -POSS.3PL.M
'their grass'

b. ward -haː
flower.PL -POSS.3SG.F
'her flowers'

c. bank -hum
bank -POSS.3PL.M
'their bank'

d. kan(i̲)z -ha
pressed.dates -POSS.3SG.F
'her pressed dates'

e. ʃaːlː -haː
carry.PTCP -OBJ.3SG.F
'carrying it (f.)'

f. maʃaːbː -haː
fan.PL - POSS.3SG.F
'her fans'

In past tense verbs, they do not epenthesise in triconsontal clusters as in (71-a), and in present tense verbs where C2 is coronal they syncopate this vowel as in (71-b)–(71-d). Syncope also occurs in fast speech regardless of the place of articulation of the C2 as in (71-e)–(71-f). The vowel that is syncopated is emboldened in the underlying form.

(71) No Epenthesis in Verbal Triconsonantal Clusters (Glover 1988, p. 223, 206, 63)

a. katab -t -ha
write.PFV -SBJ.1SG -OBJ.3SG.F
'I wrote it'

b. /jintiris/
ji -ntris
SBJ.3 -be.filled.IMPF
'it is filled'

c. /jindaxil/
ji -ndxil
SBJ.3 -be.entered.IMPF
'it is/can be entered'

d. /jistamarː/
ji -stmarː
SBJ.3 -continue.IMPF
'he continues'

e. /targamu/
targm -u
translate.PFV -SBJ.3PL.M
'they translated'

f. /jistaʕmalu/
ji -staʕml -u
SBJ.3 -use.PFV -SBJ.PL.M
'they use'

This mixed behaviour is **not** what we would expect from a VC variety. The verbal cases shown in (71) are more consistent with a C variety, but the nominal cases are not.

Stem Metathesis

Muscat does undergo stem metathesis, as shown in (72).

(72) Stem Metathesis (Glover 1988, p. 122)

a. glis -∅
sit.PFV -SBJ.3SG.M
'he sat'

b. gils -u
sit.PFV -SBJ.3PL.M
'they sat'

This is what we would expect from a VC variety.

Postgeminate High Vowel Deletion

Muscat optionally undergoes postgeminate high vowel deletion, and when it does the geminate optionally degeminates, as shown in (73) where the variants are in (73-b) and (73-c). Note that Shaaban (1977) views degemination in coastal Omani as obligatory.

(73) Postgeminate High Vowel Deletion (Glover 1988, p. 67)

a. ji -tkaːam -u
SBJ.3.M -talk.IMPF -SBJ.PL
'they m. talk'

b. jitkaːmu

c. jitkalmu

This is what we would expect from a VC variety.

Closed Syllable Shortening

Muscat does not undergo CSS-PHON. Thus, in (74-a) we see shortening of the verb stem compared to (74-b) as a case of CSS-MORPH, but in (74-c)–(74-d) the long vowel is not shortened.

(74) Closed Syllable Shortening (Glover 1988, p. 60, 135)

a. xtar -t
choose.PFV -SBJ.1SG
'I chose'

b. xtaːr -∅ -hum
choose.PFV -SBJ.3SG.M -OBJ.3PL.M
'he chose them'

c. goːhra

'jewel'

d. qandiːl -na
lantern -POSS.1PL
'our lantern'

This is what we would expect from a VC variety.

Opaque Stress and Epenthesis

Stress is postlexical in Muscat so it does not undergo opaque interactions between stress and epenthesis.

Evaluation

The data above show that Muscat appears to be a VC variety, sharing features with Iraqi and Qatari discussed above, but permits more internal CVCC syllables than expected.

2.4.7 Palestinian

From the epenthesis position shown in (75), we would expect Palestinian to behave as a VC variety.

(75) VC Epenthesis (Abu-Salim 1982a, p. 216)

katab	-∅	-i̲l	-ha
write.PFV	-SBJ.3SG.M	-DAT	-OBJ.3SG.F

'he wrote for her'

Final CC Clusters

Palestinian permits phrase final CC clusters with sonority and lexical restrictions. Where there is a rise in sonority in an underlying final CC cluster, epenthesis occurs, as shown in (76). These vowels must be epenthetic as they do not receive stress in Palestinian when followed by an affix (cf. [ˈtami̲r-ha], 'her dates'.) If this vowel was part of the underlying representation we would expect it to be stressed.

(76) Palestinian Epenthesis in Final Cluster Sonority Rise (Abu-Salim 1982a, p. 200)

a. tami̲r
   'dates'
b. ʤahi̲l
   'ignorance'
c. ʔibi̲n
   'son'

We know these are epenthetic vowels because they are not found in forms with an vowel-initial affix, as in (77). Furthermore, when followed by a consonant initial affix as in (77-c), they are not stressed, although we would expect them to be in this position.

(77) Epenthetic Vowel Evidence

a. ˈʔibn -i

'my son'

b. ˈʔibn -ak

'your (m.s.) son'

c. ˈʔibin -kum

'your (m.pl.) son'

Where the underlying cluster has a fall in sonority, there is variation in the behaviour seen. In some words, the final cluster surfaces, as in (78). Although (78-b) is a loanword, it behaves the same way as 'native' words.

(78) Palestinian No Epenthesis in Final Cluster Sonority Fall (Abu-Salim 1982a, p. 203)

a. ʤanb

'side'

b. bank

'bank'

c. ʔuxt

'sister'

However, there are cases where the final two consonants have a fall in sonority or a plateau, but still undergo epenthesis to break up the cluster, as shown in (79) and (80).

(79) Palestinian Epenthesis in Sonority Fall Cluster (Abu-Salim 1982a, p. 200)

a. kizi̠b

'lying'

b. kari̠m

'orchard'

(80) Palestinian Epenthesis in Sonority Plateau Cluster (Abu-Salim 1982a, p. 201)

a. ʕabi̱d
'slave'

b. sami̱n
'shortening'

Futhermore, there are case of free variation in whether epenthesis occurs, as shown in (81). The optionality of the epenthetic vowel is marked by the use of brackets here.

(81) Palestinian Free Variation in Final Cluster Epenthesis (Abu-Salim 1982a, p. 202)

a. dar(i̱)s
'lesson'

b. kal(i̱)b
'dog'

c. far(i̱)d
'pistol'

These restrictions on final CC clusters fit what we would expect from a VC variety.

Initial Clusters

Palestinian also permits phrase initial CC clusters, as shown in (82). There are no cases of phrase initial geminate clusters in Abu-Salim (1982a).

(82) Palestinian Initial Clusters (Abu-Salim 1982a, p. 13)

a. stalam -∅
receive.PFV -SBJ.3SG.M
'he received'

b. ktaːb -i
book -POSS.1SG
'my book'

c. mfakː
'screwdriver'

This fits what we would expect from a VC variety.

Epenthesis in Triconsonantal Clusters

Palestinian epenthesises to the left of the central consonant in a triconsonantal cluster, as shown in (83).

(83) Palestinian Epenthesis in Triconsonantal Cluster (Abu-Salim 1982a, p. 216)

katab	-∅	-i̲l	-ha
write.PFV	-SBJ.3SG.M	-DAT	-OBJ.3SG.F

‘he wrote for her’

We know this is an epenthetic vowel, as it is not found in (84) where there is no triconsonantal cluster.

(84) Epenthesis Evidence (Abu-Salim 1982a, p. 196)

katab	-∅	-l	-u
write.PFV	-SBJ.3SG.M	-DAT	-OBJ.3SG.M

‘he wrote for him’

However, if the second consonant in the triconsonantal cluster is a coronal consonant, epenthesis does not occur as Kiparsky would suggest. Rather, in these cases Palestinian permits the word internal CVCC syllable. This occurs whether the cluster is morphologically derived through affixation, as in (85-a)–(85-b), or is part of the underlying form of the word as in (85-c)–(85-d). Note that in (85-a)–(85-b) the first line of the example shows the syllable divisions, before the formative structure, gloss, and translation in the following lines.

(85) Palestinian Permitted Medial CVCC Syllables (Abu-Salim 1982a, p. 60, 22)

a. ka.tabt.lak

katab	-t	-l	-ak
write.PFV	-SBJ.1SG	-DAT	-OBJ.2SG

‘I wrote to you’

b. ?uxt.hum

?uxt	-hum
sister	-POSS.3PL

‘their sister’

c. burd.ʔaːn
‘oranges’

d. mift.xir
‘proud’

This permissibility of medial CVCC syllables if the final C is coronal is **not** what we would expect from a VC variety.

Stem Metathesis

There is no data in Abu-Salim (1982a) indicating whether Palestinian undergoes metathesis in verb stems.

Postgeminate High Vowel Deletion

Palestinian does undergo postgeminate high vowel deletion where the vowel is not syllable final. In (86-a) and (86-c) the forms where the postgeminate vowel is in the word final syllable are shown. In (86-b) and (86-d) the underlying form with the vowels is shown, before the syncopated forms, glosses and translations.

(86) Palestinian Postgeminate High Vowel Deletion (Abu-Salim 1982a, p. 165-6)

a. j -darːis
SBJ.3 -teach.IMPF
‘he teaches’

b. /jdarːisu/
j -darːs -u
SBJ.3 -teach.IMPF -SBJ.PL
‘they teach’

c. mʕalːim
‘teacher’

d. /mʕalːim-iːn/
mʕalːm -iːn
teacher(M) -PL.M
‘teachers (m.)’

This fits what we would expect from a VC variety.

Closed Syllable Shortening

Palestinian undergoes CSS-Morph in verbs (87-a)–(87-b), as well as CSS-Phon with the dative (87-c) and negation suffixes (87-d).

(87) Palestinian Closed Syllable Shortening (Abu-Salim 1982a, p. 206, 156, 149)

a. ʤaːb -∅
bring.PFV -SBJ.3SG.M
'he brought'

b. ʤib -na
bring.PFV -SBJ.1PL
'we brought'

c. /ʤaːblk/
ʤab -∅ -l -ik
bring.PFV -SBJ.3SG.M -DAT -OBJ.2SG
'he brought for you'

d. /ʤaːbʃ/
ʤab -∅ -ʃ
bring.PFV -SBJ.3SG.M -NEG
'he didn't bring'

However, outside of these cases, word medial CVːC syllables surface as shown in (88).

(88) Palestinian Word Medial CVːC Syllables (Abu-Salim 1982a, p. 150, 22, 26)

a. ʃaːf -∅ -ni
see.PFV -SBJ.3SG.M -OBJ.1SG
'he saw me'

b. ʃubːaːk -hum
window -POSS.3PL.M
'their (m.) window'

c. naːʤħ-a
successful -F
'successful (f.sg.)'

CSS-Phon occurs to repair syllables larger than CVːC. The absence of CSS-Phon fits what we would expect from a VC variety.

Opaque Stress and Epenthesis

Palestinian does undergo opaque interactions between stress and epenthesis. In (89-a) and (89-c), the attested surface forms with their translations and glosses are shown, whereas in (89-b) and (89-d) the corresponding incorrect stress placements based on surface structure are shown. This has been accounted for by Brame (1974) and others as that epenthesis occurs after stress assignment as epenthetic vowels do not seem visible for stress.

(89) Opaque Stress (Abu-Salim 1982a, p. 197, 216)

a. kaˈtab -i̲t
write.PFV -SBJ.1SG
'I wrote'

b. *ˈkatabit

c. kaˈtab -∅ -i̲l -ha
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F
'he wrote to/for her'

d. *kataˈb-i̲l-ha

This fits what we would expect from a VC variety.

Evaluation

Just as Muscat and Qatari are not fully canonically VC varieties, the data above show that while Palestinian does exhibit many of the canonical behaviours of a VC variety, it permits more medial CVCC syllables than previous work predicts.

2.4.8 Lebanese

Based on the epenthesis in the example shown in (90), we would expect Lebanese to behave as a VC variety.

(90) Epenthesis in a Triconsonantal Cluster (Haddad 1984, p. 17)
dˤarab -t -i̲l -ha
hit.PFV -SBJ.1SG -DAT -OBJ.3SG.F

'I hit for her'

Final Clusters

Lebanese only permits phrase final CC clusters if they have falling sonority. This is shown in (91), where in (91-a)–(91-c) sonority falls so the cluster surfaces; but in (91-d)–(91-f) sonority rises so an epenthetic vowel is inserted.

(91) Phrase Final CC Cluster Sonority Restrictions (Haddad 1984, p. 45, 39, 40)

a. darb
‘road’

b. xawf
‘fear’

c. nimr
‘tiger’

d. nasil
‘progeny’

e. nisir
‘eagle’

f. nidir
‘vow’

This is what we would expect from a VC variety.

Initial Clusters

Lebanese permits phrase initial CC and geminate clusters. This is shown in (92), where (92-c-d) show heteromorphemic geminate clusters. Although (92-d) is a loanword, it still behaves as expected.

(92) Phrase Initial CC Clusters (Haddad 1984, p. 26, 165, 177)

a. nzil -t
descend.PFV -SBJ.1SG
‘I descended’

b. mraːsil
‘correspondent’

c. tˤ -tˤaːwle
DEF -table
‘the table’

d. s -sinama
DEF -cinema
‘the cinema’

This is what we would expect from a VC variety.

Epenthesis in Triconsonantal Clusters

Lebanese epenthesises to the left of the central consonant in a triconsonantal cluster, as shown in (93). These must be epenthetic vowels as they are invisible to stress.

(93) Epenthesis in Medial CCC Clusters (Haddad 1984, p. 49, 17)

a. sak:ar -∅ -i̲l -na
close.PFV -SBJ.3SG.M -DAT -OBJ.1PL
‘he closed for us’

b. ħafar -∅ -i̲l -na
dig.PFV -SBJ.3SG.M -DAT -OBJ.1PL
‘he dug for us’

c. dˤarab -t -i̲l -ha
hit.PFV -SBJ.1SG -DAT -OBJ.1PL
‘I hit for her’

However, if the central consonant is a coronal consonant, it does not epenthesise. This can be seen in (94).

(94) Permitted medial CVCC syllables if C2 is coronal(Haddad 1984, p. 10)

a. mistɣil: -ni
take.advantage.of.PTCP.M -OBJ.1SG
‘he took advantage of me’

b. mistʕid:
ready.PTCP.M
‘he was ready’

This broadly follows what we would expect of a VC variety, bar the permitting of medial CVCC syllables if they end in a coronal consonant.

Stem Metathesis

Lebanese does undergo metathesis in verb stems, as shown in (95).

(95) Stem metathesis (Haddad 1984, p. 56)
ji -kitb -u
SBJ.3 -write.IMPF -SBJ.PL
'they write'

This is what we would expect from a VC variety.

Postgeminate High Vowel Deletion

There is no evidence of postgeminate high vowel deletion in Haddad (1984).

Closed Syllable Shortening

Lebanese does not undergo CSS-PHON, as shown by the surface CVːC syllables in (96).

(96) Surface CVːC Syllables (Haddad 1984, p. 10)

a. ʕaːmle
'the working class'

b. maʕaːmli
'employee'

This is what we would expect from a VC variety.

Opaque Stress and Epenthesis

Lebanese does have opaque stress, as epenthetic vowels are not stressable, as seen in (97).

(97) Opaque Stress (Haddad 1984, p. 29)

ˈʔisi̠m -na
name -POSS.1PL
'our name'

This is what we expect from VC varieties.

Evaluation

Lebanese is similar to Palestinian. The data above show that Lebanese does behave like a canonical VC variety with the exception of permitting medial CVCC syllables if they end in a coronal consonant.

2.4.9 Wadi Ramː

Based on the lack of epenthesis in the example shown in (98), we would expect Wadi Ramː to behave as a C variety.

(98) Epenthesis in a Triconsonantal Cluster (Al Mashaqba 2015, p. 120)
ʤaːb -∅ -l -na
bring.PFV -SBJ.3SG.M -DAT -OBJ.1PL
'he brought for us'

Final Clusters

Wadi Ramː permits phrase final consonantal clusters as in (99-a)–(99-d), though often obstruent-sonorant clusters are broken up through epenthesis as in (99-e)–(99-f).

(99) Phrase Final CC Clusters (Al Mashaqba 2015, p. 108, 111)

a. ʕagal -t
tie.PFV -SBJ.1SG
'I tied'

b. maɣar -t
mix.PFV -SBJ.1SG
'I mixed (something)'

c. gitˤam -t
broke.PFV -SBJ.1SG
'I broke'

d. ta -ħalb
SBJ.SG.M -milk.IMPF
'you (m.s.) milk'

e. wasi̱m
'brand'

f. ʃaʕa̱r
'hair'

This is broadly what we would expect from a C variety, but the sonority restrictions on obstruent-sonorant clusters are more consistent with VC varieties.

Initial Clusters

Wadi Ramː permits phrase initial consonantal clusters regardless of rising sonority as in (100-a)–(100-b) or falling sonority as in (100-c)–(100-d), though sonorant–obstruent clusters are optionally proceeded by a prothetic glottal stop and vowel in (100-e)–(100-f), with optionality marked by brackets. There is no evidence of initial geminate clusters in Al Mashaqba (2015).

(100) Phrase Initial Consonantal Cluster (Al Mashaqba 2015, p. 106, 107, 110)

a. draːs
   ‘harvest’

b. mraːy
   ‘mirror’

c. sbig -at
   forerun.PFV -SBJ.3SG.F
   ‘she foreran’

d. ʤbaːl
   ‘mountains’

e. (ʔi)mʕaʃːaːh
   ‘pastures’

f. (ʔi)mʕaruːf
   ‘known’

This is what we would expect from a C variety.

Epenthesis in Triconsonantal Clusters

Wadi Ramː permits CCC clusters and does not epenthesise into them, regardless of morphological complexity, as shown in (101).

(101) Surface Medial CCC Clusters (Al Mashaqba 2015, p. 110, 120, 208,)

a. ta -skn -i
SBJ.3SG.F -live.in.IMPF -SBJ.F
'you (f.s.) live in'

b. ta -nks -uw
SBJ.3SG.F -return.back.IMPF -SBJ.PL
'you (m.pl.) return back'

c. ħurmt -ah
wife -POSS.3SG.M
'his wife'

d. ʤaːb -∅ -l -na
bring.PFV -SBJ.3SG.M -DAT -OBJ.1PL
'he brought for us'

e. ʤib -t -l -ak
bring.PFV -SBJ.1SG -DAT -OBJ.1PL
'I brought you'

This is what we would expect from a C variety.

Stem Metathesis

Wadi Ramː does undergo stem metathesis, as shown in (102) where the stem shape differs between CCVC in (102-a) and CVCC in (102-b).

(102) Stem Metathesis (Al Mashaqba 2015, p. 157)

a. ʔa -ktib
SBJ.1SG -write.IMPF
'I write'

b. ja -kitb -u(w)
SBJ.3.M -write.IMPF -SBJ.PL
'they write'

This is **not** what we would expect from a C variety and is more consistent with a VC variety.

Postgeminate High Vowel Deletion

Wadi Ramː does undergo postgeminate high vowel deletion, as shown in (103).

(103) Postgeminate High Vowel Deletion (Al Mashaqba 2015, p. 122)

a. mʕalːim
‘teacher (m.’)

b. mʕalːm -iːn
teacher -PL.M
‘teachers (m.)’

c. mʕalːm -aːt
teacher -PL.F
‘teachers (f.)’

This is **not** what we would expect from a C variety and is more consistent with a VC variety.

Closed Syllable Shortening

Wadi Ramː does not undergo CSS-PHON, and permits CVːC syllables medially as shown in (104).

(104) No Closed Syllable Shortening (Al Mashaqba 2015, p. 117)
maːskiːn
‘catching (m.pl)’

This is what we would expect from a C variety.

Opaque Stress and Epenthesis

Wadi Ramː does not undergo opaque interactions between stress and epenthesis as epenthetic vowels are visible for stress. Note that unlike other varieties discussed here Wadi Ramː has iambic not trochaic feet.

(105) Stressable Epenthetic Vowels (Al Mashaqba 2015, p. 131)

a. /dihn/ → diˈhi̲n
‘fat’

b. /sˤaxl/ → saˈxa̲l
‘kid’

c. /ʤubn/ → ʤuˈbu̲n
‘coward’

Evaluation

Wadi Ramː behaves not like a canonical C variety, but instead like a VC variety that permits medial CVCC syllables. This is similar to Palestinian, Lebanese and Muscat discussed above, except in that Wadi Ramː places fewer restrictions on the type of consonants that can be present in a medial cluster. Overall it is closer in many features to the VC varieties discussed above than the C varieties this discussion now turns to.

2.4.10 Tunisian

Based on the lack of epenthesis in the triconsonantal cluster shown in (106), we would expect Tunisian to behave as a C variety.

(106) C Variety No Epenthesis (Maamouri 1967, p. 15)

ktib	-t	-ha
write.PFV	-SBJ.1SG	-OBJ.3SG.F

'I wrote it'

Final Clusters

Tunisian permits phrase final CC clusters, regardless of sonority. In (107-a) there is a sonority fall, in (107-b) there is a sonority rise, and in (107-c) there is a sonority plateau.

(107) Phrase Final CC Clusters (Maamouri 1967, p. 151, 116, 119)

a. kinz
   'treasure'
b. dihn
   'paint'
c. maqt
   'rage'

This fits what we would expect from a C variety.

Initial Clusters

Tunisian permits phrase initial CC and geminate clusters regardless of sonority. In (108-a) there is a sonority plateau, in (108-b) there is a sonority rise, and in (108-c-d) there is a sonority fall. Geminate clusters created through assimilation are shown in (109).

(108) Phrase Initial Consonant Clusters (Maamouri 1967, p. 15, 159, 151)

a. ktib -t -ha
write.PFV -SBJ.1SG -OBJ.3SG.F
'I wrote it (f.)'

b. ʃriː -tu
buy.PFV -SBJ.3PL
'you (pl.) bought'

c. mxubːi

'hidden'

(109) Phrase Initial Geminate Clusters (Maamouri 1967, p. 145, 149, 28)

a. tˤ -tˤiːr
SBJ.3SG.F -fly.IMPF
'she flies'

b. s -sriːr
DEF -bed
'the bed'

c. n -nuːr
DEF -light
'the light'

This fits what we would expect from a C variety.

Triconsonantal Cluster Epenthesis

Tunisian does not break up medial CCC clusters with epenthesis, regardless of morphological constituency or sonority. In (110-a) the cluster is heteromorphemic, whereas in (110-b)–(110-e) it is tautomorphemic. In (110-a)–(110-b) sonority falls over the first two consonants of the cluster; in (110-c)–(110-d) sonority rises; and in (110-e) there is a sonority plateau.

(110) Surface Medial CCC Clusters (Maamouri 1967, p. 15, 103, 22)

a. kalːam -t -hum
talk.PFV -SBJ.1SG -OBJ.3PL.M
'I talked to them'

b. birdgaːn
'oranges'

c. sukmt -ik
cold -POSS.2SG.M
'your cold'

d. mɣarft -u
spoon -POSS.3SG.M
'his spoon

e. maktba
'library'

However, if the cluster consists of a phrase initial definite article before a word with a initial CC cluster, epenthesis does occur, and the quality of the epenthetic vowel varies as shown in (111). Note the quality of the epenthetic vowel is merely related to the initial consonant of the noun so is inconsequential for the analyses here. We know this must be an epenthetic vowel as it is not found always with the definite article, as shown in (111-a), and without the article these nouns are found without an initial vowel.

(111) Epenthesis in Definite Article Triconsonantal Clusters (Maamouri 1967, p. 149)

a. l-maktba
'the library'

b. li-krumb
'the cabbage'

c. li-xdiːma
'the maid'

d. lu-wraqː
'the paper'

This fits what we would expect from a C variety. While there is epenthesis in the clusters with the definite article shown in (111), these are initial triconsonantal clusters, which are rarely tolerated in any language.

Stem Metathesis

Tunisian doesn't undergo stem metathesis, but syncopates the stem vowel, as is expected of a C variety. This is shown in (112), where in CV varieties we would expect [jakfiru] and in some VC varieties [jakifru].

(112) Stem Vowel Syncopation (Maamouri 1967, p. 115)
ja -kfr -u
SBJ.3 -blaspheme.IMPF -SBJ.PL.M
'they blaspheme'

Postgeminate High Vowel Deletion

Tunisian does exhibit postgeminate high vowel deletion. This is shown in (113), whereas in CV varieties we would expect a vowel after the geminate.

(113) Postgeminate High Vowel Deletion (Maamouri 1967, p. 163, 10)

a. faːrq -it
distribute.PFV -SBJ.3SG.F
'she distributed'

b. laʕːb -∅ -u
make.play.PFV -SBJ.3SG.M -OBJ.3SG.M
'he made him play'

This fits what we would expect from a C variety.

Closed Syllable Shortening

Tunisian only undergoes CSS-PHON if there is a long vowel that does not receive stress. It will also shorten long vowels in open syllables that are not stressed. This is shown in the following, where CVːC syllables are permitted if they are stressed, as in (114-a)–(114-b), but if they are unstressed as in (114-c) the vowel is shortened. Similarly, this occurs with long vowels in open syllables, where if they are stressed as in (114-d) they surface, but if they are not stressed as in (114-e) they are shortened.

(114) Closed Syllable Shortening (Maamouri 1967, p. 160, 166, 126, 159)

a. tˤaˈhuːnt -ik
mill -POSS.2SG.M
'your mill'

b. ˈʒiːb
bring.IMP.M
'bring!'

c. ʒib -ˈha -l -i
bring.IMP.M -OBJ.3SG.F -DAT -OBJ.1SG
'bring it to me'

d. ˈmsaːmir
nails

e. msaˈmir -kum
nail.PL -POSS.2PL.M
'your (pl.) nails'

That medial CVːC syllables are permitted is what we would expect from a C variety, however, that only stressed CVːC but not unstressed CVːC may surface is **not** a canonical behaviour

Opaque Stress and Epenthesis

Tunisian does not undergo opaque interactions between stress and epenthesis. This is what we would expect from a C variety.

Evaluation

The data above show that Tunisian displays the canonical behaviour of C varieties, with the exception of only permitting stressed CVː(C) syllables.

2.4.11 Algerian

Based on the lack of epenthesis in the example shown in (115), we would expect Algerian to behave as a C variety.

(115) No Epenthesis in a Triconsonantal Cluster (Bouhadiba 1988, p. 270)

ħkəm -t -l -i
catch.PFV -SBJ.2SG.M -DAT -OBJ.1SG

'You (m.s.) caught for me'

Final Clusters

Algerian permits phrase final CC clusters, regardless of sonority. In (116) they exhibit falling sonority; in (117) they exhibit a sonority plateau; and in 118 they exhibit rising sonority.

(116) Phrase Final CC Clusters Falling Sonority (Bouhadiba 1988, p. 110, 197, 311)

a. razq
'possession'

b. bənt
'girl'

c. kəlb
'dog'

(117) Phrase Final CC Clusters Sonority Plateau (Bouhadiba 1988, p. 183, 311)

a. waqt
'time'

b. sabt
'Saturday'

(118) Phrase Final CC Clusters Rising Sonority (Bouhadiba 1988, p. 110-111)

a. ħabs
'jail'

b. qatl
'killing'

c. qutr
'territory'

This is what we expect from a C variety.

Initial Clusters

Algerian permits phrase initial CC clusters as in (119-a)–(119-e) as well as initial geminate clusters as in (119-f)–(119-h). This occurs in cases of sonority plateau ((119-a)), sonority fall as in (119-b)–(119-d), and sonority fall-and-rise as in (119-e).

(119) Initial CC Clusters (Bouhadiba 1988, p. 16, 74, 383, 91, 173)

a. ktab -∅
write.PFV -SBJ.3SG.M
'he wrote'

b. nzuːm
'stars'

c. nɣlab -t
be.defeated.PFV -SBJ.1SG
'I was defeated'

d. mhatraf
dream.PTCP.M
'being dreamt (m.)'

e. ndfan -∅
be.buried.PFV -SBJ.3SG.M
'he was buried'

f. mːa
'Mother'

g. dːa -∅
take.PFV -SBJ.3SG.M
'he took'

h. dːan -∅
call.to.prayer.PFV -SBJ.3SG.M
'he called to prayer'

This is what we expect from a C variety.

Epenthesis in Triconsonantal Clusters

Algerian does not epenthesise into triconsonantal clusters, whether morphologically derived as in (120-a)–(120-b), or underlying as in (120-c)–(120-d).

(120) Medial Triconsonantal Clusters (Bouhadiba 1988, p. 270, 311–312)

a. ħkəm -t -l -i
catch.PFV -SBJ.2SG.M -DAT -OBJ.1SG
'You (m.s.) caught for me'

b. malk -na
property -POSS.1PL
'our property'

c. sansla
'chain'

d. qanqra
'toad'

This is what we expect from a C variety.

Stem Metathesis

Algerian does undergo stem metathesis, as shown in (121). Note there is no significance to the variation in epenthetic vowel here.

(121) Stem metathesis (Bouhadiba 1988, p. 184, 294)

a. /t-rgusˤ-i/
to -roksˤ -i
SBJ.3.F -dance.IMPF -SBJ.F
'you (f.s.) dance'

b. /hatraf-u/
haturf -∅ -u
dream.PFV -SBJ.3SG.M -OBJ.3SG.M
'he dreamt it'

This is **not** what we expect from a C variety. The stem metathesis is more consistent with a VC variety.

Postgeminate High Vowel Deletion

Algerian undergoes postgeminate vowel deletion, including of low vowels as seen in (122).

(122) Postgeminate Vowel Deletion (Bouhadiba 1988, p. 204, 244)

a. /ragːad-u/

rag:d -u
make.lie.down.PFV -SBJ.3PL.M

'they put to sleep'[4]

b. /nasabːagak/
n -seb:g -ek
SBJ.1SG -make.overtake.IMPF -OBJ.3SG

'I make you overtake'

This is what we expect from a C variety.

Closed Syllable Shortening

Algerian undergoes CSS-Morph as seen in (123-a), but otherwise permits medial CVːC syllables as shown in (123-b).

(123) Closed Syllable Shortening (Bouhadiba 1988, p. 152, 313)

a. ʃaf -∅
see.PFV -SBJ.3SG.M
'he saw'

b. ʃaf -na
see.PFV -SBJ.1PL
'we saw'

c. ħmaːr -na
donkey -POSS.1PL
'our donkey'

This is what we expect from a C variety.

Opaque Stress and Epenthesis

Algerian does not have opaque stress and epenthesis interactions.

Evaluation

Similar to Tunisian, Algerian is close to a canonical C variety. The non-canonical behaviour exhibition is the presence of stem metathesis, not found in Tunisian, which is more consistent with a VC variety.

[4] Translation not available in Bouhadiba (1988), translation here is mine

2.4.12 Moroccan Casablanca

Based on the epenthesis in the example shown in (124), we would expect Moroccan Casablanca to behave as a VC variety.

(124) Epenthesis in a Triconsonantal Cluster (Boudlal 2001, p. 173)

faxə̲r -na
coal -POSS.1PL

'our coal'

Final Clusters

Moroccan Casablanca does permit phrase final CC clusters. These occur with sonority falls as in (125-a) and sonority plateaux as in (125-b). However, in nouns, they can only occur with sonority falls — thus in (125-c)–(125-d), the schwa falls between the second and third consonants rather than the first and second, as otherwise the resulting clusters *fz and *ml would not fit the required sonority profile.

(125) Phrase Final CC Cluster (Boudlal 2001, p. 47, 49)

a. dˤərb

'hitting'

b. ktəb -t
write.PFV -SBJ.1SG
'I wrote'

c. qfəz

*qəfz

'cage'

d. nməl

*nəml

'ants'

Note that for this variety, the presence or absence of an epenthetic vowel can reflect lexical categories rather than phonotactic requirements. This can be seen in (126), where both examples have the same consonantal segments, but differ in the placement of the epenthetic vowel.

(126) Lexical Differences in Final Cluster Permissibility (Boudlal 2001, p. 279)

a. dˤərba

'a hit'

b. dˤarəb

'hitting'

These restrictions on final clusters are what we would expect from a VC variety.

Initial Clusters

Moroccan Casablanca permits phrase initial CC clusters as in (127-a)–(127-b), as well as phrase initial geminates, whether true geminates as in (127-e) or fake geminates created through assimilation as in (127-c)–(127-d).

(127) Phrase Initial CC cluster (Boudlal 2001, p. 47, 72, 82)

a. lʕəb -na
play.PFV -SBJ.1PL
'we played'

b. sˤfər

'yellow'

c. s -sma (>l-sma)
DEF -sky
'the sky'

d. d -dir (>t-dir)
SBJ.2SG -do
'you do'

e. dːa

'he took'

This fits what we would expect from a VC variety.

Epenthesis in Triconsonantal Clusters

Moroccan Casablanca permits medial CCC syllables as shown in (128-a)–(128-b) but does epenthesise some as part of the morphological template (128-c)–(128-d) rather than due to phonotactic restrictions.

(128) Medial CəCC syllables (Boudlal 2001, p. 85, 153, 172, 173)

a. ʒərʒr -u
trail.PFV -SBJ.3PL.M
'they trailed'

b. kərkb -u
roll.PFV -SBJ.3PL.M
'they rolled'

c. saməħ-na
forgive.PFV -SBJ.1PL
'we forgave'

d. faxər -na
coal -POSS.1PL
'our coal'

This is **not** what we would expect from a VC variety, but is seen in C varieties.

Stem Metathesis

Moroccan Casablanca does not undergo stem metathesis nor does it syncopate the stem vowel, as shown in (129).

(129) No Metathesis or Stem Vowel Deletion(Boudlal 2001, p. 110)

jə -dˤrəb -∅ -na
SBJ.3 -hit.IMPF -SBJ.SG -OBJ.1PL

'he hits us'

This is **not** what we would would expect from a VC or C variety.

Postgeminate High Vowel Deletion

There is no data on postgeminate high vowel deletion in Boudlal (2001).

Closed Syllable Shortening

Moroccan Casablanca does not undergo CSS-PHON as it does not have long vowels.

Opaque Stress and Epenthesis

There are no opaque stress and epenthesis interactions.

Evaluation

Moroccan Casablanca is further from a canonical C variety than Tunisian. The data above show that while Moroccan Casablanca does display some of the features of a VC variety in terms of initial and final clusters, it does not fit in the case of stem metathesis or epenthesis in triconsonantal clusters, where it behaves more like a C variety.

2.4.13 Libyan Tripoli

Epenthesis into triconsonantal clusters is optional in Libyan Tripoli, as shown in (130) with optionality marked by brackets. Where it does occur, it is consistent with a VC variety, but to not occur is more consistent with a C variety.

(130) Epenthesis in a Triconsonantal Cluster (Yoda 2005, p. 70)

qal-(i̠)t-l -u
say.PFV -SBJ.1SG -DAT -OBJ.3SG.M

'she said to him'

Final Clusters

Libyan Tripoli permits phrase final consonantal clusters as shown in (131), with optional epenthesis if it is a final triconsonantal cluster as in (131-c).

(131) Phrase Final Consonantal Clusters (Yoda 2005, p. 22, 99, 96)

a. kəlb

'dog'

b. xərːəf -t
tell.PFV -SBJ.1SG
'I told'

c. ma qəl -t -(ə̠)ʃ
NEG say.PFV -SBJ.1SG -NEG
'I did not say'

This is consistent with either VC or C varieties.

Initial Clusters

Libyan Tripoli permits phrase initial consonant clusters as shown in (132). However there is some free variation in prothesis as shown in (132-c-d), marked by brackets.

(132) Phrase Initial Consonantal Clusters (Yoda 2005, p. 92, 52, 98, 42)

a. ktʧəb -∅
write.PFV -SBJ.3SG.M
'he wrote'

b. stħəm -∅
bathe. -SBJ.3SG.M
'he bathed'

c. (ə̲)klab

'dogs'

d. (ə̲)mʃa -∅
go.PFV -SBJ.3SG.M
'he went'

There are a few cases of initial geminate clusters. These tend to be underlying clusters, as in (133-a). Where a (fake) geminate emerges due to assimlation of the definite article, prothesis occurs as in (133-b) (though for the non-assimilated definite article it is optional as in (133-c)). If the word starts with /l/ then epenthesis occurs as in (133-d).

(133) Phrase Initial Geminate Clusters (Yoda 2005, p. 99, 106, 97)

a. mːali -ja
parent.PL -POSS.1SG
'my parents'

b. l-tˤasˤa > *tˤtˤasˤa > ətˤtˤasˤa

'the cup'

c. (ə̲)l-qadˤi

'the judge'

d. lə-lːuzir

'to the minister'

This is consistent with either VC or C varieties.

Epenthesis in Triconsonantal Clusters

Libyan Tripoli does not epenthesise into all triconsonantal clusters. Where it does, it epenthesises to the left of the central consonant as in (134-a)–(134-b). Optionality is again marked with brackets. The vowel before the dative in (134-a) must be epenthetic as it is not found in other cases such as (134-b).

(134) Medial CCC Clusters (Yoda 2005, p. 103, 70, 105, 104, 57)

a. katʃb -ətʃ -ə̠l -kəm
write.PFV -SBJ.3SG.F -DAT -OBJ.2PL
'she wrote to you (pl.)'

b. qal -(ə̠)t -l -u
say.PFV -SBJ.3SG.F -DAT -OBJ.3SG.M
'she said to him'

c. kəlb -kəm
dog -POSS.2PL
'your (pl.) dog'

d. xəbz -na
bread -POSS.1PL
'our bread'

e. dərkb -u
hit.the.drum.PFV -SBJ.3PL.M
'they hit the drum'

f. sənsla
'necklace'

This is a combination of canonical VC and C behaviours.

Stem Metathesis

Libyan Tripoli does undergo stem metathesis, as shown in (135).

(135) Stem metathesis (Yoda 2005, p. 117)

a. qtʃəl -∅
kill.PFV -SBJ.3SG.M -OBJ.3SG.F
'he killed'

b. qətʃl -∅ -u
kill.PFV -SBJ.3SG.M -OBJ.3SG.M

'he killed him'

This is more consistent with VC varieties.

Postgeminate High Vowel Deletion

Libyan Tripoli undergoes postgeminate high vowel deletion, as shown in (136) where the vowel after the geminate is syncopated if followed by another syllable.

(136) Postgeminate High Vowel Deletion (Yoda 2005, p. 104)

a. /j-ʕəlːimu/
j -ʕəlːm -u
SBJ.3 -teach.IMPF -SBJ.PL.M

'they (m.) teach'

b. msəkːrin

'closed'

c. qəlːəq -∅
bore.PFV -SBJ..M
'he bored (someone)'

d. qəlːq -u
bore.PFV -SBJ.3PL.M
'they (m.) bored (someone)'

This is consistent with either VC or C varieties.

Closed Syllable Shortening

Libyan Tripoli does not have long vowels, but does have a length alternation between V and ə. In verbs, this alternation occurs as in (137-a)–(137-b) as the result of CSS-Morph, but this is not the case in nouns, as shown in (137-c)–(137-d) where word medial CVC syllables surface.

(137) Closed Syllable Shortening (Yoda 2005, p. 162, 203, 111)

a. bas -∅
kiss.PFV -SBJ.3SG.M
'he kissed'

b. bəs -ʧ
kiss.PFV -SBJ.1SG
'I kissed'

c. sanyəʧ

'field'

d. sakn -in
live.PTCP -PL
'living (pl.)'

This is consistent with either VC or C varieties.

Opaque Stress and Epenthesis

Libyan Tripoli does not undergo opaque interactions between stress and epenthesis.

Evaluation

Just as with Moroccan Casablanca and Algerian, Libyan Tripoli is a less canonical C variety than Tunisian. Much of the behaviour of Libyan Tripoli is consistent with either VC or C varieties, but the stem metathesis is more common with VC varieties, and the triconsonantal cluster epenthesis is a mix of both.

2.4.14 Rufaidah

Based on the epenthesis in the example shown in (138), we would expect Rufaidah to behave as a CV variety.

(138) Epenthesis in a Triconsonantal Cluster (Prochazka Jr. 1988, p. 213)
gul -ti̱ -l -i
say.PFV -SBJ.2SG -DAT -OBJ.1SG

'you told me'

Final Clusters

Rufaidah permits phrase final consonantal clusters regardless of sonority. In (139-a)–(139-c) sonority rises; in (139-d) sonority falls; and sonority plateaus in (139-e).

(139) Phrase Final Consonantal Clusters (Prochazka Jr. 1988, p. 30-32)

a. ðahn -∅
wake.PFV -SBJ.3SG.M
'he woke up'

b. ʃarb -∅
drink.PFV -SBJ.3SG.M
'he drank'

c. zaʕl -∅
be.upset.PFV -SBJ.3SG.M
'he became upset'

d. labs -∅
wear.PFV -SBJ.3SG.M
'he wore'

e. rakb -∅
ride.PFV -SBJ.3SG.M
'he rode'

This is what we would expect from a CV variety.

Initial Clusters

Rufaidah permits phrase initial consonantal clusters. There is no data available in Prochazka Jr. (1988) as to the presence or absence of initial geminate clusters.

(140) Phrase Initial Consonantal Clusters (Prochazka Jr. 1988, p. 42-46)

a. tkalːam -∅
speak.PFV -SBJ.3SG.M
'he spoke'

b. txaːsˤam -∅
quarrel.PFV -SBJ.3SG.M
'he had a quarrel'

c. j -zarːig
SBJ.3SG.M -become.blue.IMPF
'he becomes blue'

This is **not** what we would expect from a CV variety, but is more consistent with VC or C varieties.

Epenthesis in Triconsonantal Clusters

Epenthesis into a tri-consonantal cluster varies. It does not occur in (141-a), but does in (141-b)–(141-c), where the vowel is inserted to the right.

(141) Medial Triconsonantal Clusters broken by Epenthesis (Prochazka Jr. 1988, p. 200, 213, 216)

a. bint -kum
daughter -POSS.2PL
'your daughter'

b. gul -ti̲ -l -i
say.PFV -SBJ.2SG -DAT -OBJ.1SG
'you told me'

c. ʔiʃir -li̲ -na
buy.IMP -DAT -OBJ.1PL
'buy for us!'

The epenthesis that occurs is consistent with a CV variety, but the lack of epenthesis in other cases is more consistent with a C variety.

These must be epenthetic vowels as they are not found in the paradigms unless there is a triconsonantal cluster. For example, (142-a) surfaces without a following vowel, and (142-b) does not have one before the dative.

(142) Epenthesis Evidence (Prochazka Jr. 1988, p. 77, 211)

a. gul -t
say.PFV -SBJ.2SG
'you told'

b. ʔastar -at -l -ih
buy.PFV -SBJ.3SG.F -DAT -OBJ.3SG.M
'she bought for him'

Stem Metathesis

Triconsonantal clusters are also created optionally by syncope rather than undergoing stem metathesis. The vowel that is optionally syncopated is marked with brackets.

(143) Medial Triconsonantal Clusters Created by Syncope (Prochazka Jr. 1988, p. 33-37)

a. ju -kt(u)b -uːn
SBJ.3.M -write.IMPF -SBJ.PL
'they write'

b. ju -ðr(u)b -uːn
SBJ.3.M -hit.IMPF -SBJ.PL
'they hit'

c. ju -sk(u)b -uːn
SBJ.3.M -pour.IMPF -SBJ.PL
'they pour'

This is more consistent with C varieties.

Postgeminate High Vowel Deletion

Postgeminate high vowel deletion is optional, as shown in (144). The vowel that is optionally syncopated is marked with brackets.

(144) Postgeminate High Vowel Deletion(Prochazka Jr. 1988, p. 46)

a. j -salː(i)m -uːn
SBJ.3.M -hand.IMPF -SBJ.PL
'they (m.) hand'

b. j -kasː(i)r -uːn
SBJ.3.M -smash.IMPF -SBJ.PL
'they (m.) smash'

That this occurs is not consistent with a CV variety, but more common in C varieties.

Closed Syllable Shortening

Internal CVːC syllables do not undergo shortening beyond subject induced CSS-Morph. In (145-a)–(145-b), the CVːC syllable surfaces, but in (145-c) it is shortened and raised.

(145) Closed Syllable Shortening (Prochazka Jr. 1988, p. 198, 77–78)

a. beːt -hum
house -POSS.3PL
'their house'

b. saːr -∅
go.PFV -SBJ.3SG.M
'he went'

c. sir -t
go.PFV -SBJ.1SG
'I went'

This is **not** what we would expect from a CV variety, but is more consistent with a C variety.

Opaque Stress and Epenthesis

Rufaidah does not undergo opaque interactions between stress and epenthesis.

Evaluation

The data above show that Rufaidah appears to behave like a C variety, sharing features with Tunisian and Algerian. However, where epenthesis occurs it follows CV patterns, not the expected VC pattern.

2.4.15 Rwaili

Based on the epenthesis in the example shown in (146), we would expect Rwaili to behave as a CV variety.

(146) Epenthesis in a Triconsonantal Cluster (Prochazka Jr. 1988, p. 189)

simiʕ -ti̱ -kam
hear.PFV -SBJ.1SG -OBJ.2PL

'I heard you (pl.)'

Final Clusters

Rwaili permits phrase final consonantal clusters. This is shown in (147).

(147) Phrase Final Consonantal Clusters (Prochazka Jr. 1988, p. 28, 78)

a. kitab -t
write.PFV -SBJ.1SG
'I wrote'

b. gil -t
say.PFV -SBJ.1SG
'I said'

c. ʃil -t
carry.PFV -SBJ.1SG
'I carried'

This is what we would expect from a CV variety.

Initial Clusters

Rwaili permits phrase initial consonantal clusters, as shown in (148). There is no data in Prochazka Jr. (1988) about the presence or absence of initial geminate clusters.

(148) Phrase Initial Consonantal Clusters (Prochazka Jr. 1988, p. 28-29, 193)

a. ktib -at
write.PFV -SBJ.3SG.F
'she wrote'

b. ðbaħ -at
kill.PFV -SBJ.3SG.F
'she killed'

c. bgir -ih
cow.PL -POSS.3SG.M
'his cows'

This is **not** what we would expect from a CV variety, and is more consistent with a C or VC variety.

Epenthesis in Triconsonantal Clusters

Rwaili breaks up triconsonantal clusters by epenthesising to the right. This is shown in (149).

(149) Epenthesis in Triconsonantal Clusters(Prochazka Jr. 1988, p. 212)
gaːl -∅ -la -ha
say.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F

'he told her'

This must be an epenthetic vowel as it is not found in forms without a triconsonantal cluster, as in (150).

(150) Epenthesis Evidence (Prochazka Jr. 1988, p. 212)

gal -at -l -ih
say.PFV -SBJ.3SG.F -DAT -OBJ.3SG.M

'she told him'

The epenthesis is what we would expect from a CV variety.

Stem Metathesis

Rwaili does not undergo stem metathesis but syncopates a vowel to create a triconsonantal cluster, as shown in (151).

(151) Triconsonantal Cluster Created by Syncope (Prochazka Jr. 1988, p. 34)

a. ta -ktb -uːn
SBJ.2 -write.IMPF -SBJ.PL.M
'you (m.pl.) write'

b. ja -ðrb -uːn
SBJ.3 -hit.IMPF -SBJ.PL.M
'they (m.) hit'

c. ta -ʃrub -uːn
SBJ.2 -drink.IMPF -SBJ.PL.M
'you (m.pl.) drink'

This is **not** what we would expect from a CV variety, and is more consistent with a C variety.

Postgeminate High Vowel Deletion

Rwaili does undergo postgeminate high vowel deletion, so CVCː syllables surface as shown in (152).

(152) Postgeminate High Vowel Deletion (Prochazka Jr. 1988, p. 50, 46)

a. kasˤːm -u
break.IMP -SBJ.PL.M
'break (m.pl.)!'

b. t -badːl -iːn
SBJ.2 -change.IMPF -SBJ.PL.M
'you (pl.) change'

c. t -garːb -iːn
SBJ.2 -draw.near.IMPF -SBJ.PL.M
'you (pl.) draw near'

This is **not** what we would expect from a CV variety, and is more consistent with a C or VC variety.

Closed Syllable Shortening

Rwaili shortens internal CVːC syllables. Thus, we see in (153-a) there is a heavy open syllable [saː], but in (153-b) the corresponding syllable is closed and the vowel has been shortened to [saf].

(153) Closed Syllable Shortening (Prochazka Jr. 1988, p. 47)

a. j -saːfir
SBJ.3 -travel.IMPF
'he travels'

b. t -safr -iːn
SBJ.2 -travel.IMPF -SBJ.SG.F
'you (f.s.) travel'

This is what we would expect from a CV variety.

Opaque Stress and Epenthesis

Rwaili does not undergo opaque interactions between stress and epenthesis. This is what we would expect from a CV variety.

Evaluation

Rwaili is similar to Rufaidah, but behaves less canonically than the other varieties discussed here, as it does not behave like a canonical CV or C variety. If we analyse it as a CV variety, the final clusters, epenthesis, closed syllable shortening and transparent stress meet expectations; but the initial clusters, metathesis, and postgeminate high vowel deletion do not. If we analyse it as a C variety, the final and initial clusters, metathesis, post-geminate high vowel deletion and stress meet expectations; but the closed syllable shortening does not. It is perhaps closer to a Cv variety such as Sˤanʕaːniː and Makːan under Watson's (2007) analysis.

2.4.16 Haːyil

Based on the epenthesis in the example shown in (154), we would expect Haːyil to behave as a CV variety.

(154) Epenthesis in a Triconsonantal Cluster (Prochazka Jr. 1988, p. 185)
simiʕ -ti̲ -hin
hear.PFV -SBJ.2SG -OBJ.3PL.F
'you heard them (f.pl.)'

Final Clusters

Haːyil permits phrase final consonantal clusters, as shown in (155).

(155) Phrase Final Consonantal Clusters (Prochazka Jr. 1988, p. 28, 77–78)

a. kitab -t
write.PFV -SBJ.1SG
'I wrote'

b. gil -t
say.PFV -SBJ.1SG
'I said'

c. ʃuf -t
see.PFV -SBJ.1SG
'I saw'

This is what we would expect from a CV variety.

Initial Clusters

Haːyil also permits phrase initial consontal clusters regardless of sonority. In (156-a) there is a sonority plateau; in (156-b) there is a sonority fall; and in (156-c-d) there is a sonority rise. There is no data in Prochazka Jr. (1988) with regard to initial geminate clusters. Note the 3rd person feminine singular subject marker [-eh] is different in this variety to the others where it is often [-at].

(156) Phrase Initial Consonantal Clusters (Prochazka Jr. 1988, p. 28-29, 193)

a. ktub -eh
write.PFV -SBJ.3SG.F
'she wrote'

b. ðbaħ -eh
kill.PFV -SBJ.3SG.F
'she killed'

c. mal -ah
camel -POSS.3SG.F
'her camel'

d. ɣnimta -k
sheep -POSS.2SG
'your sheep'

This is **not** what we would expect from a CV variety, but is more typical of a C or VC variety.

Epenthesis in Triconsonantal Clusters

Haːyil epenthesises into triconsonantal clusters to the right optionally. Thus, we see epenthesis occurs in (157-a)–(157-b) but not (157-c).

(157) Triconsonantal Cluster Epenthesis (Prochazka Jr. 1988, p. 189, 216–217)

a. simiʕ -ti̠ -kam
hear.PFV -SBJ.1SG -OBJ.2PL
'I heard you'

b. gaːl -at -la̲ -ham
say.PFV -SBJ.3SG.F -DAT -OBJ.3PL
'she told them'

c. gil -t -l -ah
say.PFV -SBJ.2SG -DAT -OBJ.3SG.F
'you told her'

The direction of epenthesis suggests a CV variety, but the optionality is not canonical, and is closer to a C variety.

Stem Metathesis

However, triconsonantal clusters are also optionally created by syncope, as shown in (158) where CVCC syllables surface.

(158) Triconsonantal Cluster by Syncope (Prochazka Jr. 1988, p. 33-35)

a. ji -staʕml -uːn
SBJ.3 -use.IMPF -SBJ.PL
'they use'

b. t -aktb -uːn
SBJ.2 -write.IMPF -SBJ.PL
'you (pl.) write'

c. ʔuɣsl -i
wash.IMP -SBJ.2SG.F
'wash (f.s.)!'

This is **not** what we would expect from a CV variety, but is more typical of a C variety.

Postgeminate High Vowel Deletion

Haːyil exhibits postgeminate high vowel deletion, as seen in (159) where CVCː syllables surface.

(159) Postgeminate High Vowel Deletion (Prochazka Jr. 1988, p. 46)

a. t -kasːr -iːn
SBJ.2 -smash.IMPF -SBJ.PL.F
'you (f.pl.) smash'

b. t -badːl -iːn
SBJ.2 -change.IMPF -SBJ.PL.F
'you (f.pl.) change'

c. t -garːb -iːn
SBJ.2 -draw.near.IMPF -SBJ.PL.F
'you (f.pl.) draw near'

This is **not** what we would expect from a CV variety, but is more typical of a C variety.

Closed Syllable Shortening

Haːyil shortens CVːC syllables if they are unstressed, as seen in (160) where the stressed vowel in (160-a) is long, but when it is no longer stressed in (160-b) due to affixation, it is shortened.

(160) Closed Syllable Shortening (Prochazka Jr. 1988, p. 47)

a. j -ˈsaːfir
SBJ.3 -travel.IMPF
'he travels'

b. t -safˈr -iːn
SBJ.2 -travel.IMPF -SBJ.SG.F
'you (f.s.) travel'

This is **not** quite what we would expect from a CV variety, as we would expect no CVːC at all, but the shortening of the unstressed CVːC is not typical of C or VC varieties either.

Opaque Stress and Epenthesis

Haːyil does not undergo opaque interactions between stress and epenthesis.

Evaluation

Just like Rwaili, Haːyil does not behave like a canonical CV or C variety. If we analyse it as a CV variety, the final clusters, closed syllable shortening and transparent stress meet expectations; but the initial clusters, optional triconsonantal epenthesis, metathesis, and postgeminate high vowel deletion do not. If we analyse it as a C variety, the final and initial clusters, metathesis, post-geminate high vowel deletion and stress meet expectations; but the optional triconsonantal epenthesis and closed syllable shortening do not. It is perhaps closer to a Cv variety such as Sˤanʕaːniː and Makːan under Watson's (2007) analysis.

2.4.17 Summary

I will now take each dialect group in turn to summarise the extent to which these varieties fit the existing typological expectations.

CV Varieties

The canonical expectation of a CV variety is: to permit final CC clusters; to not permit initial CC clusters and as such not syncopate high vowels from initial syllables; to prothesise a vowel before initial geminate clusters; to not undergo stem metathesis; to not undergo postgeminate high vowel deletion; to not permit word medial CVːC syllables; to not undergo opaque interactions between stress and epenthesis and to have stressable epenthetic vowels. Where CV varieties are more tolerant of medial CVːC than expected, they can be classified as Cv under Watson (2007). The extent to which the varieties discussed here fit this canonical expectation is summarised in Table 2.8 below.

Table 2.8: CV Varieties

Key: Y = Yes, N = No, FS = Fast Speech,
O = Optional, Y/N = Mixed Behaviour, - = No Data

Variety	-CC#	*#CC-	*#GG-	*CCC	*Metathesis	*Post-GG HVD	*CVːC	*Opaque Stress
Cairene	Y	Y	Y	Y	Y	Y	Y	Y
Sˤanˤaːniː	Y	N	-	FS	Y	Y	O	Y
Makːan	Y	Y	-	Y	Y	Y/N	Y	Y

These varieties behave as expected with regard to final consonantal clusters, metathesis, and lack of opaque stress. Sˤanˤaːniː permits more medial CCC and CVːC than CV varieties are expected to, and on this basis Watson (2007) identified it as a Cv variety. Whilst the variety of Makːan that Watson (2007) explores is also a Cv variety, this particular variety behaves more canonically as a CV dialect, with the only difference being variation in postgeminate high vowel deletion.

VC Varieties

The canonical expectation of a VC variety is: to not permit final CC clusters or to do so with sonority restrictions; to permit initial CC clusters or prothesise a vowel before an initial CC cluster; to permit initial geminate clusters; to undergo stem metathesis from CCVC to CVCC; to undergo postgeminate high vowel deletion; to permit internal CVːC syllables; and to undergo opaque interactions between stress and epenthesis. The extent to which the varieties discussed here fit this canonical expectation is summarised in Table 2.9 below.

Table 2.9: VC Varieties

Key: Y = Yes, N = No, F = Fast Speech, O = Optional,
Y/N = Mixed Behaviour, SR = Yes with Sonority Restrictions, - = No Data

Variety	*-CC#	#CC-	#GG-	*CCC	Metathesis	Post-Cː HVD	CVːC	Opaque Stress
Palestinian	SR	Y	-	Y/N	-	Y	Y	Y
Lebanese	SR	Y	Y	Y/N	Y	-	Y	Y
Muscat	Y/N	Y	Y	Y/N	Y	O	Y	Y
Qatari	SR	Y	Y	Y	Y	N	Y	Y
Iraqi	Y	Y	Y	Y	Y	Y	Y	Y
Wadi Ramː	SR	Y	-	N	Y	Y	Y	N

These varieties broadly behave as expected in terms of initial and final consonantal clusters, metathesis, permitting internal CVːC syllables and opaque stress and epenthesis interactions. However, there is more variation than predicted by Kiparsky (2003) and Watson (2007) in terms of medial triconsonantal clusters, which are permitted by Palestinian, Lebanese, Muscat, and Wadi Ramː.

C Varieties

The canonical expectation of a C variety is: to permit final CC clusters; to permit initial CC clusters; to permit initial geminate clusters; to syncopate the stem vowel from CCVC to CCC; to undergo postgeminate high vowel deletion; to permit medial CVːC syllables; and to not undergo opaque interactions between stress and epenthesis. The extent to which the varieties discussed here fit this canonical expectation is summarised in Table 2.10 below.

Table 2.10: C Varieties

Key: Y = Yes, N = No, FS = Fast Speech,
O = Optional, Y/N = Mixed Behaviour, - = No Data

Variety	-CC#	#CC-	#GG-	CCC	Stem Syncope	Post-GG HVD	CVːC	*Opaque Stress
Tunisian	Y	Y	Y	Y	Y	Y	Y	Y
Moroccan Casablanca	Y/N	Y	Y	Y/N	Y	Y	Y	Y
Algerian	Y	Y	Y	Y/N	Y	Y	Y	Y
Libyan Tripoli	Y	Y	Y	Y/N	Y	Y	Y	N
Rufaidah	Y	Y	-	Y/N	O	Y	Y	N
Rwaili	Y	Y	-	Y/N	O	Y	N	Y
Haːyil	Y	Y	-	Y/N	O	Y	N	Y

These varieties broadly behave as expected with regard to initial and final consonantal clusters, postgeminate high vowel deletion, and lack of opaque stress and epenthesis interactions. However, there is wide variation in behaviours with triconsonantal clusters in all varieties except Tunisian. Note that we would expect C varieties to be closer to VC than CV, and as such where epenthesis occurs for it to fit the VC position. However, in Rufaidah, Rwaili, and Haːyil epenthesis fits CV positioning.

Evaluation

From the data above it is clear that these dialects do not fit all the canonical behaviours predicted. The analysis of Kiparsky (2003) cannot account for how often CVːC syllables are found word-internally, nor the asymmetry in the permissibility of internal CVːC vs CVCC syllables. While Watson (2007) solves this through mora sharing, her approach does not account for the variable permissibility of CVCC syllables demonstrated by the data explored here. It is accepted under a Canonical Typology approach to demarcate the canonical boundaries then observe how the data fits within this canonical space. However, the extent of deviation in terms of medial CVXC syllables suggests that a revision of the canonical expectations is required. To do this, I will first lay out how different varieties tolerate CVXC syllables.

2.5 Tolerance of Medial Syllables

It is clear from the previous section that these dialects display more CVXC syllables than is predicted in the literature. Therefore, in this section I will examine each variety in turn to see whether it permits medial CVCC, CVːC or CVCː syllables.

Whether a medial CVCː syllable degeminates is debated. Qafisheh (1977) and Kenstowicz and Kisseberth (1979) and Farwaneh (2009) claim all medial CVCː are degeminated. However, Obrecht (1965) claims there is a perceptible contrast for CVC versus CVCː. Abu-Salim (1982a, p. 221) suggests the phonetic realisation is different for geminates compared to singletons, producing the Palestinian minimal pairs (161-a)–(161-b) and (161-c)–(161-d).

(161) Palestinian Geminate-Singleton Minimal Pairs (Abu-Salim 1982a, p.220-221)

a. ʔimː -na
mother -POSS.1PL
'our mother'

b. ʔim -na
remove.PFV -SBJ.1PL
'we removed'

c. ʕa rabː -na
on god -POSS.1PL
'on our god'

d. ʕarab -na
arab.PL -POSS.1PL
'our Arabs'

Similarly, Bouhadiba (1988) argues that degemination does not occur in Algerian based on the following minimal pairs (162-a)–(162-b) and (162-c)–(162-d).

(162) Algerian Geminate-Singleton Minimal Pairs (Bouhadiba 1988, p. 323)

a. kadːb -u
accuse.of.lying.PFV -SBJ.3PL
'they accused someone of lying'

b. kadb -u
lie.PFV -SBJ.3PL

‘they lied’

c. tabːʕ -at
follow.PFV -SBJ.3SG.F
‘she followed’

d. ta -bʕat
SBJ.2SG.M send.IMPF
‘you m.s. send’

Preserving the geminate is important as it is providing morphological information here. In (162-a), the gemination of the middle root consonant creates a causative meaning. Long segments play an important role in Arabic morphology, as shown in the following Lebanese examples. In (163-a), the bare root is found; in (163-b), the long vowel contributes the imperative sense; in (163-c), the geminate contributes the transitive sense; in (163-d), the long vowel creates a different sense again; and in (163-e), the geminate combined with initial dental consonant creates the passive meaning.

(163) Role of Quantity in Lebanese Morphology (Haddad 1984, p.26, 86, 280)

a. nzil -t
descend.PFV -SBJ.1SG
‘I came down, descended’

b. nzaːl
descend.IMP
‘come down!’

c. nazːˈal -∅
bring.down.PFV -SBJ.3SG.M
‘he brought down’

d. naːˈzal -it
encounter.PFV -SBJ.3SG.F
‘she taught’

e. tˈnazːal -∅
be.brought.down.PFV -SBJ.3SG.M
‘he was brought down’

There is clear evidence that long vowels can be exempt from shortening operations where they are morphologically important. In the Rural Palestinian variety described by Younes (1995), CVː undergoes shortening when immediately to the left of a stressed syllable, as shown in (164).

(164) Rural Palestinian Open Syllable Shortening (Younes 1995)

a. maˈkaːtib
   ‘offices’
b. makaˈtib-na
   ‘our offices’
c. muftaːħ
   ‘key’
d. muftaħ-eːn
   ‘two keys’

However, if the long vowel is providing morphological information, it does not undergo shortening. Thus, the long vowel is shortened in (165-b), but not in (165-d) where it is providing information about the root structure.

(165) Rural Palestinian Open Syllable Shortening Exceptions (Younes 1995)

a. ħaraːmi
   ‘thief’
b. ħaramiː-ha
   ‘her thief’
c. ħaːmi
   ‘protector’
d. ħaːmiː-ha
   ‘her protector’

Therefore, given the role of long segments in the morphology, degemination that would neutralise morphological contrasts should be dispreferred. While neutralisation does exist, these minimal pairs suggest that neutralisation does not occur.

The phonetic realisation of these tautosyllabic geminates is a different matter. Lahiri and Marslen-Wilson (1992) suggests that vocalic release is needed for the perception of geminates, which would not be present with tautosyllabic geminates. However, there is some evidence that long vowels are shorter before medial geminates than medial clusters in Lebanese Arabic, so it may be this provides the acoustic information needed to identify gemination (Khattab and Al-Tamimi 2014).

Therefore, I expect here that medial CVCː syllables do exist in some Arabic varieties, following affixation or syncope, and that where they are not permitted, they undergo epenthesis rather than degemination.

In the following, I explore the tolerance of medial CVXC syllables across the varieties in question.

2.5.1 Tunisian

Tunisian permits medial CVCC regardless of sonority. This is shown in (166) where sonority rises in (166-a)–(166-c), falls in (166-d)–(166-f), and plateaus in (167-a)–(167-b). Although (166-d) is a loanword, its behaviour matches that of other non-loanword cases. [5]

(166) Medial CVCC Sonority Rise and Fall (Maamouri 1967, p.22, 115, 103, 15, 22)

a. sukmt -ik
   cold -POSS.2SG
   'your cold'

b. ji -tfda
   SBJ.3 -be.revenged.IMPF
   'it can be revenged'

c. ja -kfr -u
   SBJ.3 -blaspheme.IMPF -SBJ.PL
   'they blaspheme'

[5] The keen phonologist may ask if in cases of sonority rise such as 166[c] these clusters are syllabified as a complex onset instead. According to Maamouri (1967, p.15), the source for this data, there are no medial CCVC or CCV syllables, which implies these are indeed syllabified as complex codas. However, even if a few examples here and in the following are syllabifed as complex onsets instead, there is sufficient data with clusters that we would expect to be complex codas based on sonority that the analysis holds. The reader is further directed to the assorted cases of final consonant clusters in varieties such as Cairene that flount the sonority hierarchy, suggesting that the SSP holds a weaker place in the phonology of at least some Arabic varieties than might be expected.

d. birdgaːn
 ‘oranges’

e. kalːam -t -hum
 talk.PFV -SBJ.1SG -OBJ.3PL
 ‘I talked to them’

f. salsla
 ‘chain’

(167) Medial CVCC Sonority Plateau (Maamouri 1967, p. 149, 22)

a. maktba
 ‘library’

b. raqbt -ik
 neck -POSS.2SG
 ‘your neck’

Tunisian permits medial CVːC syllables, as shown in (168-a)–(168-c), though still undergoes CSS-MORPH as shown in (168-d).

(168) Medial CVːC Syllables (Maamouri 1967, p. 5, 69)

a. riːħ -na
 wind -POSS.1PL
 ‘our wind’

b. ruːħ -na
 soul -POSS.1PL
 ‘our soul’

c. raːqsa
 ‘a dancer’

d. ruħ -na
 go.PFV -SBJ.1PL
 ‘we went’

Tunisian permits medial CVCː syllables, as shown in (169).

(169) Medial CVCː (Maamouri 1967, p. 10, 163)

a. laʕːb -∅ -u
 make.play.PFV -SBJ.3SG.M -OBJ.3SG.M
 ‘he made him play’

b. qulːt -ik
water.jug -POSS.2SG
'your water-jug'

c. farːq -it
distribute.PFV -SBJ.3SG.F
'she distributed'

Thus, Tunisian permits all CVXC syllables word-medially.

2.5.2 Wadi Ramː

Wadi Ramː permits medial CVCC syllables, as shown in (170).

(170) Medial CVCC Syllables (Al Mashaqba 2015, p. 110, 208)

a. ta -nks -uw
SBJ.2 -return.back.IMPF -SBJ.PL.M
'you (m.pl.) return back'

b. ħurmt -ah
wife -POSS.3SG.M
'his wife'

c. ʤib -t -l -ak
bring.PFV -SBJ.1SG -DAT -OBJ.2SG
'I brought you'

Wadi Ramː permits medial CVːC syllables as shown in (171-a)–(171-c), as well as CVːCC syllables shown in (171-d).

(171) Medial CVːC (Al Mashaqba 2015, p. 109, 112, 163, 120)

a. talafoːn -aːt -na
mobile.phone -PL.F -POSS.1PL
'our mobile phones'

b. maːsk -iːn
catch.PTCP -PL.M
'they (m.) are catching'

c. la t -ʕiːd -ha
NEG SBJ.3SG.M -do.IMPF -OBJ.3SG.F
'do not (m.s.) do this'

d. ʤaːb -l -na
bring.PFV -DAT -OBJ.1PL
'he brought for us'

Wadi Ramː permits medial CVCː syllables, as shown in (172).

(172) Medial CVCː (Al Mashaqba 2015, p. 109)
mitʕalːm -iːn
educate.PTCP -PL.M

'educated (m.pl.)'

Therefore, Wadi Ramː permits medial CVCC, CVːC and CVCː syllables without restriction, just as seen above with Tunisian.

2.5.3 Rufaidah

Rufaidah permits medial CVCC, as shown in (173), though this varies with optional stem syncope as in (173-a)–(173-c) where the brackets indicate the stem vowel that is optionally syncopated.

(173) Medial CVCC (Prochazka Jr. 1988, p. 34, 200)

a. ju -kt(u)b -uːn
SBJ.3 -write.IMPF -SBJ.PL.M
'they write'

b. ju -rb(u)tˤ -uːn
SBJ.3 -tie.IMPF -SBJ.PL.M
'they tie'

c. ji -ks(i)r -uːn
SBJ.3 -break.IMPF -SBJ.PL.M
'they break'

d. ʔibn -hum
son -POSS.3PL
'their son'

e. bint -kum
daughter -POSS.2PL
'your daughter'

Rufaidah permits medial CVːC syllables, as shown in (174).

(174) Medial CVːC (Prochazka Jr. 1988, p.50, 198)

a. gaːbl -u
meet.IMP -PL.M
'meet (m.pl.)!'

b. beːt -hum
house -POSS.3PL
'their house'

Rufaidah permits medial CVCː syllables as shown in (175), although postgeminate high vowel deletion is optional as in (175-b-c).

(175) Medial CVCː (Prochazka Jr. 1988, p. 155, 46)

a. kalːm -iː -na
speak.IMP -SG.F -OBJ.1PL
'speak (f.s.) to us!'

b. j -salː(i)m -uːn
SBJ.3 -hand.IMPF -SBJ.PL.M
'they hand'

c. j -kasː(i)r -uːn
SBJ.3 -smash.IMPF -SBJ.PL.M
'they smash'

Therefore, Rufaidah permits medial CVCC, CVːC and CVCː syllables, just as seen above with Tunisian and Wadi Ramː.

2.5.4 Rwaili

Rwaili permits medial CVCC syllables as shown in (176), though stem syncope is optional so (176-d) also surfaces.

(176) Medial CVCC (Prochazka Jr. 1988, p. 34, 50)

a. ji -staʕml -uːn
SBJ.3 -use.IMPF -SBJ.PL.M
'they use'

b. ta -ktb -uːn
SBJ.2 -write.IMPF -SBJ.PL.M
'you (pl.) write'

c. ja -ðrb -uːn
SBJ.3 -hit.IMPF -SBJ.PL.M
'they hit'

d. ta -ʃrub -uːn
SBJ.2 -drink.IMPF -SBJ.PL.M
'you (pl.) drink'

Rwaili permits medial CVːC as shown in (177). There are cases of shortening of long vowels if they are unstressed. However, these are cases of Open Syllable Shortening of unstressed long vowels before syncope (tsaːfiriːn > tsafiriːn >tsafriːn).

(177) Medial CVːC (Prochazka Jr. 1988, p. 47, 71, 51)

a. j -saːfir
SBJ.3 -travel.IMPF
'he travels'

b. t -safr -iːn
SBJ.2 -travel.IMPF -SBJ.SG.F
'you (f.s.) travel'

c. ja -azn -uːn
SBJ.3 -weigh.IMPF -SBJ.PL.M
'they weigh'

d. gaːbl -i
meet.IMP -SG.F
'meet (f.s.)!'

Rwaili permits medial CVCː syllables, as shown in (178).

(178) Medial CVCː (Prochazka Jr. 1988, p. 46)

a. t -kasːr -in
SBJ.2 -smash.IMPF -SBJ.SG.F
'you (f.s.) smash'

b. t -badːl -iːn
SBJ.2 -change.IMPF -SBJ.PL.M
'you (m.pl.) change'

c. t -garːb -iːn
SBJ.2 -draw.near.IMPF -SBJ.PL.M
'you (m.pl.) draw near'

Therefore, Rwaili permits medial CVCC, CVːC, and CVCː syllables, just as seen above with Tunisian, Wadi Ramː and Rufaidah.

2.5.5 Haːyil

Haːyil permits medial CVCC syllables as shown in (179), but stem syncope is optional, so forms without stem syncope surface as in (179-d).

(179) Medial CVCC (Prochazka Jr. 1988, p. 50, 34)

a. ji -staʕml -uːn
SBJ.3 -use.IMPF -SBJ.PL.M
'they use'

b. ta -ktb -uːn
SBJ.2 -use.IMPF -SBJ.PL.M
'you (pl.) write'

c. ʔuɣsl -i
wash.IMP -SG.F
'wash (f.s.)!'

d. ja -ɟils -uːn
SBJ.3 -sit.down.IMPF -SBJ.PL.M
'they sit down'

Haːyil permits medial CVːC syllables, but as in Rwaili unstressed Open Syllable Shortening occurs before syncope as shown in (180-d).

(180) Medial CVːC (Prochazka Jr. 1988, p. 71, 51, 47)

a. ja -aznuːn
SBJ.3 -weigh.IMPF
'they weigh'

b. gaːbl -i
meet.IMP -SBJ.F.SG
'meet (f.s.)!'

c. j -saːfir
SBJ.3 -travel.IMPF
'he travels'

d. t -safr -iːn
SBJ.2 -travel.IMPF -SBJ.F.SG
'you (f.s.) travel'

Haːyil permits medial CVCː syllables as shown in (181).

(181) Medial CVCː (Prochazka Jr. 1988, p. 46)

a. kasːr -u
smash.IMP -SBJ.M.PL
'smash (m.pl.)!'

b. t -badːl -iːn
SBJ.2 -change.IMPF -SBJ.PL

‘you (pl.) change’

c. t -gaːrb -in
SBJ.2 -draw.near.IMPF -SBJ.PL
‘you (pl.) draw near’

Therefore, Haːyil permits medial CVCC, CVːC, and CVCː syllables, just as seen above with Tunisian, Wadi Ramː, Rufaidah and Haːyil.

2.5.6 Algerian

Algerian is different to the above group.

Algerian permits medial CVCC if there is falling sonority, as shown in (182-a)–(182-c), but if sonority rises as in (182-d) it is broken by epenthesis.

(182) Medial CVCC with Sonority Restrictions (Bouhadiba 1988, p. 311-312, 299)

a. sans.la
‘chain’

b. qanq.ra
‘toad’

c. malk. -na
property -POSS.1PL
‘our property’

d. *hatr.f -∅ -u

hat<u>u</u>rf -∅ -u
dream.PFV -SBJ.3SG.M -OBJ.3SG.M
‘he dreamt it’

Algerian permits medial CVːC without restriction as shown in (183), though otherwise still undergoes CSS-MORPH as seen in (183-d).

(183) Medial CVːC (Bouhadiba 1988, p. 313, 152)

a. ħmaːr -na
donkey -POSS.1PL
‘our donkey’

b. saːbga
'overtaken (f.s.)'

c. ʃaːf -∅ -na
see.PFV -SBJ.3SG.M -OBJ.1PL
'he saw us'

d. ʃaf -na
see.PFV -SBJ.1PL
'we saw'

Algerian permits medial CVCː without restriction, as seen in (184).

(184) Medial CVCː (Bouhadiba 1988, p. 296)
ʃafː -∅ -na
inspire.pity.PFV -SBJ.3SG.M -OBJ.1PL
'he inspired pity in us'

Therefore, Algerian permits CVːC and CVCː without restriction, but does place restrictions on medial CVCC. This is a different pattern to that found in the varieties described above.

2.5.7 Palestinian

Palestinian follows the same pattern as Algerian. It breaks up most triconsonantal clusters to CvCC, thus avoiding internal CVCC syllables, as shown in (185).

(185) Epenthesis in Medial CVCC (Abu-Salim 1982a, p. 216)
katab -∅ -i̲l -ha
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F
'he wrote for her'

However, if C2 of the cluster is a coronal consonant, no epenthesis occurs regardless of morphological status so a CVCC syllable surfaces, as shown in (186) where(186-a)–(186-b) have an underlying CCC cluster whereas in (186-c)–(186-d) the CCC cluster is morphologically derived through affixation. While (186-a) is a loanword, it behaves the same way as templatic morphological forms such as (186-b) and (186-d).

(186) Medial CVCC with coronal C2 (Abu-Salim 1982a, p. 22, 60)

a. burd.ʔaːn
'oranges'

b. mift.xir
'proud'

c. ʔuxt. -hum
sister -POSS.3PL.M
'their sister'

d. katab -t -l -ak
write.PFV -SBJ.1SG -DAT -OBJ.2SG
'I wrote to you'

While (186-a) is a loanword, it behaves the same way as templatic morphological forms such as (186-b) and (186-d). However, some loanwords behave slightly differently to native words. For example, Loanwords can also avoid epenthesis regardless of the place of articulation of the C2 as shown in (187), though given the shared place of articulation in this particular example it is not surprising that epenthesis does not occur.

(187) Medial CVCC Loanword Exception (Abu-Salim 1982a, p. 22)
komb.juː.tar
'computer'

Palestinian permits medial CVːC syllables, as shown in (188).

(188) Medial CVːC syllables (Abu-Salim 1982a, p. 22, 26, 149)

a. baːl.tˤo
'coat'

b. ʃub.baːk. -hum
window -POSS.3PL.M
'their (m.) window'

c. naːʤ.ħ -a
successful -SG.F
'successful (f.s.)'

d. ʃaf -uː -l -na
see.PFV -SBJ.3PL -DAT -OBJ.1PL
'they (m.) saw (sth.) for/to us'

Palestinian permits medial CVCː syllables

(189) Medial CVCː Syllables (Abu-Salim 1982a, p. 219)

a. ʔimː -na
mother -POSS.1PL
'our mother'

b. sitː -kum
grandmother -POSS.2PL
'your (pl.) grandmother'

c. mfakː -hum
screw.driver -POSS.3PL
'their screwdriver'

Thus, there is a difference in how Palestinian treats quantity, with CVːC and CVCː syllables permitted without restriction, and CVCC permitted only if C2 of the cluster is coronal. This is clearly seen in the following pair where a C2 nasal is only permitted as part of a geminate cluster, not part of a non-homorganic cluster.

(190) Quantity Difference in Palestinian (Abu-Salim 1982a, p. 38, 196)

a. ʔimː -hum
mother -POSS.3PL
'their mother'

b. ʔibi̱n -hum
son -POSS.3PL
'their son'

While the nature of the restrictions on CVCC syllables is different to Algerian, Palestinianis following the same broad distribution.

2.5.8 Lebanese

Lebanese does not permit most medial CVCC syllables, as shown in (191). However, if the C2 is coronal, this CVCC syllable does surface, as we see in (192). Note that quadriconsonantal clusters are still broken up by epenthesis even in the presence of coronal consonants.

(191) No Medial CVCC (Haddad 1984, p. 40)

a. sakːar -∅ -i̲l -na
close.PFV -SBJ.3SG.M -DAT -OBJ.1PL
‘he closed for us’

b. ħafar -∅-i̲l -na
dig.PFV -SBJ.3SG.M -DAT -OBJ.1PL
‘he dug for us’

(192) Medial CVCC if C2 is coronal (Haddad 1984, p. 17, 10)

a. dˤarab -t -i̲l -ha
hit.PFV -SBJ.1SG -DAT -OBJ.3SG.F
‘I hit for her’

b. mistɣilː -ni
take.advantage.of.PTCP.M -OBJ.1SG
‘he took advantage of me’

c. mistʕidː
ready.PTCP.M
‘he was ready’

Lebanese permits medial CVːC syllables, as shown in (193).

(193) Medial CVːC (Haddad 1984, p. 10)

a. ʕaːmle

‘working class’

b. maʕaːmli

‘employee’

Lebanese permits medial CVCː syllables as shown in (194), both in the underlying representation as in (194-a)–(194-b) and when created through affixation as in (194-c).

(194) Medial CVCː (Haddad 1984, p. 158)

a. ħabː -∅ -na
love.PFV -SBJ.3SG.M -OBJ.1PL
‘he loved us’

b. sitː -na
grandmother -POSS.1PL
‘our grandmother’

c. ʃatːat -t -na
spread.apart.PFV -SBJ.2SG.M -OBJ.1PL
'you (m.s.) spread us apart'

Thus, Lebanese permits medial CVːC and CVCː syllables without restriction, but CVCC only if C2 is coronal. This is the same broad distribution as shown in Algerian and Palestinian, though the nature of the CVCC restrictions is more similar to Palestinian than Algerian.

2.5.9 Muscat

Muscat only permits CVCC in nouns in the case of a sonorant-obstruent cluster as in (195-a)–(195-c), so epenthesises into obstruent-obstruent clusters as in (195-d), unless C2 is a dentoalveolar fricative as in (195-e), where epenthesis is optional as marked by the brackets. Although (195-c) is an English loanword, it is not behaving differently from other items here.

(195) Sonority Restrictions on Medial CVCC in Nouns (Glover 1988, p. 60, 224–225)

a. bint -ha
daughter -POSS.3SG.F
'her daughter'

b. ward -ha
flower.PL -POSS.3SG.F
'her flowers'

c. bank -hum
bank -POSS.3PL.M
'their bank'

d. ʕaʃu̲b -hum
grass -POSS.3PL.M
'their grass'

e. kan(i̲)z -ha
pressed.date.PL -POSS.3SG.F
'her pressed dates'

Muscat permits CVCC in verbs if C2 is coronal, as shown in (196).

(196) Medial CVCC in Verbs with Coronal C2 (Glover 1988, p. 221, 206)

a. katab -t -ha
write.PFV -SBJ.1SG -OBJ.3SG.F
'I wrote it'

b. ji -ntris
SBJ.3SG.M -be.filled.IMPF
'it is filled'

c. ji -ndxil
SBJ.3SG.M -be.entered.IMPF
'it is/can be entered'

d. ji -stmarː
SBJ.3SG.M -continue.IMPF
'he continues'

Otherwise, it is only permitted in fast speech, as shown in (197), when the stem vowel is syncopated as marked by the brackets.

(197) Medial CVCC in Fast Speech (Glover 1988, p. 63)

a. targ(a)m -u
translate.PFV -SBJ.3PL.M
'they translated'

b. ji -staʕm(a)l -u
SBJ.3-use.IMPF -SBJ.3PL.M
'they use'

Muscat permits medial CVːC as shown in (198-a)–(198-c), but undergoes standard CSS-MORPH as shown in (198-d).

(198) Medial CVːC (Glover 1988, p. 60)

a. xtaːr -∅ -hum
choose.PFV -SBJ.3SG.M -OBJ.3PL.M
'he chose them'

b. goːhra

'the jewel'

c. qandiːl-na
lantern -POSS.1PL
'our lantern'

d. xtar -t
choose.PFV -SBJ.1SG
'I chose'

Muscat permits medial CVCː and CVːCː syllables, as shown in (199).

(199) Medial CV(ː)Cː (Glover 1988, p. 60)

a. maʃaːbː -ha
fan.PL -POSS.3SG.F
'her (hand) fans'

b. ʃaːlː -ha
carry.PTCP -OBJ.3SG.F
'carrying it'

Therefore, Muscat permits CVːC and CVCː without restriction, but does place restrictions on medial CVCC. This continues the same pattern as Algerian, Palestinian and Lebanese, with similar coronal restrictions as Palestinian and Lebanese but still some differences.

2.5.10 Sˤanʕaːniː

Sˤanʕaːniː permits medial CVCC in fast speech only, as shown in (200), where the stem vowel that is syncopated in fast speech is marked in brackets.

(200) Medial CVCC in Fast Speech (Watson 2002, p. 115)

a. ji -kt(a)sib
SBJ.3SG.M -earn.IMPF
'he earns'

b. ji -ft(a)hin
SBJ.3SG.M -rest.IMPF
'he rests, feels better'

c. ift(a)han -∅
rest.PFV -SBJ.3SG.M
'he rested, felt better'

Sˤanʕaːniː exhibits variation in tolerance of medial CVːC syllables. Epenthesis is optional after medial CVːC with the possessive suffix as shown in (201).[6]

(201) Optional Epenthesis after Medial CVːC Syllables with the Possessive (Watson 2002, p. 69)

[6]Watson (2002) notes epenthesis is particularly common with /k/ initial suffixes, that is, 2[nd] person possessives.

a. kitaːb(a) -naː
book -POSS.1PL
'our book'

b. ji -zuːr(a) -haː
SBJ.3SG.M -visit.IMPF -OBJ.3SG.F
'he visits her'

c. bayn(u) -hum
among -OBJ.3PL
'among them'

CSS-Morph does occur with the dative as in (202) but is optional with negation as in (203).

(202) Dative Closed Syllable Shortening (Watson 2002, p. 67)

a. /raːħ-l-ih/
raħ -∅ -l -ih
go.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M
'he went off'

b. /gaːl-l-iː/
gal -∅ -l -iː
say.PFV -SBJ.3SG.M -DAT -OBJ.1SG
'he said to me'

c. /jiguːl-l-naː/
ji -gul -la -naː
SBJ.3.M -say.IMPF -DAT -OBJ.1PL
'he says to us'

d. /jibiːʕ-l-ak/
ji -biʕ -l -ak
SBJ.3.M -sell.IMPF -DAT -OBJ.3SG.M
'he sells to you (m.s.)'

Note the optional long vowel is marked with brackets in (203), but this is distinct from the initial clitic [maː] which is optional independent of vowel length in the verb.

(203) Optional Negation Closed Syllable Shortening (Watson 2002, p. 67)

a. (maː) ka(ː)n -∅ -ʃ
NEG be.PFV -SBJ.3SG.M -NEG
'he was not'

b. (maː) dir -i(ː)t -ʃ
NEG be.PFV -SBJ.1SG -NEG
'I didn't know'

Sˤanʕaːniː also has variation in the acceptance of medial CVCː syllables. Postgeminate high vowel deletion does occur as shown in (204-a)–(204-c), but epenthesis as in (204-d) is optional. Note in (204-a)–(204-c) the underlying form is shown on the first line of each example, with the apparently degeminated surface form that is glossed on the second line.

(204) Medial CVCː syllables (Watson 2002, p. 73, 64)

a. /ragːaʕuː/
ragʕ -uː
sew.PFV -SBJ.3PL.M
'they (m.) sewed'

b. /jilabːisuː/
ji -labs -uː
SBJ.3 -dress.IMPF -SBJ.PL.M
'they (m.) dress (s.o.)'

c. /jiɟamːaʕayn/
ji -ɟamʕ -ayn
SBJ.3 -collect.IMPF -SBJ.PL.F
'they (f.) collect'

d. ħaɟː -a̲na
pilgrim -POSS.1PL
'our pilgrim'

Thus, Sˤanʕaːniː does permit medial CVCː and CVːC syllables (albeit optionally), but medial CVCC is only permitted in fast speech. As such, Sˤanʕaːniː seems to be following the same distribution as Algerian, Palestinian, Lebanese and Muscat, but the nature of the restrictions on CVCC is different.

2.5.11 Iraqi

Iraqi does not permit medial CVCC syllables, as shown by the epenthesis in (205).

(205) No Medial CVCC (Majdi 1988, p. 195)

kitab -i̱t -l -a
write.PFV -SBJ.1SG -DAT -OBJ.3SG.M

'she wrote to him'

However, it does permit medial CVːC syllables, as shown in (206).

(206) Medial CVːC (Erwin 1963, p. 20)

beːt -kum
house -POSS.2PL

'your house'

It also permits medial CVCː syllables, as shown in (207).

(207) Medial CVCː (Majdi 1988, p. 19, 136)

a. badːl -at
change.PFV -SBJ.3SG.F
'she changed'

b. t -badːl -iːn
SBJ.3.F -change.IMPF -SBJ.3SG.F
'you (f.) change'

c. dabːr -at
arrange.PFV -SBJ.3SG.F
'she arranged'

Therefore, Iraqi permits CVːC and CVCː without restriction, but not medial CVCC. This is the same general pattern as described for Algerian, Palestinian, Lebanese, Muscat and SˤanʕaːniØ, but it is more conservative in that no CVCC is permitted medially. While this set of varieties do all differ in terms of the nature of the restriction placed on CVCC syllables, the core pattern persists throughout, and is distinct from the first group (Tunisian,Wadi Ramː, Rufaidah, Rwaili, Haːyil) where there were no restrictions.

2.5.12 Qatari

Qatari follows a different pattern to the varieties discussed so far.

It does not permit medial CVCC. Where these could be created morphologically, assimilation of the /h/ of non-subject pronominal formatives to the preceding consonant occurs, as shown in (208). See Section 4.6.4 and Glover (1988) for discussion of this.

(208) No Medial CVCC (Al-Sulaiti 1993, p. 155)

/niʃad-t-ha/
niʃa -t -ta
ask.PFV -SBJ.1SG -OBJ.3SG.F

'I asked her'

Qatari permits medial CVːC without restriction, as shown in (209). Note the assimilation of [h] also occurs in (209-b).

(209) Medial CVːC (Al-Sulaiti 1993, p. 113, 155, 233)

a. xanaːfs -iːn
beetle -PL
'beetles'

b. /gmaːʃ -hum/
gmaːʃ -ʃum
pearl.PL -POSS.3PL

'their pearls'

c. ktaːb-na
book -POSS.1PL
'our book'

Qatari does not permit medial CVCː, so does not undergo postgeminate high vowel deletion as shown in (210-a) and epenthesises after geminates as shown in (210-b)–(210-c).

(210) No Medial CVCː (Al-Sulaiti 1993, p. 328, 241)

a. wi -j -tikalːim -oːn
and -SBJ.3 -talk.IMPF -SBJ.PL
'and they talk'

b. kilː- i̲hum
all -OBJ.3PL
'all of them'

c. radː -∅ -i̲ha
return.PFV -SBJ.3SG.M -OBJ.3SG.M
'he returned it'

Therefore, Qatari permits medial CVːC but not CVCː or CVCC. In not permitting CVCː at all, it is distinct from the previous varieties discussed which permit syllables with long segments without restriction.

2.5.13 Makːan

Makːan continues a similar pattern to Qatari.

Makːan does not permit medial CVCC syllables. It will epenthesise to break up underlying CVCC syllables as in (211-a)–(211-b) or block the syncope if it would create CVCC syllables as in (211-c)–(211-d).

(211) Makːan No Medial CVCC (Kabrah 2004, p. 69, 124)

a. kalba -kum
dog -POSS.2PL
'your dog'

b. ruħ -ta -lː -u
go.PFV -SBJ.1SG -DAT -OBJ.3SG.M
'I went to him'

c. muʃkila
*muʃkla
'problem'

d. *tiksru
ti -ksir -u
SBJ.3SG.F -break.IMPF -OBJ.3SG.M
'she breaks it'

Makːan displays mixed behaviour with regard to CVːC syllables. It will use epenthesis to break up underlying CVːC syllables, as in (212-a)–(212-b), or undergo Closed Syllable Shortening as in (212-c). However, it will create surface CVːC syllables through syncope, as in (213).

(212) Epenthesis Breaking Underlying Medial CVːC (Kabrah 2004, p. 61, 80)

a. /baːb-ha/

baːba -ha
door -POSS.3SG.F
'her door'

b. ʃaːla -∅ -ha
carry.PFV -SBJ.3SG.M -OBJ.3SG.
'he carried her'

c. /maktuːblu/

maktub -l -u
be.written.M -DAT -OBJ.3SG.M

'it was destined/written for him'

(213) Surface medial CVːC created by Syncope (Kabrah 2004, p. 123)

a. /sˤaːħib-i/

sˤaːħb -i
friend.M -POSS.1SG

'my male friend'

b. /sˤaːħibatu/

sˤaːħbat -u
friend.F -POSS.3SG.M

'his girl friend'

c. /jisaːfiru/

ji -saːfr -u
SBJ.3 -travel.IMPF -SBJ.PL

'they travel'

Makːan does not permit medial CVCː syllables, so epenthesises after clusters as in (214-a) and blocks syncope that would create them as in (214-b)–(214-c).

(214) Medial CVCː (Kabrah 2004, p. 69, 124)

a. katab -ta -lːa -hum
write.PFV -SBJ.1SG -DAT -OBJ.3PL
'I wrote for them'

b. mudarːis -i
teacher -POSS.1SG
'my teacher'

c. *mudarːsi

'my teacher'

Therefore, Makːan permits some CVːC syllables, but not CVCC or CVCː. This is similar to Qatari, in that CVCC and CVCː syllables are treated the same way and separate from CVːC syllables, although there are more restrictions on the latter group in Makːan than Qatari.

2.5.14 Cairene

Cairene is similar to Qatari and Makːan.

Cairene does not permit medial CVCC syllables, so breaks up triconsonantal clusters with epenthesis to the right of the central consonant in the cluster. This is shown in (215).

(215) No Medial CVCC (Watson 2002, p. 64)

ʔul -ti̱ -l -ak
say.PFV -SBJ.1SG -DAT -OBJ.2SG

'I told you'

Cairene does not permit medial CVːC syllables, so when these are created by affixation they undergo closed syllable shortening. In (216-a) and (216-c) we see the word-final CVːC syllables, but in (216-b) and (216-d) these have become word medial through affixation so undergo shortening.

(216) No Medial CVːC (Watson 2002, p. 66)

a. baːb
'door'

b. bab -kum
door -POSS.2PL
'your (pl.) door'

c. joːm
'day'

d. jom -ha
day -POSS.3SG.F
'her day'

Cairene does not permit medial CVCː syllables, so follows them with an epenthetic vowel, as shown in (217-a), and does not undergo postgeminate high vowel deletion, as shown in (217-b).

(217) No Medial Tautosyllabic CVCː (Watson 2002, p. 64)

a. kulːu -hum
all -OBJ.3PL.M
'all of them'

b. ti -kalːim -i
SBJ.2 -speak.IMPF -SBJ.SG.F
'you (f.s.) speak'

Thus, Cairene does not permit any CVXC syllables word medially. While this may initially seem a different pattern to Qatari and Makːan, crucially the CVCC and CVCː syllables are being treated the same. That no CVːC syllables are permitted finishes a continuum from Qatari with Makːan in the middle as to the permissibility of this type.

2.5.15 Moroccan Casablanca

Moroccan Casablanca behaves differently to the varieties discussed above.

Moroccan Casablanca permits medial CəCC if there is a fall in sonority, as shown in (218).

(218) Medial CəCC with sonority fall (Boudlal 2001, p. 94, 85)

a. sənsla

'zip'

b. tənbr -i
stamp -POSS.1SG
'my stamp'

c. ʒərʒr -u
trail.PFV -SBJ.3PL.M
'they trailed'

If there is a sonority rise, there is metathesis instead from CVCC to CCVC. This is particularly clear in (219-b), where the syllable structure has changed in the plural compared to the singular form in (219-a). However, this is not solely a matter of plural templates. Nouns in this template take the form CəCC unless this would create a sonority rise, at which point the underlying form is then CCəC. This is not metathesis exactly, but rather reflects a constraint on the underlying form of the words.

(219) Metathesis to avoid medial CVCC with sonority rise (Boudlal 2001, p. 93-94)

a. məsləm
   Muslim

b. msəlm -in
   Muslim -PL
   ‘Muslims’

c. mtˤərqa
   ‘hammer’

Moroccan Casablanca does not have any long vowels (Vː) but does have a length opposition between short vowels (V) and schwa (ə), where schwa does not contribute any mora as demonstrated in (220).

(220) Moroccan Casablanca Schwa vs Short Vowel Quantity Difference (Boudlal 2001, p. 109)

a. məl.ˈjun
   ‘1 million’

b. ˈməl.məl
   ‘he shook’

The best way to represent this asymmetry between length and weight is using X-slots, where ‘short’ vowels are associated to two X-slots and one mora, whereas schwa is associated to one X-slot and contributes no mora. In this way, the behaviour of short vowel reduction to schwa, and where consonant clusters and short vowels are permitted, can be accounted for in terms of numbers of X-slots.

In cases that would undergo CSS-Morph in other varieties, the short vowel is reduced to a schwa as shown in (221-c).

(221) Medial CVC (Boudlal 2001, p. 172)

a. samħ -u
forgive.PFV -SBJ.3PL
'they forgave'

b. bas -∅ -na
kiss.PFV -SBJ.3SG.M -OBJ.1PL
'he kissed us'

c. bəs -t
kiss.PFV -SBJ.1SG
'I kissed'

Medial CəCC is permitted in Moroccan Casablanca, as shown in (222). There is no data to show whether medial CVCː or CVCC syllables are permitted or not in Boudlal (2001).

(222) Medial CəCː (Boudlal 2001, p. 184)

a. ʔaməzːru
'Amezrou'

b. məzːriwi
'from Amezrou'

Thus, Moroccan Casablanca appears to permits CəCC and CəCː syllables, but not CVːC, and it is unclear if CVCː or CVCC is permitted. This is a different pattern to those described in the varieties above.

2.5.16 Libyan Tripoli

Libyan Tripoli exhibits a similar pattern to Moroccan Casablanca.

Libyan Tripoli permits medial CəCC syllables, as shown in (223).

(223) Medial CəCC (Yoda 2005, p. 104)

a. kəlb -kəm
dog -POSS.2PL

‘your (pl.) dog’

b. xəbz -na
bread -POSS.1PL
‘our bread’

c. dərkb -u
hit.drum.PFV -SBJ.3PL.M
‘they hit the drum’

d. sənsla

‘necklace’

e. bərʧqal

‘oranges’

There are some cases of CVCC, but epenthesis does optionally occur to break these up, as shown in (224) where optionality is marked by brackets. That some instances of CVCC are permitted may be due to coronality of the final consonant, but there is insufficient data to be conclusive on this point.

(224) Optional Epenthesis to break CVCC (Yoda 2005, p. 70, 107)

a. qal -(ə̲)t -l -u
say.PFV -SBJ.3SG.F -DAT -OBJ.3SG.M
‘she said to him’

b. katʃb -ətʃ -(ə̲)l -kəm
write.PFV -SBJ.3SG.F -DAT -OBJ.2PL
‘she wrote to you (pl.)’

Like Moroccan Casablanca, Libyan Tripoli does not have any long vowels (Vː) but does have a length opposition between short vowels (V) and schwa (ə), where schwa does not contribute any mora. This is clear in (225). If schwa and short vowels contributed weight equally, we would not expect any difference in stress between CəCCəC and CVCCVC words. However, in (225) we find just that — in (225-a), the vowels are both schwa and receive initial stress, whereas in (225-b) they are both short vowels and receive final stress. CəCCəC behaves like CVCV, so has the same initial stress as in (225-c). This can best be accounted for by the coda consonant providing weight by position, but not the schwa, thus making both CəCCəC and CVCV words with two monomoraic syllables.

(225) Libyan Tripoli Short Vowel and Schwa Weight Asymmetry (Yoda 2005, p. 111-112)

a. ˈtˤəbːəħ
call.IMP.M
'call (m.)!'

b. sakˈn -in
live.PTCP -M.PL
'living (m.pl.)'

c. ˈħuʃa
'house'

The best way to represent this asymmetry between length and weight is using X-slots, where 'short' vowels are associated to two X-slots and one mora, whereas schwa is associated to one X-slot and contributes no mora. In this way, the behaviour of short vowel reduction to schwa, and where consonant clusters and short vowels are permitted, can be accounted for in terms of numbers of X-slots.

Medial CVC syllables are permitted, as shown in (226).

(226) Medial CVC (Yoda 2005, p. 203, 111)

a. sanyəʧ
'field'

b. sakˈn -in
live.PTCP -M.PL
'living (m.pl.)'

Libyan Tripoli permits medial CəCː as shown in (227), though Yoda (2005, p. 104) states degemination can occur in free variation.

(227) Medial CəCː (Yoda 2005, p. 104)

a. j -ʕəlːm -u
SBJ.3 -teach.IMPF -SBJ.PL.M
'they (m.) teach'

b. msəkːr -in
close.PTCP -PL
'closed'

c. qəlːq -u
bore.PFV -SBJ.3PL.M
'they bored (s.o.)'

Therefore, Libyan Tripoli permits medial CəCC and CəCː without restriction, CVCC optionally, but not CVCː. This is similar to what is found in Moroccan Casablanca.

2.5.17 Summary

From the data above, it is clear that there is variation in the acceptance of medial CVXC syllables. This variation is summarised in Table 2.11 below.

Table 2.11: Tolerance of Medial CVXC Syllables

Key: Y = Yes, N = No, - = Not Applicable,
N_E = No with Exceptions, Y_E = Yes with Exceptions

Variety	CVCC	CVːC	CVCː
Wadi Ramː	Y	Y	Y
Tunisian	Y	Y	Y
Rufaidah	Y	Y	Y
Rwaili	Y	Y	Y
Haːyil	Y	Y	Y
Palestinian	N_E	Y	Y
Sˤanʕaːniː	N_E	Y	Y
Lebanese	N_E	Y	Y
Muscat	N_E	Y	Y
Algerian	N_E	Y	Y
Iraqi	N	Y	Y
Qatari	N	Y	N
Makːan	N	N_E	N
Cairene	N	N	N
Moroccan Casablanca	Y_E	-	Y
Libyan Tripoli	Y	-	Y

To date, the literature has claimed that the only CVXC syllable found word medially is CVːC but only in some varieties. Contrary to these claims, a much broader range of CVXC syllables is found word medially in this dataset.

Futhermore, there are clear groupings in tolerance of these syllables. Wadi Ramː, Tunisian, Rufaidah, Rwaili, and Haːyil permit CVCC, CVːC and CVCː syllables medially without restriction. Palestinian, Sˤanʕaːniː, Lebanese, Muscat, Algerian, and Iraqi permit CVːC and CVCː syllables without restriction, but have a range of restrictions on CVCC syllables. Qatari, Makːan, and Cairene do not permit medial CVCC and CVCː, but have a range of restrictions on medial CVːC. Moroccan Casablanca and Libyan Tripoli do not have long vowels, but rather a distinction between short vowels and schwa, so do permit CVCː and CVC word medially, but differ in placing restrictions on CVCC.

From this distribution, we see variation in how CVːC, CVCC and CVCː syllables pattern together. Where not all of these syllables are permitted word medially, in some dialects CVCː and CVːC pattern together (Palestinian, Sˤanʕaːniː, Lebanese, Muscat, Algerian, Iraqi, Moroccan Casablanca, Libyan Tripoli) but in others CVCː patterns with CVCC (Qatari, Makːan, Cairene). Notably, what is not found is varieties that permit CVːC and CVCC but not CVCː.

The geographical distribution of this behaviour is not particularly clear. Rufaidah, Rwaili, and Haːyil are all spoken in Saudi Arabia, with Wadi Ramː in Jordan near the Saudi border. However, Tunisian behaves in the same way and is much further away in North Africa; and there are assorted varieties much closer that behave differently. A similar problem can be found with the next set — Palestinian, Lebanese and Iraqi are fairly close together, and indeed often grouped together in typologies of Arabic; however, Sˤanʕaːniː is further away in Yemen, Muscat is similarly much further south, and Algerian is further away in North Africa. Qatari, Makːan and Cairene are not quite as widely geographically distributed as some groups, but are hardly close. Moroccan Casablanca and Libyan Tripoli are both North African varieties, but separated by Tunisia and Algeria, which behave differently. Therefore, it is not clear what extralinguistic factors motivate this distribution. It may be that it is inherited with patterns of population movement accounting for the vast geographical distribution, but such an analysis requires a larger dataset and in depth historical information which is not available within the scope of this study.

However, even if the historical development of this pattern requires further work than is within the scope of this study, a phonological analysis of the synchronic distribution is warranted. In the next section, I develop a proposal for TOLERANCE.

2.6 New Proposal

Arabic varieties differ in terms of what types of syllables they permit, and this variation is wider and more systematic than has been previously recognised. Existing typological proposals fail to account for the medial CVXC patterns displayed in this chapter. By CVXC I refer to syllables where the X can be a vowel or consonant, thus subsuming CVːC, CVCː and CVCC. Whilst adjunction to mora has attempted to account for the presence of medial CVːC syllables, whether it can also be used for CVCː syllables is unclear. In any case, this doesn't account for the presence of some restricted CVCC syllables.

I propose that a bifurcated typology is needed to account for the distribution of syllable structure, as the behaviours of medial superheavy syllables and word-edge consonant clusters do not correlate. I name the two branches of this typology TOLERANCE and REPAIR. In this section I will first lay out the nature of TOLERANCE and how this accounts for the varieties explored here, before demonstrating that REPAIR is also needed, and that this parameter governs both the direction of epenthesis and the presence of word-edge consonantal clusters.

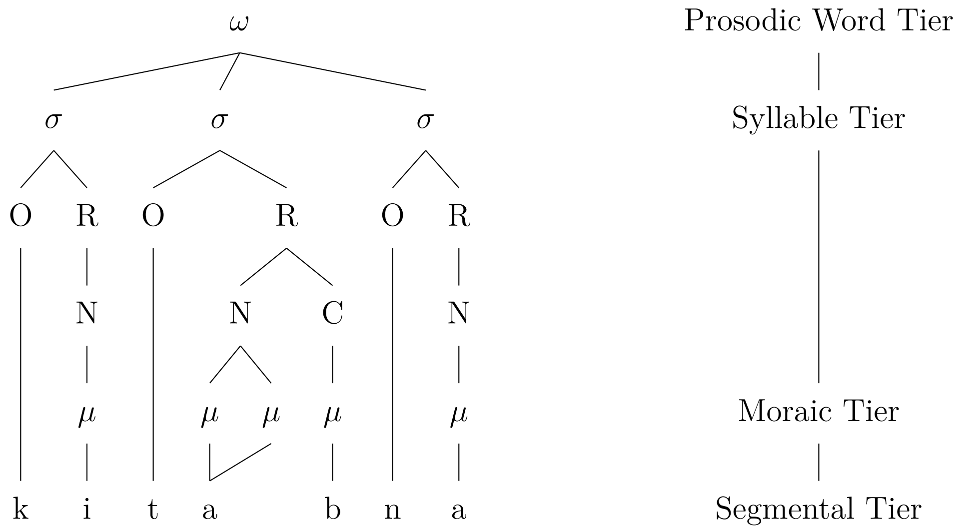
2.6.1 Tolerance

In Section 2.5, I demonstrated that these varieties permit more CVXC syllables than predicted in the literature. These findings were summarised in Table 2.11.

From this we see that where there are asymmetries in which type of CVXC syllable is permitted, CVːC and CVCː pattern together (). Rather than use mora adjunction, I propose that restrictions on syllable structure at different tiers of the Prosodic Hierarchy can better account for these patterns. This approach not

Figure 2.8: Syllable Structure for [kitaːb-na] with Moraic and Segmental Tiers

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora



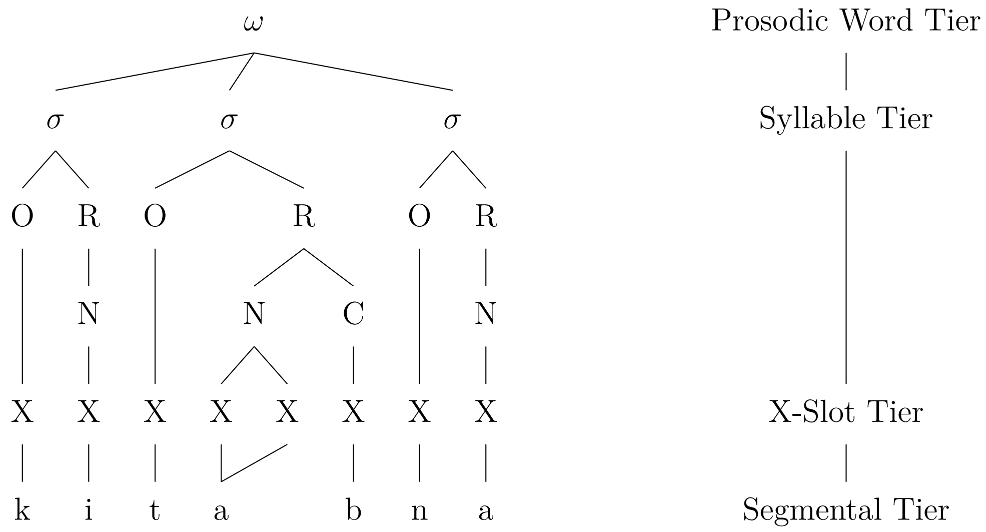
only correctly predicts that CVːC and CVCː should pattern together, as they both involve long segments, but also has the flexibility to permit extra segments in syllables if they fit variety-specific restrictions. As such, this approach fits the aims of Canonical Typology in demarcating the canonical space, but also provides the mechanisms that allow individual varieties to be analysed.

Recall that we can represent syllable structure as in the diagram for [kitaːb-na], ‘our book’, in Figure 2.8. Here, the long vowel has two moras associated to one segment. Exploiting this asymmetry in representation is key to the proposal here.

Varieties that permit long segments can have rime restrictions on the *segmental* tier, thus permitting long or short segments because there is no difference between them on this tier, whereas varieties that do not permit long segments can have (rime) restrictions on the *moraic* tier, as here length is represented. However, some varieties require different representations for weight and length. In this dataset, Libyan Tripoli and Moroccan Casablanca have vowels that are associated to one mora but two X-slots, as argued above, and elaborated on in the next chapter. Their X-slot based syllable structure is depicted in Figure 2.9 below.

Figure 2.9: Syllable Structure for [kitab-na] with X-slot and Segmental Tiers

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, X = X-slot



The varieties that allow long consonants or clusters but not long vowels have rime restrictions on the *X-slot* tier. This assumes that the short vowels in these varieties occupy two X-slots, as motivated above.

This division of syllable restrictions on the segmental, moraic and X-slot tiers accounts for the broad distribution of medial CVXC syllables found, and as such demarcates the canonical space. Varieties differ in whether they permit two or three elements in their particular tier of syllable structure restrictions. However, to make this representation sufficiently flexible it requires a mechanism for how the non-canonical structures arise — in particular, the rare but still surfacing CVCC cases.

Thus, I propose that in addition to a canonical two or three segments, moras or X-slots, individual varieties can also permit an extra element in their medial syllable rime according to variety-specific restrictions, thus allowing for some restricted consonantal clusters. These restrictions can include coronality, fast speech, and sonority. As such, this typology allows for mechanisms to account for the data in each and every variety, rather than merely providing a canonical set of features that not all fit.

This analysis predicts that varieties with restrictions on the segmental tier will not undergo vowel shortening or epenthesis to split up the syllable as there is no difference on the segmental level between CVC and CVːC. However, there is a difference on the moraic tier, so we would expect shortening/epenthesis to occur in varieties with moraic restrictions.

In subsection 2.6.2 I illustrate how this works in practice, starting with the varieties with segmental restrictions, then moraic, then X-slot.

2.6.2 Analysis of Sixteen Varities

2.6.2.1 Algerian

Algerian permits medial CVːC and CVCː without restriction as in (228-a)–(228-b), but permits CVCC only if there is falling sonority as in (228-c)–(228-d).

(228) Algerian Syllable Restrictions (Bouhadiba 1988, p. 313, 311, 294)

a. ħmaːr. -na
donkey -POSS.1PL
'our donkey'

b. ʃafː -∅ -na
inspire.pity.PFV -SBJ.3SG.M -OBJ.1PL
'he inspired pity in us'

c. qanq.ra
'toad'

d. hatu̲rf. -∅ -u
dream.PFV -SBJ.3SG.M -OBJ.3SG.M
'he dreamt it'

Therefore, Algerian has restrictions on the *segmental* tier, permitting two segments in the rime with an extra third if it fits sonority restrictions.

2.6.2.2 Palestinian

Palestinian permits CVːC and CVCː syllables without restriction, but only permits CVCC if C2 is coronal, as shown in (229-a), (229-b) and (229-c)–(229-d) respectively.

(229) Palestinian Medial Syllable Restrictions (Abu-Salim 1982a, p. 22, 53, 196)

a. ʃubːaːk -hum
window -POSS.3PL.M
'their window'

b. ʔimː. -na
mother -POSS.1PL
'our mother'

c. ʔibi̠n. -hum
son -POSS.3PL.M
'their son'

d. ʔuxt. -hum
sister -POSS.3PL.M
'their sister'

Thus Palestinian has syllable restrictions on the *segmental* tier, permitting two segments in the rime, and a third only if it is coronal.

2.6.2.3 Lebanese

Lebanese permits medial CVːC and CVCː syllables without restrictions as in (230-a-b), but only permits CVCC syllables if C2 is coronal as in (230-c)–(230-d).

(230) Lebanese Medial Syllable Restrictions (Haddad 1984, p. 10, 158, 40)

a. ʕaːm.le
'working class'

b. sitː -na
grandmother -POSS.1PL
'our grandmother'

c. sakːar -∅ -i̠l. -na
close.PFV -SBJ.3SG.M -DAT -OBJ.1PL
'he closed for us'

d. mist.ɣilː -ni
take.advantage.of.PTCP.M -OBJ.1SG
'he took advantage of me'

Thus, Lebanese has syllable structure restrictions on the *segmental* tier, permitting two segments in the rhyme, with a third if it is coronal, just as found for Palestinian.

2.6.2.4 Iraqi

Iraqi permits medial CVːC and CVCː syllables without restriction, but no CVCC syllables, as shown in (231).

(231) Iraqi Syllable Restrictions (Erwin 1963, p. 20) (Majdi 1988, p. 19, 195)

a. beːt. -kum
house -POSS.2PL
'your (pl.) house'

b. badː.l -at
change.PFV -SBJ.3SG.F
'she changed'

c. ki.ta.b -i̱t. -l -a
write.PFV -SBJ.1SG -DAT -OBJ.3SG.M
'she wrote to him'

Given this asymmetry, Iraqi has restrictions on the *segmental* tier, permitting two segments in the rime. Unlike many other varieties explored here, it does not permit a third segment in the rime in any situation.

2.6.2.5 Wadi Ramː

Wadi Ramː permits medial CVːC, CVCː, CVCC and CVːCC syllables, as shown in (232).

(232) Wadi Ramː Syllable Restrictions (Al Mashaqba 2015, p. 112, 198, 208, 120)

a. maːs.k -iːn
catch.PTCP -PL.M
'they (m.) are catching'

b. mitʕalː.m -iːn
educate.PTCP -PL.M
'educated (m.pl.)'

c. ʤib -t. -l -ak
bring.PFV -SBJ.1SG -DAT -OBJ.2SG
'I brought to you'

d. ʤaːb -l. -na
bring.PFV -DAT -OBJ.1PL
‘he brought to us’

This can best be accounted for with restrictions on the segmental tier. If the restrictions were on the moraic tier, CVːCC would require four mora to be produced. If four mora were permitted in the syllable rime, there is no reason why unattested medial CVCCC syllables could not be produced. Furthermore, if trimoraic syllables are cross-linguistically rare and controversial, quadrimoraic syllables are unlikely and such a proposal would require strong evidence.

Therefore, Wadi Ramː has restrictions on the *segmental* tier, permitting three segments in the rime.

2.6.2.6 Muscat

Muscat has lexical category and sonority based restrictions on CVCC syllables, whereby they can only surface in nouns in the case of a sonorant–obstruent cluster or optionally where C2 is a dentalveolar fricative as in (233-a-b), but in verbs can surface if C2 is coronal as in (233-c); otherwise CVCC can only surface in fast speech as in (233-d) or not at all as in (233-e).

(233) Muscat CVCC restrictions (Glover 1988, p. 223-225, 63)

a. ward. -ha
flower.PL -POSS.3SG.F
‘her flowers’

b. kan(i̠)z. -ha
pressed.dates.PL -POSS.3SG.F
‘her pressed dates’

c. ka.tab -t. -ha
write.PFV -SBJ.1SG -OBJ.3SG.F
‘I write it’

d. targ.m -u
translate.PFV -SBJ.3PL.M
‘they translated’ [Fast Speech]

e. ʕa.ʃu̠b. -hum
grass -POSS.3PL.M
‘their grass’

However, CVːC, CVCː, and CVːCː syllables are permitted without restriction, as shown in (234).

(234) Muscat CVːC, CVCː and CVːCː Syllables (Glover 1988, p. 135, 206, 60)

a. goːh.ra
'jewel'

b. stamarː -∅
continue.PFV -SBJ.3SG.M
'he continued'

c. maʃaːbː -ha
fan.PL -POSS.3SG.F
'her fans'

Therefore, Muscat has restrictions on the *segmental* tier, permitting two segments in the rime, with a third if it meets sonority, lexical category, or speech tempo restrictions.

2.6.2.7 Sˤanʕaːniː

Sˤanʕaːniː permits CVːC syllables — though does insert an epenthetic vowel after them in free variation as shown in (235-a). Similarly, it permits CVCː syllables through postgeminate high vowel deletion, but does optionally epenthesise after them as shown in (235-b). However, CVCC syllables are only permitted in fast speech, as shown in (235-c). As above, optionality is marked with brackets.

(235) Sˤanʕaːniː Mixed Medial CVXC Permissions (Watson 2002, p. 69, 73, 115)

a. kitaːb(a̲) -na
book -POSS.1PL
'our book'

b. rag(ː).ʕ -u
sew.PFV -SBJ.3PL.M
'they (m.) sewed'

c. ift(a̲)han -∅
rest.PFV -SBJ.3SG.M
'he rested, felt better'

While CVːC and CVCː are optionally broken by epenthesis, that they are permitted at all suggests that Sˤanʕaːniː has restrictions on the *segmental* tier, permitting two segments in the rime with a third only in fast speech.

2.6.2.8 Makːan

Makːan behaves differently to the varieties with segmental restrictions discussed above. Makːan does not permit CVCC syllables, CVCː syllables, or underlying CVːC syllables, but does permit CVːC syllables created by syncope.

(236) Makːan Syllable Restrictions (Kabrah 2004, p. 69, 61, 123)

a. kalba -kum
dog -POSS.3PL
'your dog'

b. katab -ta -lla -ha
write.PFV -SBJ.1SG -DAT -OBJ.3SG.F
'I wrote to/for her'

c. baːba -ha
door -POSS.3SG.F
'her door'

d. sˤaːħb -i
friend.M -POSS.1SG
'my male friend'

Therefore, Makːan has restrictions on the *moraic* tier, permitting two moras with a third optionally if it is part of a branching nucleus.

2.6.2.9 Qatari

Qatari permits medial CVːC without restriction, but not CVCː or CVCC, as shown in (237).

(237) Qatari Syllable Restrictions (Al-Sulaiti 1993, p. 233, 144, 241)

a. ktaːb. -na
book -POSS.1PL
'our book'

b. [niʃatːa]

niʃad -t -ha
ask.PFV -SBJ.1SG -OBJ.3SG.F

‘I asked her’

c. radː -∅ -i̲ha
return.PFV -SBJ.3SG.M -OBJ.3SG.M
‘he returned it’

Since there is no asymmetry between closed syllables with long segments and those without, Qatari has restrictions on the *moraic* tier, permitting two moras in a syllable, but allowing a third if it is part of a branching nucleus, just as in Makːan.

2.6.2.10 Cairene

Cairene does not permit any medial CVXC syllables, as shown in (238), where the final CVːC in (238-a) is shortened when word-medial in (238-b), and the CVCː and CVCC in (238-c) and (238-e) are followed by epenthetic vowels when word-medial.

(238) Cairene No Medial CVXC (Watson 2002, p. 64-66)

a. baːb

‘door’

b. bab -kum
door -POSS.3PL
‘your (pl.) door’

c. kulːu̲ -hum
all -OBJ.3PL.M
‘all of them’

d. bint

‘daughter’

e. bin.ti̲-na
daughter -POSS.1PL
‘our daughter’

Thus, Cairene has syllable restrictions on the *moraic* tier, permitting two moras only, therefore no surface word medial CVXC syllables.

2.6.2.11 Tunisian

Tunisian permits medial CVːC syllables, CVCː and CVCC syllables regardless of sonority as shown in (239).

(239) Tunisian CVːC, CVCː and CVCC Syllables (Maamouri 1967, p. 5, 163, 22)

a. riːħ. -na
wind -POSS.1PL
'our wind'

b. farː.q -it
distribute.PFV -SBJ.3SG.F
'she distributed'

c. sukm.t -ik
cold -POSS.2SG
'your cold'

However, Tunisian does not permit CVːCC syllables, as shown in (240).

(240) Tunisian No CVːCC Syllables (Maamouri 1967, p. 131)

a. *xaːl-k

b. xaː.li̱k
uncle -POSS.2SG
'your uncle'

If Tunisian had restrictions on the segmental tier, we would not expect epenthesis to occur in (240-b), as on the segmental level there is no difference between CVCC and CVːCC. However, this difference can be seen on the moraic level instead.

Thus, Tunisian has restrictions on the *moraic* tier, permitting three moras in the rime. Bar the number of moras in the rime, this is the same pattern as in Cairene above.

2.6.2.12 Rufaidah

Rufaidah permits medial CVːC, medial CVCC albeit with optional stem syncope, and medial CVCː syllables though with optional postgeminate high vowel deletion, as shown in (241-a), (241-b-c) and (241-d) respectively.

(241) Rufaidah Syllable Restrictions (Prochazka Jr. 1988, p. 198, 200, 32, 46)

a. beːt. -hum
house -POSS.3PL.M
'their house'

b. ʔibn. -hum
son -POSS.3PL.M
'their son'

c. ju -kt(u)b -uːn
SBJ.3 -write -SBJ.PL.M
'they write'

d. j -salː(i)m -uːn
SBJ.3 -hand -SBJ.PL.M
'they hand'

Therefore, Rufaidah has restrictions on the *moraic* tier, permitting three moras, as there is no asymmetry in the treatment of long segments and singleton clusters. This is the same as in Tunisian above.

2.6.2.13 Rwaili

Rwaili permits medial CVːC, medial CVCC though with optional stem syncope, and medial CVCː syllables, as shown in (242).

(242) Rwaili Syllable Restrictions (Prochazka Jr. 1988, p. 51, 34, 41)

a. gaːbl -i
meet.IMP -SG.F
'meet (f.s.)!'

b. ta -ktb -uːn
SBJ.2 -write.IMPF -SBJ.PL.M
'you (pl.) write'

c. t -badːl -iːn
SBJ.2 -change.IMPF -SBJ.PL.M
'you (pl.) change'

Therefore, Rwaili has restrictions on the *moraic* tier, permitting three moras, as there is no asymmetry between the types of CVXC syllable, just as in the varieties above.

2.6.2.14 Haːyil

Haːyil permits medial CVːC, medial CVCC though with optional stem syncope, and medial CVCː syllables, as shown in (243).

(243) Haːyil Syllable Restrictions (Prochazka Jr. 1988, p. 51, 34, 50)

a. gaːbl -i
meet.IMP -SG.F
'meet (f.s.)!'

b. ta -ktb -uːn
SBJ.2 -write.IMPF -SBJ.PL.M
'you (pl.) write'

c. kasː.r -u
break.IMP -PL.M
'break (m.pl.)!'

Therefore, Haːyil has restrictions on the *moraic* tier, permitting three moras.

2.6.2.15 Moroccan Casablanca

Moroccan Casablanca and Libyan Tripoli are different to the varieties discussed above, and require a different analysis.

Moroccan Casablanca permits medial CəCC if there is a fall in sonority as in (244-a), as well as medial CəCː as shown in (244-b).

(244) Moroccan Casablanca Medial CəCC and CəCː (Boudlal 2001, p. 94, 184)

a. səns.la
'zip'

b. məzː.ri.wi
'from Amezrou'

Moroccan Casablanca permits CVC but there is no evidence of CVCː or CVCC being permitted, and indeed there are cases of vowel reduction to CəCC, as shown in (245).

(245) Medial CVC(X) (Boudlal 2001, p. 172)

a. bas. -∅ -na
kiss.PFV -SBJ.3SG.M -OBJ.1PL

'he kissed us'

b. bəs -t
kiss.PFV -SBJ.1SG
'I kissed'

Government Phonology analyses of Moroccan Casablanca syllable structure denote [ə] as filling one timing slot but short vowels as filling two (Lahrouchi and Ridouane 2016; Lahrouchi 2018). There is a disparity between length and weight for Moroccan Casablanca, as while /v/ may fill two timing slots, it only contributes one mora, whereas /ə/ contributes no mora as summarised in Table 2.12 below. The impact of short vowels and schwa on stress is demonstrated in (246), where the first two words both contain two closed syllables, but there is a stress difference where (246-a) ends in CVC but (246-b) ends in CəC. We could still hypothesise that schwa carries one mora and short vowels carry two at this stage. However, for such an analysis to hold we would expect final stress on (246-c), assuming two bimoraic syllables. However, we find initial stress here. Furthermore, the amoracity of the schwa has further support in that the schwa never occurs in open syllables, which can be accounted for as a prohibition on amoraic syllables rather than monomoraic syllables (Bensoukas 2012), and that in this way verbs such as [kərkəb], 'roll', and [sədd], 'close', both fit a bimoraic template consistently rather than being quadrimoraic and trimoraic respectively (Boudlal 2001).

Table 2.12: Moroccan Casablanca Length and Weight Disparity

	Timing Slots	Mora
V	2	1
ə	1	0

(246) Moroccan Casablanca Schwa vs Short Vowel Quantity Difference (Boudlal 2001, p. 109, 311, 115)

a. məl.ˈjun

'1 million'

b. ˈməl.məl

‘he shook’

c. ˈwa.lu

‘nothing’

Diagrams for this proposed structure can be seen below.

Figure 2.10: Moroccan Casablanca Syllable Structure with X-Slots

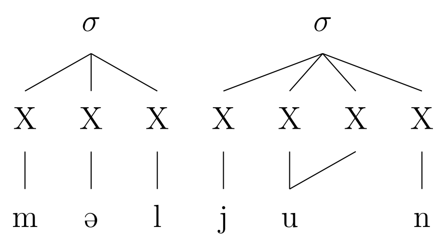
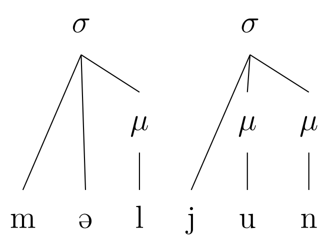


Figure 2.11: Moroccan Casablanca Moraic Structure



Given this disparity, the syllable restrictions for Moroccan Casablanca are best expressed as restrictions on the *X-slot* tier. Moroccan Casablanca permits three X-slots in the syllable rime, though often places sonority restrictions on the final slot. This accounts for the lack of overlong vowels that would take up four timing slots if ‘short’ vowels take two, as well as the reduction of short vowels to schwa when a consonant is added as shown in (246).

2.6.2.16 Libyan Tripoli

Libyan Tripoli permits medial CəCC, CVCC with optional epenthesis and CəCː as shown in (248), but not CVːC syllables as, like Moroccan Casablanca, it has no bimoraic vowels. This is clear from (247). If schwa and short vowels contributed weight equally, we would not expect any difference between CəCCəC and CVCCVC words. However, in (247) we find just that — in (247-a), the vowels are both schwa and receive initial stress, whereas in (247-b) they are both short vowels and receive final stress. CəCCəC behaves like CVCV, so has the same initial stress as in (247-c). This can be best accounted for by the coda consonant providing weight by position, but not the schwa, thus making both CəCCəC and CVCV words with two monomoraic syllables.

(247) Libyan Tripoli Short Vowel and Schwa Weight Asymmetry (Yoda 2005, p. 111-112)

a. ˈtˤəbːəħ
call.IMP.M
'call (m.)!'

b. sakˈn -in
living -M.PL
'living (m.pl.)

c. ˈħuʃa

'house'

(248) Libyan Tripoli Syllable Restrictions (Yoda 2005, p. 57, 70, 104)

a. sənsla

'necklace'

b. qal -(ə̱)t -l -u
say.PFV -SBJ.3SG.F. -DAT -OBJ.3SG.M
'she said to him'

c. msəkː.r -in
close.PTCP -PL
'closed'

Like Moroccan Casablanca, Libyan Tripoli displays an asymmetry between length and weight. Therefore, its syllable restrictions are best expressed in terms of *X-slots*. Thus, Libyan Tripoli permits three X-slots in a syllable rime, with an optional consonantal fourth. This accounts for the optional epenthesis in (248-b), which reflects the optional incorporation of the fourth X-slot [t] into the syllable.

2.6.3 Evaluation

The principle behind assigning syllable tolerance restrictions is that moraic restrictions are the default option, and chosen where there is no asymmetry in the tolerance of CVXC syllables. Segmental level restrictions are required where there is asymmetry in CVXC tolerance, or where required to rule out unattested medial CVCCC syllables, as in many of the Levantine varieties. X-slot restrictions are needed where there is an asymmetry between weight and length, as in Moroccan Casablanca and Libyan Tripoli.

The syllable tolerance restrictions motivated above are summarised in Table 2.13 below.

Table 2.13: Summary of Syllable Tolerance Restrictions

	Tolerance Type	No. Elements	Extra Element Requirement
Algerian	Segmental	2	Sonority
Palestinian	Segmental	2	Coronal
Lebanese	Segmental	2	Coronal
Iraqi	Segmental	2	-
Wadi Ramː	Segmental	3	-
Muscat	Segmental	2	Sonority, Lexical, Fast Speech
Sˤanʕaːniː	Segmental	2	Fast Speech
Makːan	Moraic	2	Branching Nucleus
Qatari	Moraic	2	Branching Nucleus
Cairene	Moraic	2	-
Tunisian	Moraic	3	-
Rufaidah	Moraic	3	-
Rwaili	Moraic	3	-
Haːyil	Moraic	3	-
Moroccan Casablanca	X-slot	3	-
Libyan Tripoli	X-slot	3	Consonantal

This analysis accounts for the closed syllable shortening, postgeminate deletion, metathesis and some of the epenthesis behaviours discussed in Kiparsky (2003) where these varieties displayed greater variation than previously predicted. However, it doesn't account for the domain edge cluster behaviour.

To account for the full range of phonological behaviours identified in Kiparsky (2003), we need both syllable restrictions at different levels of prosodic structure as motivated above and a way to repair violations to syllable structure that accounts for medial epenthesis as well as domain edge clusters. This is because the distribution of varieties across the groups with X-slot, segmental and moraic restrictions does not covary with the epenthesis position and domain edge behaviour discussed in the literature.

In the following Table 2.14, varieties are grouped by their Tolerance Type. The Epenthesis Type of Kiparsky is distributed across Tolerance types rather than covarying with it.

Table 2.14: Tolerance of medial CVXC Syllables, Type of Tolerance Restriction and Epenthesis Type

Key: Y = Yes, N = No, N_E = No with Exceptions, Y_E = Yes with Exceptions, Segmental = Segmental Restrictions, Mora = Moraic Restrictions, X-slot = X-slot Restrictions, - = Not Applicable

Variety	CVCC	CVːC	CVCː	Tolerance Type	Epenthesis Type
Wadi Ramː	Y	Y	Y	Segmental	VC
Tunisian	Y	Y	Y	Mora	C
Rufaidah	Y	Y	Y	Segmental	C
Rwaili	Y	Y	Y	Segmental	C
Haːyil	Y	Y	Y	Segmental	C
Palestinian	N_E	Y	Y	Segmental	VC
Sˤanʕaːniː	N_E	Y	Y	Segmental	CV
Lebanese	N_E	Y	Y	Segmental	VC
Muscat	N_E	Y	Y	Segmental	VC
Algerian	N_E	Y	Y	Segmental	C
Iraqi	N	Y	Y	Segmental	VC
Qatari	N	Y	N	Mora	VC
Makːan	N	Y_E	N	Mora	CV
Cairene	N	N	N	Mora	CV
Moroccan Casablanca	R	-	-	X-slot	C
Libyan Tripoli	Y	-	Y	X-slot	C

Therefore, rather than adding extra dialect groups, we need two axes of variation: Tolerance of medial CVXC syllables, and Repair of illicit syllables. As such, this splits the criteria that Kiparsky identified in half; metathesis, closed syllable shortening and postgeminate high vowel deletion reflect the TOLERANCE of medial CVXG syllables, whereas initial and final clusters along with epenthesis position reflect REPAIR. Next I will sketch out the initial REPAIR analysis, though this will be further clarified in the next chapter.

2.6.4 Repair

Why should domain edges behave differently to word-medial syllables? This phenomenon is widespread cross-linguistically. While there are languages such as Manam which have symmetrical word internal and final codas[7] (Buckley 1998; Piggott 1999), many languages have more consonants at the right edge than word-medially, and these final consonants can be extrametrical, therefore invisible, for stress assignment. For example, English permits one (and occasionally two) consonants in word medial codas as seen in (249-a) and (249-b); whereas it permits up to four word finally if there is a word level suffix as in (249-c).

(249) English Codas

a. kɒn.sə.næn.tl̩
   consonantal

b. ɛmp.ti
   empty

c. θaʊ.zəntθs
   thousandths

Different approaches have been taken to account for word-final coda behaviour, including limiting sequences to only one place of articulation, and different syllable representations including appendices, defective syllables, attachment to higher prosodic constituents, extraprosodicity and non-moraic consonants.

I propose that REPAIR divides varieties into ONSET and CODA varieties according to whether resyllabified illicit consonants form the onset or coda of a new syllable

[7] Manam codas only contain nasals and these contribute a mora for stress word medially and word finally.

following epenthesis. Just as under Broselow (1992), CODA varieties are associated with promiscuous syncope that can create initial clusters, whereas in ONSET varieties the syncope will not create initial clusters. As such, the distribution of initial and final consonantal clusters is associated with REPAIR as summarised in Table 2.15 below.

Table 2.15: Initial Summary of REPAIR

Key: Y = Yes, N = No

	ONSET	CODA
Epenthesis Position	C<u>V</u>	<u>V</u>C
Syncope Type	Non-Promiscuous	Promiscuous
Initial #CC- Permitted	N	Y
Final -CC# Permitted	Y	N

While this may be the canonical distribution of word-edge clusters, it is clear from the analysis of varieties in Section 2.4 that there is variation in this distribution, ranging from Iraqi where no word-edge clusters are permitted, through to Tunisian where clusters at either edge are permitted. In previous work, varieties like Tunisian have been considered 'C' varieties (Kiparsky 2003; Watson 2007) or 'Extreme Coda' type (Farwaneh 2009). However, I have demonstrated that Rufaidah, Rwaili and Haːyil are 'C'-esque varieties that still resyllabify extrasyllabic consonants as onsets post epenthesis. Therefore, I propose that there is a continuum between the canonical ONSET and CODA variety and the 'Extreme' ONSET and 'Extreme' CODA variety, as summarised in Table 2.16 below.

Table 2.16: Second Summary of REPAIR

Key: Y = Yes, N = No

	ONSET	Extreme ONSET	CODA	Extreme CODA
Epenthesis Position	C<u>V</u>		<u>V</u>C	
Syncope Type	Non-Promiscuous		Promiscuous	
Initial #CC-	N	Y	Y	Y
Final -CC#	Y	Y	N	Y

2.7 Conclusion and Next Steps

In this chapter I have argued that previous work on Arabic syllable typology does not adequately account for the range of variation observed, in particular with regards to the tolerance of medial CVXC syllables. That CV:C and CVC: syllables, both with long segments, tend to pattern together suggests that a representation that takes quantity into account is needed.

I have proposed that a bifurcated typology is required, consisting of Tolerance of medial CVXC syllables at the segmental, moraic, or X-slot tiers, as well as Repair of illicit syllables in terms of epenthesis direction (Coda or Onset) and the licensing of word edge extrasyllabic segments. This typology is summarised in the following tables, Table 2.17 for Tolerance and Table 3.4 for Repair.

Table 2.17: Summary of Tolerance

Tier of Syllable Restriction	Moraic	Segmental	X-slot
No. of Elements	2–3		
Extra Elements	variety specific including coronality, sonority and speech tempo		

Table 2.18: Summary of Repair

Key: Y = Yes, N = No

	Onset	Extreme Onset	Coda	Extreme Coda
Epenthesis Position	CV		VC	
Syncope Type	Non-Promiscuous		Promiscuous	
Initial #CC-	N	Y	Y	Y
Final -CC#	Y	Y	N	Y

The advantage of this approach is that is does not require any extra representational machinery — no adjunction to mora, no semisyllables. Rather, it assumes that the same syllabification process for CVC syllables occurs in larger medial syllables according to the tier at which syllabification occurs. It also allows for variation and restrictions on CVCC syllables in a flexible way to account for the behaviour of each individual variety, rather than solely demarcating the boundaries of the canonical space.

Furthermore, rather than needing to add increasing numbers of dialect groups, such as the Cv group, we can analyse these as ONSET varieties with segmental restrictions.

However, exactly how to represent these word edge clusters has not been discussed here. Furthermore, Kiparsky (2003) demonstrated that stress did not covary with his dialect types. Given the extension of the typology I have motivated here, it is appropriate to revisit the issue of stress to explore what evidence it provides for the structure of final CVXC syllables and how it fits into the phonological typology of Arabic.

3
The Relationship between Stress and Syllable Structure

Contents

3.1 Introduction

In the previous chapter, I proposed a new typology of Arabic syllable structure. This requires two parts, TOLERANCE and REPAIR. Under REPAIR, ONSET varieties and EXTREME CODA varieties permit word final CC clusters.

However, the structure of these final CC clusters is unclear, and has been the subject of extensive debate in the literature. In this chapter, I explore the metrical stress analysis of these sixteen varieties to see what evidence stress can provide for the structure of these final CVXC syllables, and to what extent any of these parameters are correlated with the syllable typology motivated here.

3.1.1 Previous Approaches to Superheavy Syllables

The syllable inventory in Figure 3.1 below suggests a three–way weight distinction, assuming each element in the coda carries a mora. Therefore, CVːC and CVCC would be trimoraic, as argued by Broselow (1976, p. 34). This is typologically unusual, as not only are three-way oppositions relatively rare (Trubetzkoy 1939; Broselow, Chen, and Huffman 1997; Hayes 1995), languages often avoid trimoraic syllables.

Figure 3.1: Arabic Syllable Types

Key: C = Consonant, V = Vowel

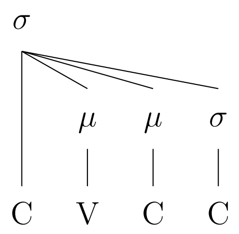
Light	CV	kataba	‘he wrote’
Heavy	CVː	saːlim	‘intact’
	CVC	man	‘who’
Superheavy	CVːC	qaːl	‘he said’
	CVCC	nasr	‘prose’
	CVCː	madd	‘he extended’

From Azzabi (2001, pp. 112–113)

Broselow (1992, p. 10) proposed the Bimoraicity Constraint, arguing that the optimal syllable in Arabic is bimoraic and that phonological processes conform to this constraint. This follows arguments made by McCarthy and Prince (1990a) for Classical Arabic, and is further supported by Golston and Riad (1997) who demonstrated that Classical Arabic metrical positions in poetry were maximally bimoraic.

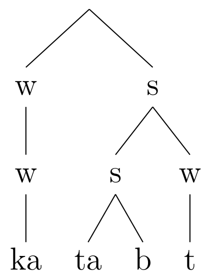
Accordingly there have been a range of proposals to avoid analysing CVXC syllables as trimoraic. Hayes (1995) proposes word final consonant extrametricality, incorporating it into the syllable after stress assignment. McCarthy (1979, p. 453) and McCarthy and Prince (1990a) propose that a syllable node could dominate another syllable node associated with the final C, as shown in Figure 3.2 below.

Figure 3.2: McCarthy (1979) Approach to Superheavy Syllables



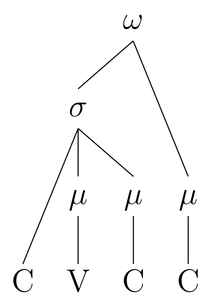
Halle and Vergnaud (1979) propose a final branching foot with the final C as the weak element, as shown in Figure 3.3 below.

Figure 3.3: Halle and Vergnaud (1979) Approach to Superheavy Syllables



Kenstowicz (1986) and Kiparsky (2003) suggest that the word final C and its mora could attach directly to the word node, as in Figure 3.4 below.

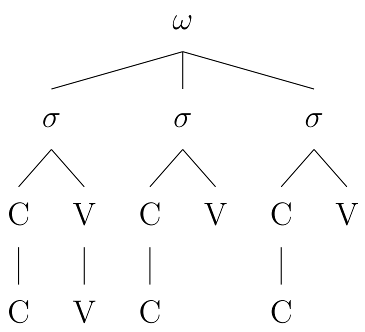
Figure 3.4: Kiparsky (2003) Approach to Superheavy Syllables



Aoun (1979), Angoujard (1981; 1990) and Selkirk (1981) view the final C as the onset to a syllable with an empty nucleus, as shown in Figure 3.5 below. These approaches make reference to the theory of CV Phonology, in which CV is the only possible syllable structure. Similarly, Broselow (1992) argues that the final

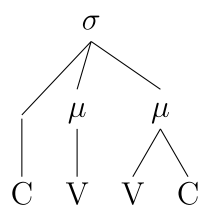
C can form the onset or coda of a degenerate syllable without a nucleus, but this is not explicitly within the CV Phonology framework.

Figure 3.5: Selkirk (1981) Approach to Superheavy Syllables



Finally, Broselow et al. (1992; 1995; 1997) and Watson (2002; 2007) propose that the final C shares a mora with a preceding segment rather than projecting its own mora, as shown in Figure 3.6 below.

Figure 3.6: Broselow Approach to Superheavy Syllables



In this chapter I demonstrate that stress, and in particular extrametricality, can shed light on what type of syllable representation best accounts for the data. First in Section 3.2, I will lay out the development of approaches to stress, culminating in the Metrical Stress Theory of Hayes (1995) which forms the basis of my analysis. Then, I explore my dataset of sixteen varieties in Section 3.3. I find that syllable extrametricality cannot make a final CVXC syllable extrametrical, nor can consonant extrametricality make a final Cː extrametrical. Therefore, in Section 3.4 I argue that CVXC constitutes two syllables where the final C is part of a second syllable with a catalectic mora as the nucleus. In Section 3.5 I show that extrametricality in general correlates with Tolerance, giving further support to this manner of grouping varieties together.

3.2 Approaches to Stress and Extrametricality

3.2.1 What is Stress?

We can define stress as ‘the means of marking relative prominence within various organisational groupings of metrical units’ (Lahiri 2001, p. 1347). It is a perceptual matter that allows us to distinguish between the noun ‘*re*cord’ and the verb ‘re*cord*’ in English. Although there are phonetic correlates cross-linguistically, such as pitch, amplitude, vowel quality and duration, there is no isometric mapping, thus it appears to be a phonological construct. Stress can be predictable based on the underlying phonological structure of the word[1], and it displays the following phonological characteristics: one single strongest stress for a word or phrase (culminativity); hierarchical structure with multiple degrees of stress (Liberman and Prince 1977, p. 262); and no assimilation of stress to adjacent syllables (McCarthy 2003; Hayes 1995).

3.2.2 Stress Analyses

This section summarises the following approaches to phonological stress analysis: Chomsky and Halle (1968), Liberman and Prince (1977), Prince (1983), Halle and Vergnaud (1987), Hayes (1995), and Optimality Theory (OT) approaches.

3.2.2.1 Sound Pattern of English

Chomsky, Halle, and Lukoff (1956) found that some aspects of stress was predictable by regular phonological rules. Chomsky and Halle (1968) expanded this further to showcase rules predicting English stress including phrasal stress contours.

Unlike other features discussed in Chomsky and Halle (1968), stress is not binary, with native speakers able to distinguish between multiple degrees of stress.

[1]However, this is not globally true. In some languages, stress falls on a fixed syllable, as in Polish where it generally falls on the penultimate syllable (cf. [ˈkɒmin], ‘chimney’, [kɒˈminɛk], ‘fireplace’); morphology can play a part, as in the English pair of the noun ‘*re*cord’ and the verb ‘re*cord*’, even if a phonological analysis that makes reference to lexical categories can be given; in others, it is lexically fixed. However, for the Arabic varieties discussed here, stress is predominately predictable based on phonological structure, though Chapter 4 will discuss cases of apparent exceptions to this phonologically predictable stress.

To account for this, stress was not marked as [± stress], but with numerical degrees, as in Figure 3.7 below.

Figure 3.7: SPE Stress Values

Value	Interpretation
[0stress]	stressless
[1stress]	primary stress
[2stress]	secondary stress
[3stress]	tertiary stress
[*n*stress]	etc.

Chomsky and Halle (1968, pp.16-17) proposed the Stress Subordination Convention (SSC), recognising that stress shifted as morphological and syntactic operations occurred: 'when primary stress is placed in a certain position, then all other stresses in the string under consideration at that point are automatically weakened by one'. However, this theory had limitations. Syllables appear to be stressed not because of their own composition, but rather their position in the word or phrase. These special properties of stress led to further developments in the field.

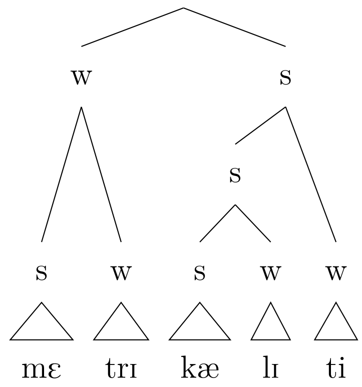
3.2.2.2 Liberman and Prince (1977)

Liberman (1975) and Liberman and Prince (1977) developed Metrical Phonology to analyse prosodic phenomena. They proposed two hierarchical structures: metrical trees[2] and metrical grids. Metrical trees illustrate relative prominence based on constituent structure; metrical grids depict rhythm (Liberman and Prince 1977, p. 392).

In the metrical tree, stress is portrayed as a binary relationship between two constituents, one weak (W) and one strong (S). The strongest unit is dominated by S nodes at every level, and this unit receives primary stress. In the tree in Figure 3.8 below, this is the syllable /kæ/.

[2] Note Rischel (1972) originally proposed stress could be depicted through a morphosyntactic tree.

Figure 3.8: Metrical Tree: ‘metricality’



Metrical grids depict hierarchies of syllable prominence. In Figure 3.9, the syllable /kæ/ has the most asterisks, so receives primary stress. Note that the number of asterisks corresponds to the number of [s] nodes in the tree in Figure 3.8 above.

Figure 3.9: Metrical Grid: ‘metricality’

		*		
*		*		
*	*	*	*	*
mɛ	trɪ	kæ	lɪ	ti

The association of asterisks to syllables should not be taken to mean Liberman and Prince (1977) viewed stress as a feature or property of syllables. Rather, nuclei bear stress by being associated to an autosegmental tier. Since they portrayed degrees of stress through the tree and the grid, in phonological rules they used the notation [±stress].

This iteration of metrical phonology and its notation were taken up widely in the literature, with extensive debate over whether to use trees and grids (Liberman and Prince 1977; Hayes 1981), or merely a reduced version of grids (Prince 1983; Selkirk 1984).

3.2.2.3 Prince (1983)

Prince (1983) advocated only using the grid, and accounting for binary patterns of stress with a perfect grid:

Figure 3.10: Prince (1983) Perfect Grid

Feet	*Left-to-Right(LR)*					*Right-to-Left(LR)*				
peak first	x		x		x	x		x		x
	x	x	x	x	x	x	x	x	x	x
	σ	σ	σ	σ	σ	σ	σ	σ	σ	σ
trough first		x		x			x		x	
	x	x	x	x	x	x	x	x	x	x
	σ	σ	σ	σ	σ	σ	σ	σ	σ	σ

Prince (1983) also proposed the *End Rule*, whereby the leftmost or rightmost element with the most grid marks gains a higher grid mark, and therefore primary stress. The effect of heavy syllables on stress is attained by projecting grid marks from them before applying a perfect grid.

3.2.2.4 Halle and Vergnaud (1987)

Halle and Vergnaud (1987) proposed a parametrised framework for lexical stress based on the left- or right-headedness of constituents; boundedness of the domain; direction of constituent construction; number of levels; and the existence of any extrametrical units.

They impose the following conditions (Halle and Vergnaud 1987, p. 15):

Exhaustivity Condition: Every element in a given line must be included within a constituent.

Maximal Constituency: Each constituent must include as many elements as possible within bounds of metrical theory.

Faithfulness Condition: 'Each constituent has an obligatory head, represented by a grid element at the next-higher level' (Kager 2007, p. 199).

Recoverability Condition: The location of the head and of constituent boundaries must be recoverable from each other.

Halle and Vergnaud (1987) developed bracketed grid notation. They represent prominence and constituency within one structure, as in Figure 3.11 which shows their analysis of 'metricality'.

Figure 3.11: Halle and Vergnaud (1987) Metrical Grid: 'metricality'

		x		
x		x		
(x	x)	(x	x)	x
mɛ	trɪ	kæ	lɪ	ti

Note Halle and Vergnaud, unlike Liberman and Prince (1977) and Prince (1983), do make use of the foot, as demonstrated by brackets showing foot constituency.

3.2.2.5 Hayes (1995)

Hayes (1995) built on Halle and Vergnaud's (1987) parametrised approach to include feet from a pre-set inventory. He assumes the following small universal binary foot inventory, following from McCarthy's (1986) work on templatic morphology and Hayes' (1987) typology of stress systems.

Syllabic Trochee: Strong-Weak syllable pairs regardless of syllable structure

Moraic Trochee: Strong-Weak syllable pairs accounting for syllable structure

Iamb: Weak-Strong syllable pairs accounting for syllable structure

He proposes that stress can be predicted with the parameters in Figure 3.12.

Figure 3.12: Hayes (1995) Metrical Stress Theory Parameters

Quantity Sensitive	Yes		No	
Foot Type	Trochee		Iamb	
Direction of Parsing	Left		Right	
End Rule	Left		Right	
Extrametricality	Yes		No	
Extrametrical Unit	Consonant	Syllable	Mora	Foot

That is, dialects vary in terms of whether they are sensitive to syllable weight in building feet (quantity sensitivity); what types of feet are built (foot type); whether feet are built starting at the left or right edge of the word (parsing directionality); whether main stress occurs towards the left or right edge (end rule); whether a unit is consistently ignored for stress assignment (extrametricality); and if so, what that unit is. Hayes (1995) uses the following notation: the stressed component of a foot is marked by 'x', the non-stressed component is marked by a dot '·', and the foot is marked by brackets '()'. This notation shall be adopted throughout this chapter.

3.2.2.6 OT Approaches

With the advent of Optimality Theory (OT) (Smolensky 1993; Kager 1999), these approaches have been translated into constraint-based models (Crowhurst and Hewitt 1994; Eisner 1997; Kenstowicz 1996; Elenbaas 1999; Kager 1999; Kager 2007; McCarthy 2003).

Whereas generative phonology derives surface forms through the application of linearly ordered rules that are language specific but may have parallels in other languages, OT and its successors use universal violable constraints on surface forms, with language specific rankings that reject candidates to select the optimal surface form with fewest violations.

OT approaches can model both foot-based approaches (Kager 2007) and grid-based theories with syllabic constituents (Gordon 2002). Both subtypes attempt to account for Boundedness, Headedness and Weight sensitivity (Al-Mohanna 2004). Figure 3.13 summarises the constraints used by Kager (2007).

Figure 3.13: Kager 2007 OT Stress Constraints

Constraint		Meaning
Ft-Type=Trochee	Ft-Type=Iamb	Foot type
Parse-Syl		Syllables must be parsed into feet
All-Ft-R	All-Ft-L	Feet must align to right or left edge of prosodic word
Non-Finality		Last syllable is extrametrical
Align-Head-L	Align-Head-R	Stressed syllable is towards left or right end of the word
Align-PrWd-Left	Align-PrWd-Right	Word must begin or end with a foot

3.2.2.7 Summary

Most of the approaches above make recourse either to the foot or the grid. Grid based approaches are better at analysing systems that have initial and final stress, such as Tauya[3] (MacDonald 1990) where the initial syllable and odd-numbered syllables from the right are stressed. However, various segmental processes are sensitive to the foot, so a foot-based analysis provides a more coherent system (Gordon 2014, p. 159).

The analyses motivated here will use Hayes (1995)'s Metrical Stress Theory. This is preferred to OT-based accounts as other Arabic phenomena, such as the homorganicity constraint on root consonants,[4] contradict the Richness of the Base principle[5] of OT (Prince and Smolensky 1993). For further discussion of the challenges of applying constraint based models to this research, see Section 5.3.1.

[3]A Trans-New Guinea Rai Coast language spoken in Madang Province of Papua New Guinea

[4]It has been noted by the lexicographer al-Khaliːl (d.791) and scholars including Cantineau (1946), Greenberg (1950), Bachra (2001) and Frisch, Pierrehumbert, et al. (2004), that Arabic avoids roots with homorganic consonants; however this restriction does not apply to all consonants in the word. Thus *T-T-K is not a viable or attested root, but the combination surfaces in the word [tatakalːam], 'you speak'. Therefore, this seems to be a restriction on the base.

[5]The idea that any restrictions should come from the constraints used to identify the optimal candidate, rather than having restrictions on possible inputs

3.2.3 Degenerate Feet and Extrametricality

The above theories find that some syllables are systematically ignored for stress assignment (extrametricality) while others can create monosyllabic light feet (degenerate feet).

3.2.3.1 Extrametricality

Introduced by Liberman and Prince (1977) as a type of extraprosodicity[6], Hayes (1982) argued it permitted a unified approach to the parsing of medial and final syllables across different lexical categories in English, where the nouns and adjectives have syllable extrametricality (see (250)) whereas verbs have consonant extrametricality (see (251)). Note the use of <> to denote the extrametrical element.

(250) Syllable Extrametricality in English Nouns and Adjectives

a. ˈre<cord>

b. eˈter <n+al>

(251) Consonant Extrametricality in English Verbs

a. reˈcor<d>

b. deˈvelo<p>

Subsequently, it was formalised as an invisibility to all phonological rules at a given level, represented by Inkelas (1989) as a mismatch between segmental and prosodic constituency. It only occurs at domain edges, and this peripherality is enshrined as a key element by both Hayes (1995, p. 57) and Kager (1995b). Hayes places the following restrictions on extrametricality: only phonological constituents can be extrametrical; only constituents on a domain edge can be extrametrical; the unmarked edge is the right edge; and it is blocked if the entire domain would become extrametrical (otherwise referred to as the Principle of Nonexhaustivity).

[6]This covers all phenomena that violate prosodic rules, whereas extrametricality only refers to violations of stress assignment

The corresponding Optimality Theory approach is NON-FINALITY (Prince and Smolensky 1993), which was introduced on three levels, depending on whether the focus is syllables, feet, or both.

Figure 3.14: OT: Non-Finality

NON-FIN 1	The head of the PrWd does not fall on the word-final syllable *From Prince and Smolensky (1993, p. 40)*
NON-FIN 2	The head foot of the PrWd must not be final *From Prince and Smolensky (1993, p. 43)*
NON-FIN 3	No head of a PrWd is final in the PrWd *From Prince and Smolensky (1993, p. 52)*

3.2.3.2 Degenerate Feet

The degenerate foot is a foot that does not contain the maximum number of syllables, that is, in a quantity sensitive language, a monomoraic foot, and in a quantity insensitive language, a monosyllabic foot. Some languages avoid degenerate feet; some allow them in peripheral positions or under main stress (Kiparsky 1991; Hayes 1995; Kager 2007).

Hayes (1995) accepts degenerate feet, suggesting a parameter for strong prohibition of degenerate feet, weak prohibition of degenerate feet, and not permitted. To remove degenerate feet from metrical theory altogether, Kiparsky (1991) proposed catalexis.

3.2.4 Catalexis

Metrical verse can contain both hypermetrical syllables at line ends (see (252-a)) and missing weak syllables (see (252-b)). Kiparsky (1977) notes both are permitted throughout time in English verse, as depicted in the examples below where (252-a) demonstrates hypermetricality and (252-b) catalexis.

(252) English Hypermetricality and Catalexis Kiparsky (1991, p. 3)

a. There are | more things | in heaven | and earth | Hora<tio>

b. Hail to | thee, blithe | spirit //
bird thou | never | wert ∅

Catalexis is also found across Classical Arabic Poetic metres, including metron internally in the *basi:t* meter, and in every metron in the *mutaqa:rib* meter (Golston and Riad 1997).

That domain edges can be longer or shorter by one metrical constituent is not limited to poetry, but also occurs in phonology more generally (Giegerich 1985; Burzio 1987). Inspired by poetry, Kiparsky (1991) proposed catalexis as the logical counterpart to extrametricality. Just as extrametricality ignores segmentally overt material for prosodic operations, catalexis includes segmentally empty material for prosodic operations. This segmentally empty material can be a catalectic mora or a catalectic syllable, and the presence of a catalectic syllable does not entail the presence of a catalectic mora. Both extrametricality and catalexis are subject to peripherality constraints, in that they can only act at the domain edge.

This proposal accounts for the generalisation that trochaic systems with subminimal words can have final degenerate stressed feet, whereas languages without subminimal words do not, a generalisation confirmed by Kager (1995a) in an extensive typological survey. The benefit of catalexis is that it eliminates degenerate feet and gives a universal minimal word — bimoraic in quantity sensitive systems, and bisyllabic in quantity insensitive systems.

Green (1995) demonstrates this solves the initial and final stress in Maranungku[7] as shown in (253) below, where the catalectic syllable is marked with square brackets. Catalexis permits the final stress to be part of a bisyllabic trochee.

(253) Maranungku Initial and Final Stress (Green 1995, p. 21)

(σ	σ)	(σ	σ)	(σ	[σ])
ˈlang	ka	ˌra	te	ˌti	

The key diagnostic for the presence of catalexis is the presence of subminimal words, as is found in Ono[8]. The secondary diagnostic is the presence of degenerate feet, such

[7] A Western Daly language spoken in Northern Australia.
[8] A Trans-New Guinea language spoken in the Morobe Province, Papua New Guinea.

as the final secondary stress in odd numbered polysyllables in Diyari[9] (Kager 1995b).

When accounting for these stress patterns, catalexis is presented as either always ‘on’ or always ‘off’. In Tübatulabal[10], it is always ‘on’ regardless of whether the final syllable is V, CV, or CVC. Stress falls on long vowels and every other light syllable, which appears to be a trochaic pattern, except that the final syllable is also stressed regardless of quantity or that of the preceding syllable, as shown in (254) (Kiparsky 1991, p. 12–13).

(254) Tübatulabal Trochees but with Final Stress (Kiparsky 1991, pp. 12-13)

a. tˈaahˈawilˈa
   ‘the summer (obj.)’
b. haˈniilˈa
   ‘the house (obj.)’
c. ˈeeyˈeeˈu
   ‘he became ashamed’
d. wɪtˈayhatˈal
   ‘the Tejon Indians’
e. ˈwɪtayhˈatalˈaabacˈu
   ‘away from the Tejon indians’
f. ˈɨmbˈɨywˈɨbaʔˈat
   ‘he is wanting to roll string on his thigh ’

As argued by Kager (1989) and Hayes (1995), if degenerate feet are prohibited Tübatulabal can be accounted for as right-to-left moraic trochees, with the exception that the final syllable is stressed regardless of its quantity or that of the preceding syllable. Kiparsky (1991) argues that the solution cannot be an iamb, as this could not stress the initial syllable in (254-f) nor the antepenultimate in (254-a). However, if there is a final catalectic mora, the final stress is a consequence of the trochaic stress. Word final surface short open syllables are then underlyingly long, and

[9] A Pama-Nyungan language spoken in South Australia
[10] A Uto-Aztecan language spoken in Sierra Nevada, California, US.

word-final surface closed syllables are then underlyingly bisyllabic. Therefore, the metrical structure of Tübatulabal is as shown in (255).

(255) Tübatulabal Right-to-Left Moraic Trochees with Mora Catalexis

a. (tˈaa)(hˈawi)(lˈa ·)
   ‘the summer (obj.)’
b. ha(ˈnii)(lˈa ·)
   ‘the house (obj.)’
c. (ˈee)(yˈee)(ˈu ·)
   ‘he became ashamed’
d. wɪ(tˈay)ha(tˈal ·)
   ‘the Tejon Indians’
e. (ˈwɪtay)(hˈata)(lˈaa)ba(cˈu ·)
   ‘away from the Tejon indians’
f. (ˈɨm)(bˈɨy)(wˈɨba)(ʔˈat ·)
   ‘he is wanting to roll string on his thigh ’

However, catalexis does not have to be always on. Kiparsky (1991, p. 17) notes it can be a ‘last resort’ strategy to avoid subminimal words only. Likewise, Kager (1995a) recognises that it can behave as a morpheme in Toba Batak[11]. As such it can account for the stress shift in the passive (ˈlapu ‘to smear’ vs laˈpu ‘be smeared’). Similarly for Tongan[12], Kager (1995a) views the bare definite to be constituted by a catalexic mora to account for the stress shift that is also seen with overt enclitics. Further restricted uses for catalexis include lexical specification for catalexis or extrametricality to account for marked stress patterns in Dutch (Nouveau 1994), in principle cross-linguistically (Peperkamp 2004), and its presence on subminimal clitics in Catalan (Torres-Tamarit and Pons-Moll 2019). Much of this research where catalexis is not ‘always on’ argues for catalexis functioning for morphological purposes or for solving exceptional marked stress cases. However

[11] A Malayo-Polynesia language spoken in North Sumatra, Indonesia.
[12] A Malayo-Polynesian language spoken in Tonga.

there has been limited research considering catalexis as a consistent strategy for phonological repair below the level of repairing degenerate feet.

3.3 Analysis of Stress in Arabic Varieties

Watson (2011a, pp. 2–4) makes the following generalisations about Arabic stress, noting that all varieties: are quantity sensitive; assign stress to a final CVːC or CVCC; do not assign stress to a final CVC; and are sensitive to morphological structure in stress assignment. Most are trochaic, though some Western and Bedouin varieties including Cyrenaican (Mitchell 1960) and Wadi Ramː (Al Mashaqba 2015) are iambic, as shown in (256).

(256) Cyrenaican Bedouin Iamb (Watson 2011a, p. 3)

a. nuˈxal
   ‘palm-trees’

b. biˈna
   ‘he built’

There are broadly four patterns of stress assignment in this dataset. For each, I will give a description of the pattern with representative data from one variety, before exploring the metrical stress analysis of the varieties with that pattern.

3.3.1 Pattern 1

Palestinian, Algerian, Rufaidah, Lebanese, Iraqi, Makːan, Muscat, Qatari, and Sˤanʕaːniː share the following broad pattern of stress assignment, although there are Sˤanʕaːniː is slightly different as will be discussed below. Using data from Palestinian as a representative example, we can describe the pattern as follows:

1. Stress final superheavy syllable,

2. Else stress penultimate heavy syllable,

3. Else stress the antepenultimate

(257) Final Superheavy Stress (Abu-Salim 1982a, p. 59)

a. baˈnaːt
   ‘girls’

b. dukːˈaːn
   ‘store’

c. kaˈtabt
   ‘I wrote’

d. maˈħalː
   ‘place’

(258) Heavy Penultimate Stress (Abu-Salim 1982a, p. 60)

a. ˈtaːjir
   ‘merchant’

b. ˈmaktab
   ‘office’

c. maˈkaːtib
   ‘offices’

(259) Antepenultimate Stress (Abu-Salim 1982a, p. 60)

a. ˈkatabu
   ‘they wrote’

b. ˈmasˤtˤasa
   ‘ruler’

c. ˈʃaːfato
   ‘she saw him’

I will now explore the metrical stress analyses of the varieties that share this pattern, demonstrating that these varieties share moraic trochees, End Rule Right and syllable extrametricality. There is some variation in parsing directionality, but data is not available to discern this parameter for all varieties, and the items where a difference is found constitutes a very small set. As such, both parsing directionalities have been included in this same general pattern.

3.3.1.1 Palestinian

According to Hayes (1995), Palestinian displays moraic trochees built left-to-right with End Rule Right and both foot and consonant extrametricality.

Indeed, it must be moraic trochees to account for the stress pattern in the following two words:

(260) Moraic Trochees (Abu-Salim 1982a, p. 88, 60)

a. daˈras-na
   ‘we studied’

b. ˈkatabu
   ‘they wrote’

Since both words contain three syllables and stress falls on different syllables, Palestinian must be quantity sensitive. Furthermore, while (260-a) is ambiguous as to whether it is an iamb or a trochee as demonstrated in Figures 3.15–3.16, (260-b) must be a trochee as demonstrated in Figures 3.17–3.18.

Figure 3.15: Palestinian [darasna] with Moraic Trochee

	X	
	(x)	
da	ˈras	na

Figure 3.16: Palestinian [darasna] with Iamb

	X	
(.	x)	
da	ˈras	na

Figure 3.17: Palestinian [katabu] with Correct Moraic Trochee

X		
(x	.)	
ˈka	ta	bu

Figure 3.18: Palestinian [katabu] with Incorrect Iamb L−R

	X	
(.	x)	
*ka	ˈta	bu

It must be End Rule Right as in words with two feet such as [dukːˈaːn], ‘shop’, the stress is on the right hand edge (Abu-Salim 1982a, p. 60). In Figure 3.19

we see End Rule Right gives the attested final stress whereas End Rule Left in Figure 3.20 gives unattested initial stress.

Figure 3.19: Palestinian [dukːaːn] with Correct End Rule Right

	X
(x)	(x)
duk	kaːn

Figure 3.20: Palestinian [dukːaːn] with Incorrect End Rule Left

X	
(x)	(x)
*duk	kaːn

However, I dispute the extrametricality analysis Hayes proposes. Hayes argues that <u>two</u> units of extrametricality are necessary, as there will be times when one does not apply, but the other does. Foot extrametricality does not cover all final feet in Palestinian — rather, only the final foot when the word ends with LLLL or HLL, such as [ˈʃadʒaratu], 'his tree', and [ˈmadrasa], 'school' (Watson 2011a, p. 11). Consonant extrametricality is required in all other cases, such as [ˈkatab], 'he wrote'. In doing so, this accounts for why CVXC feet attract stress, unlike heavy feet. In Figure 3.21, foot extrametricality prevents stress from falling in the final foot of [ˈmadrasa], 'school', constructed of two light syllables, so it must fall on the initial foot instead: [*mad*].

Figure 3.21: Foot Extrametricality Analysis of Palestinian I

X		
(x)	<(x	.)>
mad	ra	sa

However, when the word ends with a consonant (so does not end in LLLL or HLL), consonant extrametricality applies instead. This makes the final syllable in [katab], 'he wrote', light rather than heavy (Abu-Salim 1982a, p. 135). As such, it can form a foot with the initial light syllable, which then gains initial stress, as we see Figure 3.22.

Figure 3.22: Foot Extrametricality Analysis of Palestinian II

X	
(x	.)
ka	ta<b>

For final CVXC syllables such as [katabt], 'I wrote' (Abu-Salim 1982a, p. 49), when consonant extrametricality applies it makes the final consonant extrametrical, but here still leaves a heavy syllable, which can form a foot by itself and receive stress, as we see in Figure 3.23.

Figure 3.23: Foot Extrametricality Analysis of Palestinian III

	X	
	(x)	
ka	tab	<t>

This is typologically unusual in terms of having two units of extrametricality and in requiring consideration of the broader environment. However, the data can be accounted for by a single unit of extrametricality: the syllable. This prevents a final LL foot from being footed, as we see in Figure 3.25, so the word [madrasa] (Watson 2011a, p. 11) receives initial stress.

Figure 3.24: Syllable Extrametricality Analysis of Palestinian I

X		
(x)		
ˈmad	ra	< sa >

In the case of a bisyllabic word such as [ˈkatab], 'he wrote' (Abu-Salim 1982a, p. 135), the residue is stressed as we see in Figure 3.25, so the word gains initial stress.

Figure 3.25: Syllable Extrametricality Analysis of Palestinian II

X	
(x)	
ˈka	< tab>

In order to accept syllable extrametricality, we need to account for CVXC syllables. These consistently receive stress in Palestinian. If no restrictions are placed on syllable extrametricality, and these are assumed to be trimoraic syllables,[13] then why are they not affected by extrametricality? We could place restrictions on syllable extrametricality, as Hayes does with foot extrametricality. However, trimoraic syllables are uncommon cross-linguistically, and assorted attempts have been made to reanalyse these syllables as underlyingly bimoraic instead (Broselow 1976; McCarthy and Prince 1990b). I will propose a representation below that will make the final element of these apparently CVXC syllables part of a separate syllable, thus preventing the syllable from being peripheral and thus subject to extrametricality. This novel analysis is preferable because it allows stress placement to occur due to the interaction of phonological structures and rules that exist across the dialect, rather than requiring exceptional architecture to solve analytic problems.

Syllable extrametricality is not the only alternative analysis that has been proposed for Palestinian. Steriade (1991) argues for mora extrametricality. Regardless of the theoretical quandaries that such a unit poses, mora extrametricality cannot account for data in Palestinian. Whilst it does correctly predict antepenultimate stress for [ʕalˈlamato], 'she taught him' (Abu-Salim 1982a, p.60), as we see in Figure 3.26, it incorrectly predicts penultimate stress rather than initial stress for [*ʕalˈlamat], 'she taught' as shown in Figure 3.27.

Figure 3.26: Mora Extrametricality Correct: [ʕalˈlamato]

	X		
(x)	(x	·)	
$\mu\mu$	μ	μ	$<\mu>$
ʕal	ˈla	ma	to

Figure 3.27: Mora Extrametricality Incorrect: [*ʕ]alˈlamat

	X	
(x)	(x	·)
$\mu\mu$	μ	$\mu <\mu>$
*ʕal	ˈla	mat

However, syllable extrametricality correctly predicts both cases, as shown in Figures 3.28 and 3.29 below:

[13] If we assume every coda element carries mora through weight by position.

Figure 3.28: Syllable Extrametricality
Correct: [ʕalˈlamato]

	X		
(x)	(x	·)	
$\mu\mu$	μ	μ	μ
ʕal	ˈla	ma	<to>

Figure 3.29: Syllable Extrametricality
Correct: [ˈʕ]allamat

X		
(x)		
$\mu\mu$	μ	$\mu\mu$
ʕal	ˈla	<mat>

Given syllable extrametricality, it is difficult to identify parsing directionality for Palestinian unless Classical Arabic words are used, as sequences of three light syllables plus a final syllable are not found in colloquial words. If, like Hayes (1995) we do accept using dialectal pronunciations of Classical vocabulary, we can ascertain parsing directionality as left-to-right with words such as [ˈʃaʤaratun], ‘a tree (acc.)’. As shown in Figures 3.30–3.31, Right-to-Left parsing incorrectly predicts antepenultimate stress rather than the attested initial stress.

Figure 3.30: Left-to-Right Parsing
Correct: [ˈʃaʤaratun]

X			
(X	.)		
ˈʃa	ʤa	ra	<tun>

Figure 3.31: Right-to-Left Parsing
Inorrect: [ˈʃaʤaratun]

	X		
	(x	.)	
*ʃa	ˈʤa	ra	<tun>

Therefore, Palestinian has moraic trochees, built left-to-right, with end rule right and syllable extrametricality. However, this syllable extrametricality does not seem to apply in case of CVXC syllables, such as those in (261).

(261) Final Superheavy Syllables Exempt from Syllable Extrametricality (Abu-Salim 1982a, p. 59)

a. dukːˈaːn
‘shops’

b. kaˈtabt
‘I wrote’

c. maˈħalː

‘place’

Superheavy syllables are often exempt from extrametricality cross-linguistically. However, in 3.4 I propose an explanation for why this happens.

3.3.1.2 Algerian

Algerian requires moraic trochees built left-to-right with End Rule Right and syllable extrametricality, similar to the other varieties discussed in Pattern 1.

It must be quantity sensitive to account for the different stress on the words in (262) that have the same number of syllables.

(262) Quantity Sensitivity (Bouhadiba 1988, p. 216, 225)

a. masaˈbuːga

‘overtaken (passive participle, f.s.)’

b. ˈtasabagu

‘you overtake him’

It must be moraic trochees to account for initial stress on [ˈsabagu], ‘they overtook’ (Bouhadiba 1988, p.221), as shown in Figures 3.32–3.33.

Figure 3.32: Algerian [sabagu] with Correct Moraic Trochee

X		
(x	.)	<>
ˈsa	ba	gu

Figure 3.33: Algerian [sabagu] with Incorrect Iamb

	X	
(.	x)	<>
*sa	ˈba	gu

It needs to be End Rule Right to account for cases with more than one foot, such as [masaˈbuːga], ‘overtaken’ (Bouhadiba 1988, p. 216). This is shown in Figures 3.34–3.35, where End Rule Right in Figure 3.34 predicts the attested penultimate stress, whereas End Rule Left in Figure 3.35 predicts unattested initial stress.

Figure 3.34: Algerian [masaˈbuːga] with Correct ERR

		X	
(x	.)	(x)	
ma	sa	buː	ga

Figure 3.35: Algerian [masabuːga] with Incorrect ERL

X			
(x	.)	(x)	
*ma	sa	buː	ga

There must be syllable extrametricality to account for the initial stress in [ˈtasabagu], 'you overtake him' (Bouhadiba 1988, p. 225). This is shown in Figures 3.36–3.37, where syllable extrametricality in Figure 3.36 predicts the attested initial stress, whereas no extrametricality in Figure 3.37 predicts unattested penultimate stress.

Figure 3.36: Algerian [ˈtasabagu] with Correct Syllable Extrametricality

X			
(x	.)		<>
ˈta	sa	ba	gu

Figure 3.37: Algerian [*tasaˈbagu] without Extrametricality

		X	
(x	.)	(x	.)
*ta	sa	ˈba	gu

Similarly, the initial stress in [ˈtasabagu], 'you overtake him', suggests that the parsing directionality is left-to-right, as right-to-left parsing with syllable extrametricality would predict antepenultimate stress as shown in figures 3.38–3.39.

Figure 3.38: Algerian [ˈtasabagu] with Syllable Extrametricality and Left-to-Right Parsing

X			
(x	.)		<>
ˈta	sa	ba	gu

Figure 3.39: Algerian [*tasaˈbagu] with Syllable Extrametricality and Right-to-Left Parsing

	X		
	(x	.)	<>
*ta	ˈ sa	ba	gu

Final CVXC syllables do receive stress despite final syllable extrametricality, as shown in (263). There is no data in Bouhadiba (1988) as to the stressing of final CVCː syllables, but there is no reason to believe they will behave differently to CVːC and CVCC syllables here.

(263) Algerian Final Superheavy Stress (Bouhadiba 1988, p. 215, 221)

a. saˈbag -t
overtake.PFV -SBJ.1SG
‘I overtook’

b. ʃaːˈf -uː -h
see.PFV -SBJ.3PL -OBJ.3SG.M
‘they saw him’

3.3.1.3 Rufaidah

Rufaidah requires moraic trochees built left-to-right with End Rule Right and syllable extrametricality.

It must be moraic trochees as this is the only way to get initial stress on a light syllable in words like [ˈkatabaw], ‘they wrote’ (Prochazka Jr. 1988, p. 21), in Figures 3.40–3.41. Moraic trochees predict the attested initial stress in Figure 3.40, whereas iambs predict unattested medial stress in Figure 3.41.

Figure 3.40: Rufaidah [ˈkatabaw] with Correct Moraic Trochee

X		
(x	.)	<>
ˈka	ta	baw

Figure 3.41: Rufaidah [ˈkatabaw] with Incorrect Iamb

	X	
(.	x)	<>
*ka	ˈta	baw

It must be left-to-right parsing to account for the behaviour in sequences of light syllables, such as in [ʔasˈtalam-at-ih], ‘she received it’ (Prochazka Jr. 1988, p. 20), in Figures 3.42–3.43 . Left-to-right parsing predicts the attested preantepenultimate stress in Figure 3.42, whereas right-to-left parsing predicts unattested antepenultimate stress in Figure 3.43.

Figure 3.42: Rufaidah [ʔasˈtalam-at-ih] with Correct L−R Parsing

	X			
(x)	(x	.)		<>
ʔas	ˈta	la	ma	tih

Figure 3.43: Rufaidah [ʔasˈtalam-at-ih] with Incorrect R−L Parsing

		X		
(x)		(x	.)	<>
*ʔas	ta	ˈla	ma	tih

End Rule Right is needed to account for words with more than one foot, such as [ʔasˈtalam-at], ‘she received’ (Prochazka Jr. 1988, p. 20), in Figures 3.44–3.45. End Rule Right (with syllable extrametricality, motivated below), predicts the attested antepenultimate stress in Figure 3.44, whereas End Rule Left predicts unattested initial stress in Figure 3.45.

Figure 3.44: Rufaidah [ʔasˈtalamat] with Correct ERR

	X		
(x)	(x	.)	<>
ʔas	ˈta	la	mat

Figure 3.45: Rufaidah [ʔasˈtalamat] with Incorrect ERL

X			
(x)	(x	.)	<>
*ˈʔas	ta	la	mat

Given End Rule Right, extrametricality is needed to account for not stressing final CVC syllables in words such as [ˈʔaħtaram], ‘he respected’ (Prochazka Jr. 1988, p. 21), in Figures 3.46–3.48. Syllable extrametricality predicts the attested initial stress in Figure 3.46, whereas consonant extrametricality predicts unattested medial stress in Figure 3.47, and no extrametricality predicts unattested final stress in Figure 3.48.

Figure 3.46: Rufaidah [ˈʔaħtaram] with Correct Syllable Extrametricality

X		
(x)		<>
ˈʔaħ	ta	ram

Figure 3.47: Rufaidah [ˈʔaħtaram] with Incorrect Consonant Extrametricality

	X	
(x)	(x	.)
*ʔaħ	ˈta	ra$< m >$

Figure 3.48: Rufaidah[ˈʔaħtaram] with Incorrect No Extrametricality

		X
(x)		(x)
*ʔaħ	ta	ˈram

Despite syllable extrametricality, final CVXC syllables still receive stress, as shown in (264).

(264) Rufaidah Final CVXC Stress (Prochazka Jr. 1988, p. 21)

a. kaˈtab -t
write.PFV -SBJ.1SG
'I wrote'

b. ji -ʃraˈb -uːn
SBJ.3 -drink.IMPF -SBJ.PL.M
'they drink'

c. ʔastaˈmarː -∅
continue.PFV -SBJ.3SG.M
'he continued'

Prochazka Jr. (1988, p. 20) suggests that when the CVXC syllable is created by concatenation of a monosegmental pronominal bound object the behaviour is different. He claims that when attaching to a heavy open syllable this final syllable does receive stress as in (265-a), but when attaching to a heavy closed syllable it does not as in (265-b). It is not that the object markers do not contribute to stress, as seen by the stress difference between (265-a) with an attached object and (265-c) without.

(265) Rufaidah Final Superheavy Stress with Object Markers (Prochazka Jr. 1988, p. 20)

a. ʃaːˈf -oː -h
see.PFV -SBJ.3PL -OBJ.3SG.M
'they saw him'

b. ˈʃaːf -at -k
see.PFV -SBJ.3SG.F -OBJ.2SG
'she saw you'

c. ˈʃaːf -aw
see.PFV -SBJ.3PL
'they saw'

Given the peculiarity of this pattern, it requires further investigation, but this is beyond the scope of this thesis. However, the remaining data suggests that here as well, CVXC syllables are behaving like two syllables or a syllable plus an extrasyllabic element.

3.3.1.4 Lebanese

Lebanese requires moraic trochees with End Rule Right and syllable extrametricality. Unlike the varieties discussed above, it requires right-to-left parsing directionality instead of left-to-right parsing directionality. The only data available to determine directionality are classical items.

It must be moraic trochees as this is the only way to get initial stress on [ˈdˤarabu], ‘they hit’ (Haddad 1984, p. 159), as shown in Figures 3.49–3.50, where only moraic trochees in Figure 3.49 predicts the attested initial stress.

Figure 3.49: Lebanese [dˤarabu] with Correct Moraic Trochee L−R

X		
(x	.)	
ˈdˤa	ra	bu

Figure 3.50: Lebanese [dˤarabu] with Incorrect Iamb L−R

	X	
(.	x)	
*dˤa	ˈra	bu

End Rule Right is needed to account for stress in words with more than one foot, such as [makˈtabti], ‘my library’ (Haddad 1984, p. 19), as shown in Figures 3.51–3.52. End Rule Right correctly predicts the attested medial stress in Figure 3.51, whereas End Rule Left incorrectly predicts initial stress in Figure 3.52.

Figure 3.51: Lebanese [maktabti] with Correct ERR

	X	
(x)	(x)	
mak	ˈtab	ti

Figure 3.52: Lebanese [maktabti] with Incorrect ERL

X		
(x)	(x)	
*ˈmak	tab	ti

While syllable extrametricality or consonant extrametricality could account for the final unstressed CVC in [ˈnazːal], ‘he brought down’ (Haddad 1984, p. 19), as shown in Figures 3.53–3.54, only syllable extrametricality can account for the final unstressed CVCV in [ˈmaktabe], ‘library’ (Haddad 1984, p. 73), as shown in Figures 3.55–3.56. As argued for Palestinian above, there is no need to posit foot extrametricality for this instead of syllable extrametricality.

Figure 3.53: Lebanese [ˈnazːal] with Syllable Extrametricality

X	
(x)	<>
ˈnaz	zal

Figure 3.54: Lebanese [ˈnazːal] with Consonant Extrametricality

X	
(x)	
ˈnaz	za<l>

Figure 3.55: Lebanese [maktabe] with Syllable Extrametricality

X		
(x)		<>
ˈmak	ta	be

Figure 3.56: Lebanese [maktabe] with Incorrect Consonant Extrametricality

	X	
(x)	(x	.)
*mak	ˈta	be

There is little data that can distinguish parsing directionality if we do not make recourse to Classical words pronounced locally. If we accept such data, [daˈrabana] 'he hit us' and [ʃacaratun] 'a tree' (Haddad 1984, p. 21) suggest right to left parsing, as shown in the Figures below.

Figure 3.57: Lebanese [ʃaˈcaratun] with Correct Right-to-Left Parsing

	X		
	(x	.)	<>
ʃa	ˈca	ra	tun

Figure 3.58: Lebanese [ʃaˈcaratun] with Incorrect Left-to-Right Parsing

X			
(x	.)		<>
*ˈʃa	ca	ra	tun

Furthermore, these items support syllable extrametricality rather than foot extrametricality, as foot extrametricality would predict initial stress on [darabana] rather than the attested antepenultimate stress as shown in the figures below.

Figure 3.59: Lebanese [daˈrabana] with Correct Right-to-Left Parsing and Syllable Extrametricality

	X		
	(x	.)	<>
da	ˈra	ba	na

Figure 3.60: Lebanese [daˈrabana] with Incorrect Foot Extrametricality

X			
(x	.)	>(x	.)>
*ˈda	ra	ba	na

Therefore, Lebanese requires moraic trochees with end rule right and syllable extrametricality, just like Palestinian, however based on Classical words it appears that the parsing directionality is right-to-left rather than left-to-right.

Despite syllable extrametricality, final CVXC syllables still receive stress in Lebanese, as shown in (266).

(266) Stressed Final Superheavy Syllables (Haddad 1984, p. 19, 23)

a. nazːˈal -t
bring.down.PFV -SBJ.1SG
'I brought down'

b. naːˈzal -t
encounter.PFV -SBJ.1SG
'I encountered'

c. maktaˈb -aːt
library -PL
'libraries'

d. makaːˈtiːb
letter.PL
'letters'

That these syllables are not rendered completely extrametrical by syllable extrametricality, even in the presence of plausible alternative feet as in all of the examples in (266), suggests that these CVXC syllables are behaving as two syllables, or a syllable plus an extrasyllabic element. There is no data in Haddad (1984) on the stressing of words with final CVCː, but there is no reason to think these behave differently to our expectations that such a syllable would receive stress as seen in the other varieties explored here.

3.3.1.5 Iraqi

Iraqi requires moraic trochees with End Rule Right and syllable extrametricality, just like the other varieties above in Pattern 1. However, for Iraqi and the subsequent varieties discussed under Pattern 1 there is is insufficient data to determine parsing directionality.

Given the words in (267) both have three syllables but different stress, Iraqi must be quantity sensitive.

(267) Quantity Sensitivity (Majdi 1988, p. 61)

a. mu.[ˈ]han.dis
   ‘engineer’

b. [ˈ]fa.ha.mak
   ‘he understood you’

In order to predict initial stress on [ˈfahamak], ‘he understood you’, moraic trochees are needed as iambs would predict unattested medial stress as shown in Figures 3.61–3.62.

Figure 3.61: Iraqi [ˈfahamak] with Correct Moraic Trochees

X		
(x	.)	<>
ˈfa	ha	mak

Figure 3.62: Iraqi [ˈfahamak] with Incorrect Iamb

	X	
(.	x)	<>
*fa	ˈha	mak

It must be End Rule Right to account for cases with more than one foot, such as [musˈtaʃfa], ‘hospital’ (Majdi 1988, p. 58), as in Figures 3.63–3.64. End Rule Right correctly predicts the penultimate stress in Figure 3.63, whereas End Rule Left predicts unattested initial stress in Figure 3.64.

Figure 3.63: Iraqi [musˈtaʃfa] with Correct ERR

	X	
(x)	(x)	
mus	ˈtaʃ	fa

Figure 3.64: Iraqi [musˈtaʃfa] with Incorrect ERL

X		
(x)	(x)	
*ˈmus	taʃ	fa

Extrametricality is needed to account for the lack of stress on final CVC syllables such as [ˈmustaqil], ‘independent’, and on final CVCV feet such as [ˈkaːtiba], ‘writer (f.)’ (Majdi 1988, p. 58). While either syllable or consonant extrametricality can account for [ˈmustaqil], syllable extrametricality is required to predict the initial stress in [ˈkaːtiba] as in Figure 3.65. There is no final consonant in [kaːtiba], so consonant extrametricality would predict unattested penultimate stress as in Figure 3.66 .

Figure 3.65: Iraqi [ˈkaːtiba] with Syllable Extrametricality

X		
(x)		<>
ˈkaː	ti	ba

Figure 3.66: Iraqi [kaːtiba] with Incorrect Consonant Extrametricality

	X	
(x)	(x	.)
*kaː	ˈti	ba

Final CVXC syllables are stressed despite syllable extrametricality, as shown in (268). There are no final CVCC words marked for stress in either Majdi (1988) or Erwin (1963); however there is no reason to believe these will behave differently to the other final CVCC syllables across Arabic varieties.

(268) Iraqi Final Superheavy Stress (Erwin 1963, p. 40)

a. taʕ.ˈbaːn
 ‘tired’

b. maˈħalː
 ‘place’

3.3.1.6 Makːan

Makːan requires moraic trochees built left-to-right with End Rule Right and syllable extrametricality. Similarly to Iraqi, there is no data available in Kabrah (2004) with which to determine parsing directionality.

It must be moraic trochees as this is the only way to get initial stress on [ˈkatabu], ‘they wrote’ (Kabrah 2004, p. 101), as shown in Figures 3.67-3.68.

Figure 3.67: Makːan [katabu] with Correct Moraic Trochee

X		
(x	.)	
ˈka	ta	bu

Figure 3.68: Makːan [katabu] with Incorrect Iamb

	X	
(.	x)	
*ka	ˈta	bu

End Rule Right is needed to account for words with more than one foot, such as [naːˈdeːt], 'I called' (Kabrah 2004, p. 36), in Figures 3.69–3.70. End Rule Right, predicts the attested antepenultimate stress in Figure 3.69, whereas End Rule Left predicts unattested initial stress in Figure 3.70.

Figure 3.69: Makːan [naːˈdeːt] with Correct ERR

	X
(x)	(x)
na:	ˈde:t

Figure 3.70: Makːan [naːˈdeːt] with Incorrect ERL

X	
(x)	(x)
*ˈnaa:	de:t

Given End Rule Right, extrametricality is needed to account for not stressing final CVCV feet in words such as [ˈsaːfaru], 'they travelled' (Kabrah 2004, p. 36), in Figures 3.71–3.73. Syllable extrametricality predicts the attested initial stress in Figure 3.71, whereas consonant extrametricality predicts unattested medial stress in Figure 3.72, and no extrametricality predicts unattested medial stress in Figure 3.73.

Figure 3.71: Makːan [ˈsaːfaru] with Correct Syllable Extrametricality

X		
(x)		<>
ˈsaː	fa	ru

Figure 3.72: Makːan [ˈsaːfaru] with Incorrect Consonant Extrametricality

	X	
(x)	(x	.)
*saː	ˈfa	ru

Figure 3.73: Makːan [ˈsaːfaru] with Incorrect No Extrametricality

		X
(x)	(x	.)
*saː	fa	ˈru

There are some cases in Makːan which appear to require consonant extrametricality rather than syllable extrametricality. A selection of these is included in (269). However, the next chapter will demonstrate that this is due to a preference for affixes to attach after a stressed or stressable syllable.

(269) Makːan Apparent Consonant Extrametricality Cases (Kabrah 2004, p. 20, 61)

a. bagaˈrat -u
cow -POSS.3SG.M
'his cow'

b. naːˈdi -i
club -POSS.3SG.M
'his club'

c. ʃaːˈl -at -u
carry.PFV -SBJ.3SG.F -OBJ.3SG.M
'she carried him'

d. ram -ˈat -u
throw.PFV -SBJ.3SG.F -OBJ.3SG.M
'she threw it'

Despite syllable extrametricality, final CVXC syllables receive stress, as shown in (270).

(270) Makːan Final Superheavy Stress (Kabrah 2004, p. 143, 9)

a. kiˈtaːb

'a book'

b. ka.ˈtab -t
write.PFV -SBJ.1SG
'I wrote'

Kabrah (2004) does not include polysyllabic cases with final CVCː syllables, but states that such syllables are heavy and stressable.

Therefore, here CVXC syllables appear to behave as two syllables, or one plus an extrasyllabic element.

3.3.1.7 Muscat

Muscat requires moraic trochees with end rule right and syllable extrametricality, just like the varieties discussed above. There is no data available in Glover (1988) with which to determine the parsing directionality.

Given that the following words have stress on different syllables despite the same number of syllables, Muscat must be quantity sensitive.

(271) Muscat Quantity Sensitive Stress (Glover 1988, p. 69)

a. ˈkan.sa.l -u
cancel.PFV -SBJ.3PL
'they cancelled'

b. qa.ˈbiː.lah
'tribe'

As stress is postlexical in Muscat Glover (1988, p. 238) provides limited stress data, and of this there is minimal data upon which to determine foot type. The initial stress on [ˈʕaʃa], 'supper' (Glover 1988, p. 69), might suggest moraic trochee rather than iamb as foot type as shown in Figures 3.74–3.77.

Figure 3.74: Muscat [ˈʕaʃa] with Correct Moraic Trochees

X	
(x	.)
ˈʕa	ʃa

Figure 3.75: Muscat [ˈʕaʃa] with Incorrect Iambs

	X
(.	x)
*ʕa	ˈʃa

The syllable extrametricality that will be motivated below means that this data point is less conclusive, as it is possible that it is an iamb but here the residue is stressed after extrametricality. Based on stress behaviour in clearly iambic varieties, this is less likely, as in LL and LH words in iambic varieties such as Wadi Ramː and Cyrenaican Bedouin, syllable extrametricality fails to apply as it would exhaust the foot and final iambic stress is found instead (cf. Wadi Ramː [liˈfa], 'he came' (Al Mashaqba 2015, p. 112), Cyrenaican [biˈna] 'he built', [nuˈxal] 'palm trees' (Watson 2011a, p. 3)). Further evidence supporting a moraic trochee analysis can be seen in patterns of syncope. Muscat syncopates unstressed vowels in non-final syllables as shown in (272).

(272) Muscat Unstressed Vowel Syncope (Glover 1988, p. 62)

a. ˈqalam
'pen'

b. /qalami/
[ˈqalmi]

‘my pen’

c. ˈgaːlis

‘sitting (m.s.)

d. /gaːlisah/

[ˈgaːlsah]

‘sitting (f.s.)’

For (272-b) to syncopate the medial vowel, it must be unstressed, suggesting a moraic trochee analysis as an iambic analysis would stress the medial vowel instead as shown below.

Figure 3.76: Muscat /ˈqalami/ with Correct Moraic Trochee

X		
(x	.)	<>
ˈqa	la	mi

Figure 3.77: Muscat /ˈqalami/ with Incorrect Iamb

	X	
(.	x)	<>
*qa	ˈla	mi

It must be End Rule Right with final syllable extrametricality to account for stress on the apparently middle foot in [daʃaˈdiʃ-hum], ‘their dishdashas’ (Glover 1988, p. 70), as shown in Figures 3.78–3.80. With End Rule Right and final syllable extrametricality, the attested penultimate stress is predicted in Figure 3.78. Without extrametricality, incorrect final stress is predicted in Figure 3.79. End Rule Left on the other hand predicts unattested initial stress in Figure 3.80. While consonant extrametricality could account for the final syllable of [daʃadiʃhum] not forming a foot, syllable extrametricality is needed in Figure 3.81 to predict the correct initial stress for [kansalu], ‘they cancelled’ (Glover 1988, p. 69).

Figure 3.78: Muscat [daʃaˈdiʃhum] with End Rule Right and Syllable Extrametricality

		X	
(x	.)	(x)	<>
da	ʃa	ˈdiʃ	hum

Figure 3.79: Muscat [daʃadiʃhum] with End Rule Right and Incorrect No Extrametricality

			X
(x	.)	(x)	(x)
*da	ʃa	diʃ	ˈhum

Figure 3.80: Muscat [daʃaˈdiʃhum] with Incorrect End Rule Left

X			
(x	.)	(x)	(x)
*ˈda	ʃa	diʃ	hum

Figure 3.81: Muscat [ˈkansalu] with Syllable Extrametricality

X		
(x)		<>
kan	sa	lu

Final CVXC syllables are stressed despite syllable extrametricality, as shown in (273).

(273) Muscat Final Superheavy Stress (Glover 1988, p. 70)

a. gaˈlas -t
sit.PFV -SBJ.1SG
'I sat down'

b. luˈbaːn

'frankincense'

c. ji -nˈkabː
SBJ.3 -pour.PASS.IMPF
'it is poured, pourable'

3.3.1.8 Qatari

Qatari Arabic requires moraic trochees with End Rule Right. There is definitely extrametricality, however it is unclear from the data whether this is consonant or syllable extrametricality, and the parsing directionality is similarly unclear.

It must be moraic trochees as depicted in Figure 3.82 as it is impossible to get stress on the initial light syllable of [ˈmaθalan], 'for example' (Al-Sulaiti 1993, p. 8), with an iamb as shown in Figure 3.83.

Figure 3.82: Qatari [ˈmaθalan] with Correct Moraic Trochee

X		
(x	.)	<>
ˈma	θa	lan

Figure 3.83: Qatari [ˈmaθalan] with Incorrect Iamb

	X	
(.	x)	<>
*ma	ˈθa	lan

It must be End Rule Right to account for words with more than one foot such as [maʃˈhuːra], 'famous' (Al-Sulaiti 1993, p. 8), as shown in Figures 3.84–3.85. End Rule Right correctly predicts the attested penultimate stress in Figure 3.84, whereas End Rule Left predicts unattested initial stress in Figure 3.84.

Figure 3.84: Qatari [maʃˈhuːra] with Correct ERR

	X	
(x)	(x)	
maʃ	ˈhuː	ra

Figure 3.85: Qatari [maʃˈhuːra] with Incorrect ERL

X		
(x)	(x)	
*ˈmaʃ	huː	ra

Extrametricality is needed in this variety to account for the lack of stress on final CVC syllables such as [ˈmaθalan], 'for example' (Al-Sulaiti 1993, p. 8), despite the End Rule Right motivated above. However, whether this is syllable or consonant extrametricality is unclear based on the limited data available in Al-Sulaiti (1993). For [ˈmaθalan], either analysis could predict the attested stress if left-to-right parsing directionality is assumed, as shown in Figures 3.86–3.87.

Figure 3.86: Qatari [ˈmaθalan] with Syllable Extrametricality

X		
(x	.)	<>
ˈma	θa	lan

Figure 3.87: Qatari [ˈmaθalan] with Consonant Extrametricality

X		
(x	.)	
ma	ˈθa	la< *n* >

The only data point available that suggests one analysis over the other is [zatˤˈtˤi-hum] 'devoured them' (Al-Sulaiti 1993, p. 223). This suggests consonant extrametricality, as shown in Figures 3.88-3.89 below.

Figure 3.88: Qatari [zatˤˈtˤi-hum] with Syllable Extrametricality

X		
(x)		<>
*ˈzatˤ	tˤi	hum

Figure 3.89: Qatari [zatˤˈtˤi-hum] with Consonant Extrametricality

	X	
(x)	(x	.)
zatˤ	ˈtˤi	hu< *m* >

However, as I motivate in Section , stress is often pulled to the syllable before affixes in Arabic, so this single data point is insufficient to determine which extrametrical unit is required.

Final CVXC syllables receive stress despite extrametricality, whether syllable or consonant, as shown in (274). There is no data in Al-Sulaiti (1993) on the stressing of words with final CVCː, but there is no reason to think these behave differently to our expectations that such a syllable would receive stress.

(274) Qatari Final Superheavy Stress (Al-Sulaiti 1993, p. 8)

a. ħaˈmiːr
   ‘donkey’

b. wiˈsˤal -t
   arrive.PFV -SBJ.1SG
   ‘I arrived’

Therefore, Pattern 1 corresponds to moraic trochees with end rule right and syllable extrametricality. There appears to be variation in parsing directionality, with some left-to-right parsing varieties, one right-to-left parsing variety, and several where there is insufficient data to determine.

3.3.1.9 Pattern 1b: Sˤanʕaːniː

Sˤanʕaːniː is very similar to the pattern described under Pattern 1 above and repeated below, as illustrated with the Sˤanʕaːniː data.

1. Stress final superheavy syllable,
2. Else stress penultimate heavy syllable,
3. Else stress the antepenultimate

(275) Final Superheavy Stress (Watson 2002, p. 81)

a. makˈtuːb
   ‘letter’

b. daˈrast
   ‘I /you m.s. learnt’

(276) Heavy Penultimate Stress (Watson 2002, p. 81)

a. ˈmadrasih
   ‘school’

b. saːfarat
   ‘she travelled’

c. kaˈtabnaː
   ‘we wrote’

d. maˈkaːtib
   ‘offices’, ‘libraries’

(277) Antepenultimate Stress (Watson 2002, p. 81)

a. ˈkatab
   ‘he wrote’

b. ˈlibisat
   ‘she wore’

c. makˈtabatiː
   ‘my library’

However, these generalisations fail to apply where the penultimate or antepenultimate syllable is CVː or ends in the left leg of a geminate consonant, as shown below in (278) where stress can fall even on the pre-antepenultimate syllable as in (278-f).

(278) CVː or CVC(.ː) (Watson 2002, p. 81)

a. maˈkaːtiːb
   ‘offices’

b. muˈdarːisiːn
   ‘teachers (m.)’

c. ˈxutˤːaːf
   ‘clasp’

d. ˈxaːrijiːn
   ‘going out (m.pl.)‘’

e. darˈrast

‘I / you (ms) taught’

f. ˈhaːkaðahaː

‘like this’

Watson (2002, pp. 81–82) analyses Sˤanʕaːniː as needing moraic trochees built left-to-right with End Rule Right and both foot and consonant extrametricality, as well as a preference to stress CVː and CVCː feet. Given the following pair have the same number of syllables but different stress, it must be quantity sensitive.

(279) Quantity Sensitive (Watson 2002, p. 81)

a. ˈlibisat

‘she wore’

b. miˈgambar

‘sitting’

I agree that moraic trochees built left-to-right are needed to account for the attested stress in [ˈragabat-ih], ‘his neck’ (Watson 2002, p. 82), as shown in Figures 3.90–3.93, where only the combination of moraic trochee and left-to-right parsing directionality in Figure 3.90 predicts the attested initial stress. Note in these diagrams I mark the final syllable as extrametrical — this will be motivated below.

Figure 3.90: Sˤanʕaːniː [ragabatih] with Correct Moraic Trochee L–R

X			
(x	.)		<>
ˈra	ga	ba	tih

Figure 3.91: Sˤanʕaːniː [ragabatih] with Incorrect Iamb L–R

	X		
(.	x)		<>
*ra	ˈga	ba	tih

Figure 3.92: Sˤanʕaːniː [ragabatih] with Incorrect Moraic Trochee R–L

	X		
	(x	.)	<>
*ra	ˈga	ba	tih

Figure 3.93: Sˤanʕaːniː [ragabatih] with Incorrect Iamb R–L

		X	
	(.	x)	<>
*ra	ga	ˈba	tih

End Rule Right is needed to account for words with more than one foot, such as [makˈtabatiː], ‘my library’ (Watson 2002, p. 81). This is shown in Figures 3.94-3.95 below, as End Rule Right predicts correct antepenultimate stress in Figure 3.94 whereas End Rule Left predicts incorrect initial stress in Figure 3.95. Note that both these analyses assume the final syllable is extrametrical, which shall be motivated next.

Figure 3.94: Sˤanʕaːniː [maktabatiː] with Correct ERR

	X		
(x)	(x	.)	<>
mak	ˈta	ba	tiː

Figure 3.95: Sˤanʕaːniː [maktabatiː] with Incorrect ERL

X			
(x)	(x	.)	<>
*ˈmak	ta	ba	tiː

Watson (2002) argues for both foot and consonant extrametricality, in order to account for data such as [ˈmadrasih], ‘school’ (Watson 2002, p. 81). This is the same issue as in Palestinian and the other Pattern 1 varieties, however Watson’s implementation of extrametricality is unusual. She claims that consonant extrametricality applies first, then the resulting foot (ra.si) is made extrametrical, as shown in Figure 3.96 below.

Figure 3.96: Sˤanʕaːniː [madrasih] with Foot and Consonant Extrametricality

X		
(x)	<(x	.)>
mad	ra	si<h>

However, this is a very unusual way to apply extrametricality. By making the final consonant extrametrical, the final foot is no longer peripheral so should not be available for extrametricality. Watson (2002, p. 99) claims that with consonant extrametricality, the final foot becomes peripheral; but then later on page 101 contradicts herself, arguing that a final extrasyllabic consonant in a CVXC syllable does render the final foot non-peripheral. Syllable extrametricality instead can solve this problem without need to apply extrametricality in an unusual manner, as shown in Figure 3.97 below.

Figure 3.97: Sˤanʕaːni: [madrasih] with Syllable Extrametricality

X		
(x)		<>
mad	ra	sih

This is the same solution that I have posited for the other Pattern 1 varieties.

Sˤanʕaːni: does display an unusual (for Arabic) preference for stressing CVː and CVG syllables where the G is the first leg of an ambisyllabic geminate. As such, during affixation stress does not move in between (280-a)–(280-b) or (280-c)–(280-d) but does in (281-a)–(281-b) and (281-c)–(281-d).

(280) Sˤanʕaːni: No Stress Shift during Affixation (Watson 2002, p. 103)

a. ʔa -ˈdawːir
sbj.1sg -look.for.impf
'I look for'

b. ʔa -ˈdawːir -hum
sbj.1sg -look.for.impf -obj.3pl.m()
'I look for them (m.)'

c. ˈgaːl -at
say.pfv -sbj.3sg.f
'she said'

d. ˈgaːl -at -l -ih
say.pfv -sbj.3sg.f -dat -obj.3sg.m
'she said to him'

(281) Sˤanʕaːni: Stress Shift during Affixation (Watson 2002, p. 103)

a. ˈdˤarab -∅
hit.pfv -sbj.3sg.m
'he hit'

b. dˤaˈrab -∅ -kum
hit.pfv -sbj.3sg.m -obj.2pl.m
'he hit you (m. pl.)'

c. ˈarxasˤ

'cheaper'

d. arˈxasˤ -l -ih
cheaper -dat -obj.3sg.m
'cheaper for him'

This preference does not outrank syllable extrametricality — which applies to both CVː and CVG syllables alike, as shown in (282).

(282) Sˤanʕaːniː Syllable Extrametricality Regardless of CVX Type (Watson 2002, p. 103, 82)

a. ʔaˈsaːmiː
'names'

b. ˈlaflaf
'he collected'

Watson accounts for these patterns by suggesting a preference for feet built with underlying rather than weight by position assigned mora. Davis and Ragheb (2014) takes this further, arguing Weight by Position only applies in words without CVː or CVCː syllables. Alternatively, it may be that there is a preference to stress CVː or CVG syllables where these are present in the stem of the word.

When the final syllable is CVXC, the behaviour is mixed. If it is a CVCC syllable, it escapes final syllable extrametricality, suggesting it is behaving as two syllables, or a syllable plus an extrasyllabic element, as shown in (283).

(283) Final CVCC Stressed Despite Syllable Extrametricality (Watson 2002, p. 82)

a. gamˈbar -t
sit.PFV -SBJ.1SG
'I sat'

b. lafˈlaf -t
sit.PFV -SBJ.1SG
'I collected'

However, when the final syllable is CVːC, it is made extrametrical if there is a CVː or CVG foot left to receive stress as in (284-a)–(284-b), but it is not extrametrical in (284-c)–(284-d) where such syllables are not found.

(284) Final CVːC Not Stressed (Watson 2002, p. 81-2)

a. baˈsaː<tiːn>
  ‘gardens’

b. muˈdarːi<siːn>
  ‘teachers (m.)’

c. baˈnaːt
  ‘girls’

d. makˈtuːb
  ‘letter’

3.3.2 Pattern 2

Rwaili, Haːyil, and Wadi Ramː have a similar pattern of stress assignment to Pattern 1, however these Bedouin varieties require iambic feet rather than moraic trochees. The broad pattern of stress assignment is as follows, as summarised by Al Mashaqba (2015, pp. 111–113) illustrated here with data from Wadi Ramː. Note that element 3 in this list, the behaviour of LH/LL disyllabic words, is typically of Wadi Ramː and other Bedouin varieties including Cyrenaican and Negev Bedouin, but does not hold for Rwaili and Haːyil discussed below. It is not that either case is particularly unusual for Bedouin varieties; as Rosenhouse (2011) notes, both initial and final stress on disyllabic words are found across Bedouin varieties.

1. Stress final superheavy syllable,
2. Else stress rightmost non-final heavy syllable,
3. Else in disyllabic words with initial light syllable, stress final syllable
4. Else stress the penultimate

(285) Final Superheavy Stress (Al Mashaqba 2015, p. 111)

a. mir.ˈyaːʕ
  ‘ram leading the herd’

b. si.ˈbag -t
  forerun.PFV -SBJ.1SG
  ‘I foreran’

c. mʕalː.ˈm -iːn
teacher -PL.M
'teachers (m.pl.)

(286) Non-final Heavy Stress (Al Mashaqba 2015, p. 112)

a. ˈaː.gil

'newly pregnant'

b. ˈya -x.ti.ris
SBJ.3SG.M -be.frightened.IMPF
'he is frightened'

c. miʕ.ni.ˈgiy.yih

'pure-bred horse'

(287) Disyllabic LL/LH Final Stress (Al Mashaqba 2015, p. 112)

a. li.ˈfa -∅
come.PFV -SBJ.3SG.M
'he came'

b. da.ˈħal

'dune'

c. zi.ˈbil

'sheep muck'

(288) Penultimate Stress (Al Mashaqba 2015, p. 113)

a. fi.ˈlu. -na
mare -POSS.1PL
'our mare'

b. ʔa.ˈʕa.war

'one eyed'

c. mu.ˈtˤa.rah

'water bottle'

d. da.ˈħa.lih

'quicksand'

3.3.2.1 Wadi Ramː

Wadi Ramː requires iambs built left-to-right with End Rule Right and extrametricality. Al Mashaqba (2015) argues that foot extrametricality is required. However, I demonstrate that, just like Palestinian, syllable extrametricality is sufficient to account for this data.

Wadi Ramː must have iambs is this is the only way to get medial stress on a word starting with two light syllables, such as [daˈħalih], 'quicksand' (Al Mashaqba 2015, p. 113), assuming syllable extrametricality, as in Figures 3.99–**??**. Iambs correctly predict the attested final stress in Figure 3.99 , whereas moraic trochees would predict unattested initial stress in Figure **??**.

Figure 3.98: Wadi Ramː[daˈħalih] with Correct Iamb

	X	
(.	x)	<>
da	ˈħa	lih

Figure 3.99: Wadi Ramː[daˈħalih] with Incorrect Moraic Trochee

X		
(x	.)	<>
*ˈda	ħa	lih

End Rule Right is needed to account for words with more than one foot, such as [maːsˈkiːn], 'they (m.) are catching' (Al Mashaqba 2015, p. 112), as in Figures 3.100 -3.101. End Rule Right correctly predicts the attested final stress in Figure 3.100, whereas End Rule Left predicts unattested initial stress in Figure 3.101.

Figure 3.100: Wadi Ramː [maːskiːn] with Correct ERR

	X
(x)	(x)
maːs	ˈkiːn

Figure 3.101: Wadi Ramː [maːskiːn] with Incorrect ERL

X	
(x)	(x)
*ˈmaːs	kiːn

Given that End Rule Right is needed, the only way to predict the antepenultimate stress in [naˈʕajat-i], 'my ewe' (Al Mashaqba 2015, p. 113) , is using left-to-right parsing with syllable extrametricality, as shown in Figure 3.102. Figures 3.103–3.105 demonstrate that with right-to-left parsing and/or no extrametricality, the correct stress position cannot be predicted.

Figure 3.102: Wadi Ram: [naˈʕajat-i] with Correct Left-to-Right Parsing and Syllable Extrametricality

	X		
(.	x)		<>
na	ˈʕa	ja	ti

Figure 3.103: Wadi Ram: [naˈʕajat-i] with Incorrect Right-to-Left Parsing and Syllable Extrametricality

		X	
	(.	x)	<>
*na	ʕa	ˈja	ti

Figure 3.104: Wadi Ram: [naˈʕajat-i] with Incorrect Left-to-Right Parsing and No Extrametricality

			X
(.	x)	(.	x)
*na	ʕa	ja	ˈti

Figure 3.105: Wadi Ram: [naˈʕajat-i] with Incorrect Right-to-Left Parsing and No Extrametricality

			X
(.	x)	(.	x)
*na	ʕa	ja	ˈti

Where the word is disyllabic, syllable extrametricality does not apply and thus final stress can occur on a light or heavy syllable as shown in (289) and Figures 3.106–3.107. This is found in other Bedouin iambic varieties such as Negev and Cyrenaican, but not in Rwaili and Haːyil discussed below.

(289) Wadi Ram: Failure of Extrametricality in Disyllabic Words (Al Mashaqba 2015, p. 112)

a. liˈfa -∅
come.PFV -SBJ.3SG.M
'he came'

b. dˈħal
'dune'

c. ziˈbil
'sheep muck'

Figure 3.106: Wadi Ram: [liˈfa] with Correct Iamb

	X
(.	x)
li	ˈfa

Figure 3.107: Wadi Ram: [liˈfa] with Extrametricality

X	
(x)	<>
*ˈli	fa

There is no need to use foot extrametricality to account for this data when syllable extrametricality works instead. Below I will demonstrate that choosing syllable extrametricality also has typological benefits.

Final CVXC syllables are stressed despite syllable extrametricality, as shown in (290). There are no polysyllabic final CVCː words marked for stress in Al Mashaqba (2015); however there is no reason to believe these will behave differently to the other final CVCː syllables across Arabic varieties.

(290) Wadi Ram: Final Superheavy Stress (Al Mashaqba 2015, p. 111)

a. mirˈyaːʕ
'ram leading the heard'

b. siˈbag -t
forerun.PFV -SBJ.1SG
'I foreran'

This suggests that final CVXC syllables are behaving as two syllables, or as a syllable plus an extrasyllabic element.

3.3.2.2 Rwaili

Rwaili requires iambs with End Rule Right and syllable extrametricality to account for its stress patterns. There is no data in Prochazka Jr. (1988) to determine parsing directionality and foot type is inconclusive, however data from Al Solami (2020, pp. 104–106) strongly argues for iambs built left-to-right in this variety. Although Rwaili does require a different foot to many of the other varieties in Pattern 1, the surface patterns only differ in a very small number of cases.

Iambs would account for the medial stress on [ʔaˈkalat] 'she ate' and [ʔaˈxaðu] 'they took'Al Solami (2020, pp. 104–106), whereas moraic trochees would predict initial stress as shown below.

Figure 3.108: Rwaili [ʔaˈkalat] with Correct Iamb

	X	
(.	x)	<>
ʔa	ˈka	lat

Figure 3.109: Rwaili [ʔaˈkalat] with Incorrect Moraic Trochee

X		
(x	.)	<>
*ˈʔa	ka	lat

Figure 3.110: Rwaili [ʔaˈxadu] with Correct Iamb

	X	
(.	x)	<>
ʔa	ˈxa	ðu

Figure 3.111: Rwaili [ʔaˈxadu] with Incorrect Moraic Trochee

X		
(x	.)	<>
*ˈʔa	xa	ðu

These iambs must be built left-to-right to account for the antepenultimate stress on [ʔaˈkalitih] 'she ate it (m.)', [ʔaˈxaðitih] 'she took it', and [ʔaˈmaritih] 'she ordered him' Al Solami (2020, pp. 104–106), as right-to-left parsing predicts unattested penultimate stress as shown in Figures 3.112–3.113.

Figure 3.112: Rwaili [ʔaˈkalitih] with Correct Iamb and Left-to-Right Parsing

	X		
(.	x)		<>
ʔa	ˈka	li	tih

Figure 3.113: Rwaili [ʔaˈkalitih] with Incorrect Iamb and Right-to-Left Parsing

		X	
	(.	x)	<>
*ʔa	ka	ˈli	tih

End Rule Right is needed to account for words with more than one foot, such as [jaʃruˈbuːn], 'they drink' (Prochazka Jr. 1988, p. 21), in Figures 3.114–3.115. End Rule Right predicts the attested antepenultimate stress in Figure 3.114, whereas End Rule Left predicts unattested initial stress in Figure 3.115.

Figure 3.114: Rwaili [jaʃruˈbuːn] with Correct ERR

		X
(x)		(x)
jaʃ	ru	ˈbuːn

Figure 3.115: Rwaili [jaʃruˈbuːn] with Incorrect ERL

X		
(x)		(x)
*ˈjaʃ	ru	buːn

Given End Rule Right, extrametricality is needed to account for not stressing final CVC syllables in words such as [ˈjiʃtirin], 'they (f.) buy' Prochazka Jr. (1988, pp. 100), in Figures 3.116–3.118. Syllable extrametricality predicts the attested initial stress in Figure 3.116, whereas consonant extrametricality predicts unattested medial stress in Figure 3.117, and no extrametricality predicts unattested final stress in Figure 3.118.

Figure 3.116: Rwaili [ˈjiʃtirin] with Correct Syllable Extrametricality

X		
(x)		<>
ˈjiʃ	ti	rin

Figure 3.117: Rwaili [jiʃtirin] with Incorrect Consonant Extrametricality

	X	
(x)	(x	.)
*jiʃ	ˈti	ri<n>

Figure 3.118: Rwaili [jiʃtirin] with Incorrect No Extrametricality

		X
(x)		(x)
*jiʃ	ti	ˈrin

Unlike Bedouin Cyrenaican and Wadi Ramː, syllable extrametricality can apply in LL and LH words, giving initial stress, as shown in (291) with data from Al Solami (2020, p. 110) (cf. Wadi Ramː [liˈfa], 'he built' (Al Mashaqba 2015, p. 112), Cyrenaican [biˈna] 'he built', [nuˈxal] 'palm trees' (Watson 2011a, p. 3)).

(291) Syllable Extrametricality Produces Initial Stress in Disyllabic Words Al Solami (2020, p. 110)

a. ˈliga -∅
find.PFV -SBJ.3SG.M
'he found'

b. ˈkitab
'book'

c. ˈdifaʕ -∅
push.PFV -SBJ.3SG.M
'he pushed'

Despite syllable extrametricality, final CVCC and CVːC syllables still receive stress, as in (292).

(292) Rwaili Final Superheavy Stress (Prochazka Jr. 1988, p. 21)

a. siˈmiʕ -t
hear.PFV -SBJ.1SG
'I heard'

b. ja -ʃraˈb -uːn
SBJ.3 -drink.IMPF -SBJ.PL.M
‘they drink’

However, final CVCː syllables degeminate and are not stressed, as in (293).

(293) Rwaili Final Degemination (Prochazka Jr. 1988, p. 21)

a. ˈji -stimir
SBJ.3 -continue.IMPF
‘he continues’

b. ˈji -zim
SBJ.3 -carry.IMPF
‘you (m.s.) carry’

Therefore, where syllables retain their CVXC status, they behave like two syllables or a syllable plus an extrasyllabic element.

3.3.2.3 Haːyil

Haːyil requires iambs with End Rule Right and syllable extrametricality to account for its stress patterns. There is no data in Prochazka Jr. (1988) to determine parsing directionality.

It must be iambs as this is the only way to get penultimate stress on a light syllable in words like [ʔaˈʕatˤaʃ], ‘I become thirsty’ (Prochazka Jr. 1988, p. 37) if we assume the syllable extrametricality motivated below.

Figure 3.119: Haːyil [ʔaˈʕatˤaʃ] with Correct Iamb

	X	
(.	x)	<>
ʔa	ˈʕa	tˤaʃ

Figure 3.120: Haːyil [ʔaˈʕatˤaʃ] with Incorrect Moraic Trochee

X		
(x	.)	<>
*ˈʔa	ʕa	tˤaʃ

End Rule Right is needed to account for words with more than one foot, such as [jaʃruˈbuːn], ‘they drink’ (Prochazka Jr. 1988, p. 21), in Figures 3.121–3.122. End Rule Right predicts the attested antepenultimate stress in Figure 3.121, whereas End Rule Left predicts unattested initial stress in Figure 3.122.

Figure 3.121: Haːyil [jaʃruˈbuːn] with Correct ERR

		X
(x)		(x)
jaʃ	ru	ˈbuːn

Figure 3.122: Haːyil [jaʃruˈbuːn] with Incorrect ERL

X		
(x)		(x)
*ˈjaʃ	ru	buːn

Given End Rule Right, extrametricality is needed to account for not stressing final CVC syllables in words such as [ˈʔal-ɟimal], ‘the camel’ (Prochazka Jr. 1988, p. 21), in Figures 3.123–3.125. Syllable extrametricality predicts the attested initial stress in Figure 3.123, whereas consonant extrametricality predicts unattested medial stress in Figure 3.124, and no extrametricality predicts unattested final stress in Figure 3.125. Note in Haːyil unlike many other Arabic varieties the definite article forms a prosodic word with the noun it syntactically dominates, and as such can receive stress.

Figure 3.123: Haːyil [ˈʔal-ɟimal] with Correct Syllable Extrametricality

X		
(x)		<>
ˈʔal	ɟi	mal

Figure 3.124: Haːyil [ˈʔal-ɟimal] with Incorrect Consonant Extrametricality

	X	
(x)	(x	.)
*ʔal	ˈɟi	ma<l>

Figure 3.125: Haːyil [ˈʔal-] with Incorrect No Extrametricality

		X
(x)		(x)
*ʔal	ɟi	ˈmal

Just as in Rwaili, in disyllabic words syllable extrametricality still occurs to produce initial stress as shown in words like [ˈnisi], ‘he forgot’ (Prochazka Jr. 1988, p. 92), in Figures 3.126–3.127.

Figure 3.126: Haːyil [ˈnisi] with Syllable Extrametricality

X	
(x)	<>
ˈni	si

Figure 3.127: Haːyil [niˈsi] with No Extrametricality

	X
(.	x)
*ni	ˈsi

Despite syllable extrametricality, final CVCC and CVːC syllables still receive stress, as in (294).

(294) Haːyil Final Superheavy Stress (Prochazka Jr. 1988, p. 21)

a. kiˈtab -t
write.PFV -SBJ.1SG
'I wrote'

b. ja -ʃraˈb -uːn
SBJ.3 -drink.IMPF -SBJ.PL.M
'they drink'

However, as in Rwaili, final CVCː syllables degeminate so are not stressed, as in (295).

(295) Haːyil Final Degemination (Prochazka Jr. 1988, p. 21)

a. ˈji -stimir
SBJ.3 -continue.IMPF
'he continues'

b. ˈji -mid
SBJ.3 -spread.IMPF
'he spreads'

Therefore, where syllables retain their CVXC status, they behave like two syllables or a syllable plus an extrasyllabic element.

3.3.3 Pattern 3

Cairene and Tunisian differ from the varieties discussed in Patterns 1 and 2 above. They share the following pattern of stress assignment, as described by Watson (2002, p. 80) and illustrated here with data from Cairene:

1. Stress final superheavy syllable or CVː

2. Else stress penultimate heavy syllable,

3. Else stress the penultimate or antepenultimate, whichever is separated by 0 or an even number of syllables from either the closest preceding heavy syllable or the start of the word

(296) Final Superheavy or CVː Stress (Watson 2002, p.80)

a. kaˈtabt
   ‘I wrote’
b. kiˈtaːb
   ‘book’
c. falaˈħiːn
   ‘peasants’
d. haˈyaː
   ‘life’

(297) Heavy Penultimate Stress (Watson 2002, p.80)

a. ˈbeːtah
   ‘his house’
b. muˈdarːis
   ‘teacher’
c. darˈrasni
   ‘he taught me’

(298) Penultimate Stress (Watson 2002, p.80)

a. mudarˈrisa
   ‘teacher (f.)
b. madˈrasa
   ‘school’
c. ˈfihim
   ‘he understood’

d. kataˈbitu

‘she wrote it’

(299) Antepenultimate Stress (Watson 2002, p.80)

a. inˈkasarit

‘it (f.) was broken)’

b. yixˈtalifu

‘they differ’

c. ˈdarasit

‘she learnt’

d. ˈkataba

‘scribes’

3.3.3.1 Cairene

According to Hayes (1995), Cairene displays moraic trochees built left-to-right with End Rule Right and consonant extrametricality. This is very similar to the varies discussed under Pattern 1, except for the unit of extrametricality.

Indeed, it must be moraic trochees, based on the data in (300).

(300) Moraic Trochees (Watson 2002, p. 80)

a. daˈrasni

‘he taught me’

b. ˈkatabu

‘they wrote’

Since both words contain three syllables and stress falls on different syllables, Palestinian must be quantity sensitive. Furthermore, while (300-a) is ambiguous as to whether it is an iamb or a trochee as demonstrated in Figures 3.15–3.16, (300-b) must be a trochee as demonstrated in Figures 3.17–3.18.

Figure 3.128: Palestinian [darasni] with Moraic Trochee

	X	
	(x)	
da	ˈras	na

Figure 3.129: Palestinian [darasni] with Iamb

	X	
(.	x)	
ka	ˈtab	na

Figure 3.130: Cairene [katabu] with Correct Moraic Trochee

X		
(x	.)	
ˈka	ta	bu

Figure 3.131: Cairene [katabu] with Incorrect Iamb L−R

	X	
(.	x)	
*ka	ˈta	bu

Left-to-right parsing is also needed to account for the initial stress in [ˈkatabu], 'they wrote', as shown in Figures 3.132–3.133, where only the left-to-right parsing directionality in Figure 3.130 predicts the attested initial stress.

Figure 3.132: Cairene [katabu] with Correct L−R

X		
(x	.)	
ˈka	ta	bu

Figure 3.133: Cairene [katabu] with Incorrect R−L

		X
	(.	x)
*ka	ta	ˈbu

It must be End Rule Right as this correctly predicts the attested medial stress in [muʃˈkila], 'problem' (Watson 2002, p. 94), as shown in Figure 3.134, whereas End Rule Left incorrectly predicts initial stress as shown in Figure 3.135.

Figure 3.134: Cairene [muʃkila] with Correct End Rule Right

	X	
(x)	(x	.)
muʃ	ˈki	la

Figure 3.135: Cairene [muʃkila] with Incorrect End Rule Left

X		
(x)	(x	.)
*muʃ	ˈki	la

Cairene has final consonant extrametricality. This explains why final CVC appears light and not stressed in (301-b), whereas medial CVC is heavy and stressable in (301-a).

(301) Cairene Extrametricality (Watson 2002, p.80)

a. muˈdarris

‘teacher’

b. ˈbintik

‘your daughter’

The metrical structure of (301-b) can be seen in Figure 301 below.

Figure 3.136: Cairene [bintik]

with Consonant Extrametricality

X	
(x)	
ˈbin	ti<k>

However, it is not that syllable extrametricality is needed, as final light syllables as in [maktaba], ‘library’ (Watson 2002, p. 94), still form part of final stressable feet, as shown in Figure 3.137 below.

Figure 3.137: Cairene [maktaba] without Syllable Extrametricality

	X	
(x)	(x	.)
mak	ˈta	ba

The different in extrametrical units thus gives us the following difference:

(302) Cairene vs Palestinian

a. ˈmadrasa

school (Cairene, consonant extrametricality)

b. madˈra<sa>

school (Palestinian, syllable extrametricality)

Therefore, Cairene displays moraic trochees built left-to-right with End Rule Right and consonant extrametricality.

Final CVXC syllables receive stress in Cairene. This is not surprising in the case of CVːC and CVCC syllables, as making the final consonant extrametrical still leaves a heavy syllable that be footed and stressed, as shown in (303).

(303) Cairene Superheavy CVːC and CVCC Syllables (Watson 2002, p.80)

a. ka.ˈtab -<t>
write.PFV -SBJ.1SG
'I wrote'

b. falaˈħiːn

'peasants'

However, the case of final CVCː syllables is more informative as to the shape of these final syllables. Davis and Ragheb (2014, p. 4) cites the following minimal pair in Example (304) where the form without a final geminate in (304-a) receives initial stress whereas the form in (304-b) with a final geminate receives final stress.

(304) Cairene Final Geminate Stress (Davis and Ragheb 2014, p. 4)

a. ˈʔamal

'hope'

b. ʔaˈmalː

'more/most boring'

I assume a geminate representation where there is one consonant on the segmental tier that is associated to two moras as in Figure 3.139, rather than two separate consonants on the segmental tier as in the case of non-geminate clusters shown in Figure **??**.

Figure 3.138: Structure of Consonant Clusters

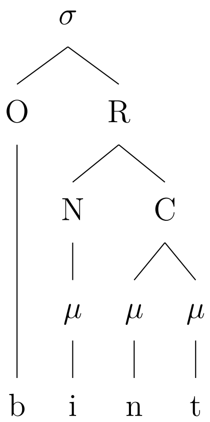
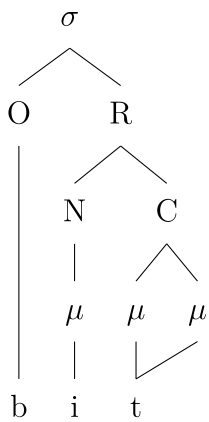


Figure 3.139: Structure of Geminates



Therefore, if the geminate cluster is present within a single final syllable, we would expect the whole geminate to be made extrametrical by final consonant extrametricality. However, if this was the case, we would not expect a stress difference between (304-a) and (304-b). This suggests that final CVXC syllables are behaving as two syllables, or a syllable plus an extrasyllabic element, with the final geminate split across the two structures and there is something present to make the final C non-peripheral.

3.3.3.2 Tunisian

Tunisian has the same stress pattern as Cairene. Tunisian requires moraic trochees built left-to-right with End Rule Right and consonant extrametricality. Given the following pair have the same number of syllables but different stress, Tunisian is quantity sensitive.

(305) Quantity Sensitivity (Maamouri 1967, p. 20-22)

a. bulaˈda:
’tiresome people’

b. ˈʃarika
‘company’

Tunisian requires moraic trochees built left-to-right as this is the only way to predict the attested initial stress in [ˈʃarika], 'company' (Maamouri 1967, p. 22), as shown in Figures 3.140–3.143, where only the combination of moraic trochee and left-to-right parsing directionality in Figure 3.140 predicts the attested initial stress.

Figure 3.140: Tunisian [ʃarika] with Correct Moraic Trochee L−R

X		
(x	.)	
ˈʃa	ri	ka

Figure 3.141: Tunisian [ʃarika] with Incorrect Iamb L−R

	X	
(.	x)	
*ʃa	ˈri	ka

Figure 3.142: Tunisian [ʃarika] with Incorrect Moraic Trochee R−L

	X	
	(x	.)
*ʃa	ˈri	ka

Figure 3.143: Tunisian [ʃarika] with Incorrect Iamb R−L

		X
	(.	x)
*ʃa	ri	ˈka

Tunisian requires End Rule Right to correctly predict final stress in [bulaˈdaː], 'tiresome people' (Maamouri 1967, p. 20), as shown in Figures 3.144–3.145. End Rule Right in Figure 3.144 predicts the attested final stress whereas End Rule Left in Figure 3.145 predicts unattested initial stress.

Figure 3.144: Tunisian [buladaː] with Correct ERR

		X
(x	.)	(x)
bu	la	ˈdaː

Figure 3.145: Tunisian [buladaː] with Incorrect ERL

X		
(x	.)	(x)
*ˈbu	la	daː

Consonant extrametricality is needed to account for why final stress occurs on the final CVː in [bulaˈdaː], 'tiresome people'(Maamouri 1967, p. 20), but not on the final CVC in [ˈmsaːmir], 'nails' (Maamouri 1967, p. 21). This is shown in Figure 3.146 and3.147.

Figure 3.146: Tunisian [buladaː] with Consonant Extrametricality

		X
(x	.)	(x)
bu	la	ˈdaː

Figure 3.147: Tunisian [msaːmir] with Consonant Extrametricality

X	
(x)	
ˈmsaː	mi< *r* >

Final CVXC syllables receive stress in Tunisian, as shown in (306), whether they are CVːC, CVCC, or CVCː.

(306) Tunisian Final Stressed Superheavy Syllables (Maamouri 1967, p. 20)

a. liˈmuːn

‘lemons’

b. qaːyˈm -iːn
rising -PL
‘rising (pl.)’

c. diˈmaʃq

‘Damascus’

d. labːˈis -t
clothe.PFV -SBJ.1SG
‘I clothed’

e. ʔaˈqalː

‘less’

f. muʕtaˈmadː

‘head of administrative district’

That consonant extrametricality still leaves a stressable heavy foot in surface CVːC and CVCC syllables is not surprising. However, (306-e)–(306-f) shown that geminate clusters are not rendered extrametrical, as shown for Cairene above. This suggests that in some way the final C of these final CVXC syllables is non-peripheral.

3.3.4 Pattern 4

Moroccan Casablanca and Libyan Tripoli share the following pattern of stress assignment, which can be described as follows illustrated with data from Moroccan Casablanca:

1. Stress final CVC syllable
2. Else, stress the penultimate syllable

(307) Final CVC Stress (Boudlal 2001, p. 13)

a. lawˈyin
   ‘wilted (m.pl)’
b. liˈmun
   ‘oranges’
c. mərmədˈnak
   ‘we somethingd you’

(308) Penultimate Stress (Boudlal 2001, p. 13)

a. rəmla
   ‘sand’
b. ˈmayəl
   ‘bent (m.)‘’
c. wəl"datni
   ‘she gave birth to me’

The only exception to this is words where the final syllable is an object suffix, such as those in (309). This exceptional behaviour will be discussed in Section 4.5.

(309) Exceptions (Boudlal 2001, p. 13)

a. ˈkalkum
   ‘all of you’
b. lawˈyinhum
   ‘they are twisting them’
c. dˤrəbhum
   ‘he hit them’

3.3.4.1 Moroccan Casablanca

Moroccan Casablanca requires moraic trochees to account for initial stress in words like [ˈwalu], ‘nothing’ (Boudlal 2001, p. 115), and medial stress on [minˈʒara], ‘sharpener’ (Boudlal 2001, p. 117). This is shown in Figures 3.148–3.149, where moraic trochees in Figure 3.148 predict the attested initial stress, whereas iambs in Figure 3.149 predict unattested final stress.

Figure 3.148: Moroccan Casablanca [walu] with Correct Moraic Trochee

```
 X
 (x   .)
 ˈwa  lu
```

Figure 3.149: Moroccan Casablanca [walu] with Incorrect Iamb

```
        X
  (.    x)
 *wa   ˈlu
```

Given that Moroccan Casablanca uses moraic trochees, right-to-left parsing directionality is needed to predict the medial stress in [liˈmuna], orange (Boudlal 2001, p. 121). This is shown in Figures 3.150–3.151, where right-to-left parsing in Figure 3.150 predict the attested medial stress, whereas left-to-right parsing in Figure 3.151 predicts unattested initial stress. Although [limuna] is a loanword, it has been integrated into the phonology sufficiently to have a productive diminutive of the form [lwimina], (Boudlal 2001, p. 293).

Figure 3.150: Moroccan Casablanca [limuna] with Correct R−L

```
       X
       (x   .)
 li   ˈmu   na
```

Figure 3.151: Moroccan Casablanca [limuna] with Incorrect L−R

```
  X
  (x    .)
 *ˈli   mu   na
```

No extrametricality is needed to account for the stress in Moroccan Casablanca, as shown by the stressed final CVC syllables in (310).

(310) Moroccan Casablanca No Extrametricality (Boudlal 2001, p. 122)

a. lawˈj -in
   wilted -PL
   ‘wilted’

b. məl.ˈjun

‘1 million’

End Rule Right is needed to account for the final stress in words with more than one syllable, such as [mandˤaˈrin], ‘clementine’ (Boudlal 2001, p. 122). This is shown in Figures 3.152–3.153, where End Rule Right in Figure 3.152 predicts the attested final stress, whereas End Rule Left in Figure 3.153 predicts unattested initial stress.

Figure 3.152: Moroccan Casablanca [mandˤaˈrin] with Correct ERR

		X
(x)		(x)
man	dˤa	ˈrin

Figure 3.153: Moroccan Casablanca [mandˤaˈrin] with Incorrect ERL

X		
(x)		(x)
*ˈman	dˤa	rin

Therefore, Moroccan Casablanca can be accounted for with moraic trochees built left to right with end rule rule and no extrametricality. The parsing directionality and lack of extrametricality is significantly different the varieties discussed above under Patterns 1 and 2.

3.3.4.2 Libyan Tripoli

As discussed in the previous chapter, Libyan Tripoli does not have long vowels, but rather short vowels and schwas. There is a distinction between these two, in that short vowels contribute one mora, but schwa contributes no mora. This is clear in (311). If schwa and short vowels contributed weight equally, we would not expect any difference between CəCCəC and CVCCVC words. However, in (311) we find just that — in (311-a), the vowels are both schwa and receive initial stress, whereas in (311-b) they are both short vowels and receive final stress. CəCCəC behaves like CVCV, so has the same initial stress as in (311-c). This can be best accounted for by the coda consonant providing weight by position, but not the schwa, thus making both CəCCəC and CVCV words with two monomoraic syllables.

(311) Libyan Tripoli Short Vowel and Schwa Weight Asymmetry (Yoda 2005, p. 111-112)

a. ˈtˤəbbəħ
call.IMP.M
'call (m.)!'

b. sakˈn -in
living -M.PL
'living (m.pl.)

c. ˈħuʃa

'house'

Last chapter I motivated the benefit of associating 'short' vowels to two timing slots in order to account for the distribution of CVXC syllables. However, this double length does not match the weight distribution. Therefore, there is an asymmetry between weight and length in this variety. As such, X-slots are needed as an additional layer of representation in addition to moras.

Libyan Tripoli requires moraic trochees built right-to-left with End Rule Right and no extrametricality.

There must be moraic trochees as it is the only way to get initial stress on words with two light syllables, such as [ˈyənsa], 'he forgets' (Yoda 2005, p. 114), as shown in Figures 3.154–3.155. Moraic trochees predict the attested initial stress in Figure 3.154, whereas iambs predict unattested final stress in Figure 3.155. Recall that schwa carries no mora so it is the coda consonant that is contributing the sole mora in the first light syllable.

Figure 3.154: Libyan Tripoli [ˈyənsa] with Correct Moraic Trochee

X	
(x	.)
ˈyən	sa

Figure 3.155: Libyan Tripoli [ˈyənsa] with Incorrect Iamb

	X
(.	x)
*yən	ˈsa

It must be right-to-left parsing to account for medial stress in words with three light syllables, such as [yəʧˈkəsːər], 'it is broken' (Yoda 2005, p. 111), as in Figures 3.156 -3.157. Right-to-left parsing predicts the attested medial stress in Figure 3.156, whereas left-to-right parsing predicts unattested initial stress in Figure 3.157.

Figure 3.156: Libyan Tripoli [yətʃˈkəsːər] with Correct R−L Parsing

	X	
	(x	.)
yətʃ	ˈkəs	sər

Figure 3.157: Libyan Tripoli [yətʃˈkəsːər] with Incorrect L−R Parsing

X		
(x	.)	
*ˈyətʃ	kəs	sər

It must be End Rule Right to account for words with more than one foot, such as [sakˈnin], 'living (m.pl.)' (Yoda 2005, p. 111), as in Figures 3.158 -3.159 . End Rule Right predicts the attested final stress in Figure 3.158, whereas End Rule Left predicts unattested initial stress in Figure 3.159.

Figure 3.158: Libyan Tripoli [saknin] with Correct ERR

	X
(x)	(x)
sak	ˈnin

Figure 3.159: Libyan Tripoli [saknin] with Incorrect ERL

X	
(x)	(x)
*ˈsak	nin

In order to get stress on a final CVC such as [saknin], 'living (m.pl.)', shown in Figure 3.158 above, there cannot be any extrametricality.

Final CəCC, CVC and CəCː receive stress in Libyan Tripoli, as shown in (312). There is no attested CVCC which fits the expectations from the previous chapter, as where final CVCC could surface, and epenthetic vowel breaks up the cluster as in (313).

(312) Libyan Tripoli Final CVXC Stress (Yoda 2005, p. 111-112, 78)

a. ma kətʃˈb -ət -ʃ
NEG write.PFV -SBJ.3SG.F -NEG
'she did not write'

b. sakˈn -in
living -M.PL.
'living (m.pl.)'

c. nəˈfəmː -∅
be.understood.PFV -SBJ.3SG.M
'he was understood'

(313) Libyan Tripoli No Final CVCC
smanəʧ
become.fat.PFV -SBJ.3SG.M
'she became fat'

As there is no extrametricality in this variety, the stressing of final CVXC syllables cannot inform as to their structure.

3.3.5 Overview

The stress analyses motivated in this section are summarised in Table 3.1 below.

Table 3.1: Stress across Arabic Varieties

Key: MT = Moraic Trochee, L–R = Left-to-Right Parsing, R–L = Right-to-Left Parsing, ERR = End Rule Right, - = No Data Available

Variety	Foot Type	Direction	End Rule	Extrametricality
Palestinian	MT	L–R	ERR	Syllable
Algerian	MT	L–R	ERR	Syllable
Rufaidah	MT	L–R	ERR	Syllable
Lebanese	MT	R–L	ERR	Syllable
Iraqi	MT	-	ERR	Syllable
Makːan	MT	-	ERR	Syllable
Muscat	MT	-	ERR	Syllable
Qatari	MT	-	ERR	Unclear
Sˤanʕaːniː	MT	L–R	ERR	Syllable
Wadi Ramː	Iamb	L–R	ERR	Syllable
Rwaili	Iamb	L–R	ERR	Syllable
Haːyil	Iamb	-	ERR	Syllable
Cairene	MT	L–R	ERR	Consonant
Tunisian	MT	L–R	ERR	Consonant
Moroccan Casablanca	MT	R–L	ERR	None
Libyan Tripoli	MT	R–L	ERR	None

We can see in this table that all varieties explored here use End Rule Right. However, there is more variation in the other parameters. The majority have moraic trochees, with the exception of Wadi Ramː, Rwaili, and Haːyil which use iambs. These iambic varieties are also the only Bedouin varieties in this dataset. The use of iambs is common across Bedouin varieties (Watson 2011a, pp.2-3), but there is some variation with Northern Jordanian Bedouin varieties using moraic trochees (Mashaqba and Huneety 2018). The parsing directionality data is mixed, with a combination of left-to-right, right-to-left, and unclear. Moroccan Casablanca and Libyan Tripoli do share the right-to-left parsing directionality in addition to the X-slot analysis motivated in the previous chapter. However, it is unlikely that the parsing-directionality is linked to the syllable structure tolerance, given that Lebanese shares the parsing directionality as well. Furthermore, given the distance between these three varieties and the differing intervening varieties (cf. the map in Figure 1.5 it is implausible to be an areal feature.) Sˤanʕaːniː is noted as a variety that behaves differently to the rest of the varieties explored here in that there is a preference to stress CVː or CVC(left leg of a geminate) syllables. This is very unusual for Arabic, however is also found in South Arabian languages such as Mehri (Watson and Al-Mahri 2018), so may be the result of substratum influence. There does seem to be greater regularlity in terms of the extrametrical unit, and in Section 3.5, I will explore how this correlates with the syllable typology motivated in the previous chapter, and find that it correlates with TOLERANCE. However, first in Section 3.4 I will discuss the key question of this chapter: what is the shape of final CVXC?

3.4 Superheavy Syllables: The Shape of Final Syllables

In all of the varieties discussed above, all final CVXC syllables receive stress regardless of the type of extrametricality they require. The only exception to this is where final geminates are degeminated as in Rwaili and Haːyil, at which point the syllable becomes accessible to extrametricality and no longer receives stress.

The key evidence from the discussion above to identify the shape of final CVXC comes in (a) that all CVXC are stressed in the face of final syllable extrametricality and (b) that final CVCː is stressed even in cases of consonant extrametricality. Whilst we could state that CVXC syllables are merely an exception to extrametricality as Hayes (1995) implies, this is inadequate. Rather, there is a solution integrating many facts of Arabic syllable structure that I will now motivate.

That all CVXC are stressed despite final syllable extrametricality suggests that the stressable nucleus of the syllable is non-peripheral — therefore a solution where the final C is not in the same syllable as the rest is required. This rules out the recursive syllable structure of McCarthy (1979) and McCarthy and Prince (1990b), as syllable extrametricality should affect both the N^1 syllable as well as the N^0 syllable. Similarly it discounts the branching nodes of Halle and Vergnaud (1979) and the adjunction to mora of Broselow et al. (1995; 1997) and Watson (2002; 2007) as in these analyses the final C is still internal to the syllable. However, a solution where the final C is not internal to the rest of the syllable, such as Kiparsky's (2003) adjunction to word would account for this, as it would make the rest of the syllable non-peripheral.

However, adjunction to word is insufficient as the stressing of final CVCː with consonant extrametricality demonstrates. Recall that geminates are represented as a single segment associated to two higher constituents, as shown in Figure 3.161, whereas in consonant clusters and singleton consonants there is a one-to-one relationship between segment and mora as shown in Figure 3.160 and Figure 3.162. We would expect that consonant extrametricality would make the entire geminate extrametrical as it would make the last consonantal segment extrametrical, with no difference between final geminate and final singleton consonants.

Figure 3.160: Structure of Consonant Clusters

Key: σ = Syllable, O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora

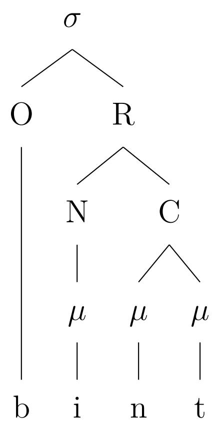


Figure 3.161: Structure of Geminates

Key: O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora

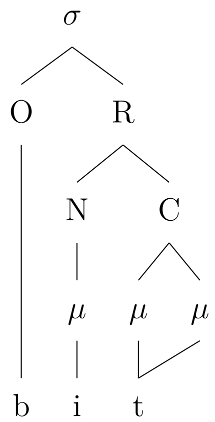
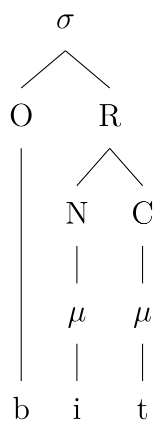


Figure 3.162: Structure of Singleton Consonants

Key: σ = Syllable, O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora



However, this is not what is observed, as shown in (314)— rather, final geminates seem somehow immune to final consonant extrametricality whereas their singleton counterparts are not.

(314) Cairene Final Geminate Stress (Davis and Ragheb 2014, p. 4)

a. ˈʔamal

‘hope’

b. ʔaˈmalː

'more/most boring'

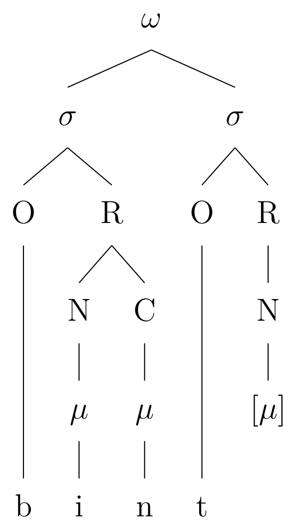
The solution, once again, is that we need this final C to be non-peripheral. This similarly rules out the CV Phonology approach of Selkirk (1981) and others, as well as Broselow's (1992) degenerate syllable analysis. To resolve this, we must make recourse to catalexis.

3.4.1 Catalexis Revisited

I propose that any material which does not conform to a bimoraic syllable template or is not tolerated medially — that is, any extrasyllabic material — is incorporated into the word as part of a syllable with a catalectic mora. As such, final CVXC cannot be made extrametrical as its structure is underlyingly CVX.C[μ], as shown in Figure 3.163.

Figure 3.163: Catalectic Mora Structure

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora



Thus the CVX is non-peripheral for syllable extrametricality and can be stressed; and similarly a final geminate is also non-peripheral as the catalectic mora fills the peripheral position. Note that while Kiparsky (1991) is noncommital about the ordering of extrametricality and catalexis, here the case of consonant extrametricality not affecting final CVXC is crucial in demonstrating that catalexis must occur first.

This solution does not account only for the stressing of final CVXC but can also account for the relationship between epenthetic position and word-edge clusters. By word-edge clusters I do not only refer to those at the right edge that are accessible to extrametricality, but those at *both* the left and the right edge. I propose that the extrasyllabic consonant can be assigned to either the onset or coda of the syllable containing a catalectic mora. Epenthesis is then not a matter of resyllabification but of associating segmental material to this catalectic mora. I propose that varieties differ in the extent to which they permit word internal catalectic moras without overt segmental material. In word medial syllables, catalectic mora cannot surface as they are non-peripheral, so these are always filled with segmental material through epenthesis. In word edge syllables, it depends on whether the catalectic mora is word internal or word peripheral. In (315) and (316), we see the position of the catalectic mora in relation to the final consonant. In Onset varieties, the catalectic mora is at the edge of the word, so is permitted to not contain overt segmental material; however in Coda varieties, it is word-internal, so if a variety does not permit this it is filled with an epenthetic vowel as overt segmental material. Note that contrary to previous work that has focused on the right-edge catalectic moras, in (315) I posit the existence of catalectic moras at the left edge as well. While it is very marked for extrametricality to affect the left edge, in Arabic there is no reason to assume that such an asymmetry exists for catalexis as well.

In the following diagrams, square brackets denote the catalectic mora, and the epenthetic vowel inserted into the catalectic mora slot continues to be marked with an underline.

(315) Initial Clusters

a. Onset Variety: *f$[\mu]$red $\rightarrow$ f$\underline{\text{i}}$red

b. Coda Variety: $[\mu]$fred

Figure 3.164: Initial Cluster Onset Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora

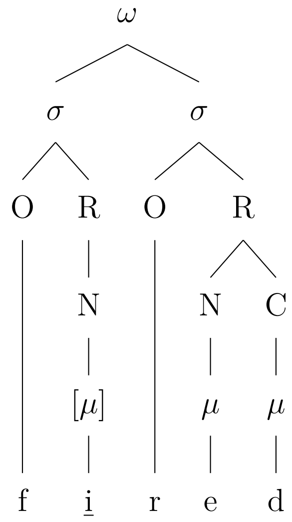
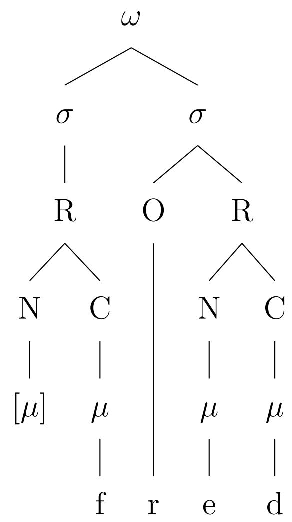


Figure 3.165: Initial Cluster Coda Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora



(316) Final Clusters

a. Onset Variety: katabt[μ]

b. Coda Variety: *katab[μ]t $\rightarrow$ katabi̱t

Figure 3.166: Final Cluster Onset Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora

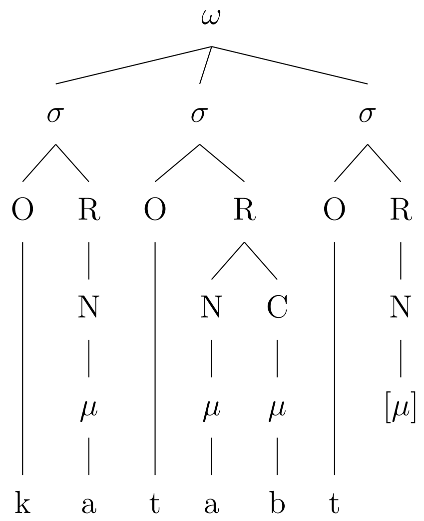
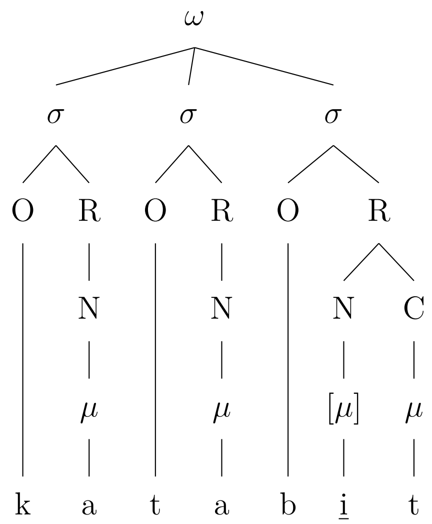


Figure 3.167: Final Cluster Coda Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora



This, therefore, accounts for the relationship between epenthesis direction and domain edge clusters. Furthermore, the distinction made in the previous chapter about 'Extreme' Coda and 'Extreme' Onset varieties can be formalised as the 'Extreme' ones permit a catalectic mora word internally in a peripheral syllable, whereas the non-extreme varieties do not. In the middle of the word, the catalexic mora must be filled by overt material as catalexic moras are only permitted in the periphery (Kiparsky 1991; Kager 1995a).

Crucially, while catalectic moras can be used to repair feet (and here, syllables), a catalectic mora cannot constitute the head of a foot. Therefore, the variation in whether epenthetic vowels are stressed is not a problem. In varieties such as Palestinian epenthetic vowels are not stressed, for example in [kaˈtabi̱lna], 'he wrote to us'. Under cyclical models such as that of Brame (1974), this is analysed as epenthesis occurring after stress assignment. In this analysis, stress is assigned in varieties such as Palestinian before the catalectic mora is associated with segmentally overt material. While segmentally empty cannot form the head of a foot. However, in varieties where it can be stressed such as Cairene, the catalectic mora is filled with overt segmental material before stress assignment, so can function as the head of a foot. Such a mechanism can also be seen in Tamil prothetic glides and schwa-insertion after monosyllabic stems in Dutch diminuitives (Kiparsky 1991).

The proposal here does not suggest that catalexis is always 'on' for Arabic; rather, it occurs when needed to repair degenerate feet and syllables. It cannot be always 'on' otherwise it would incorrectly predict the surface stress of [ˈʔalːamat] in Palestinian. As shown in Figure 3.168, a final catalectic mora would resyllabify the final /t/ as its onset, leaving a trochee and as such predicting antepenultimate stress rather than the attested initial stress. However, without a catalectic mora, syllable extrametricality gives the attested initial stress in 3.169.

Figure 3.168: Catalectic Mora with [ʔalːamat] Incorrectly Predicts Stress

	X		
(x)	(x	.)	<>
ʔal	la	ma	t[μ]

Figure 3.169: Correct Stress Analaysis of [ʔalːamat] Without Catalectic Mora

X		
(x)		<>
ʔal	la	mat

In some ways, this proposal is not dissimilar to that of Broselow (1992) who argued that the relationship between extrametricality and edge clusters could be accounted for by assigning extrasyllabic material to the onset or coda of a degenerate syllable. However, there are several advantages to my proposal. Firstly, it accounts for the exemption of final geminates from consonant extrametricality by making them non-peripheral; secondly, it does not require the creation of a new and vaguely defined 'degenerate syllable'; and thirdly, it better accounts for subminimal words — a key diagnostic for the presence of catalexis.

Particularly in the varieties of North Africa, there are words with initial geminates of the structure CːəC as well as singleton clusters of the structure CCəC. Given that the schwa does not contribute a mora, these words appear subminimal. Examples of such words are given in (317)–(319).

(317) Subminimal Examples in Moroccan Casablanca (Boudlal 2001, p. 49, 82)

a. qfəz
'cage'

b. nməl
'ants'

c. sːma < l-sma
DEF sky
'the sky'

d. dːa -∅
take.PFV -SBJ.3SG.M
'he took'

(318) Subminimal Examples in Algerian (Bouhadiba 1988, p. 204, 323)

a. mːa
'mother'

b. dːa -∅
take.PFV -SBJ.3SG.M
'he took'

c. mra
'wife'

d. mrːa

'itch'

(319) Subminimal Examples in Libyan Tripoli (Yoda 2005, p. 159-160, 238)

a. ktʧəb -∅
write.PFV -SBJ.3SG.M
'he wrote'

b. fːətˤ

'towels'

c. gdəb -∅
lie.PFV -SBJ.3SG.M
'he lied'

d. mrəd -∅
crawl.PFV -SBJ.3SG.M
'he crawled'

e. xləq -∅
create.PFV -SBJ.3SG.M
'he created'

f. wqəf -∅
stop.PFV -SBJ.3SG.M
'he stopped'

If we assume that the initial clusters form a complex onset, as in the diagram in Figure 3.170, these words appeal monomoraic since only the final consonant carries mora (given the amoraicity of the schwa). However, if we assume the catalectic mora representation as in Figure 3.171, this is no longer a problem.

Figure 3.170: Initial Complex Cluster Representation of Subminimal Words

Key: σ = Syllable, O = Onset, R = Rime,
N = Nucleus, C = Coda, μ = Mora

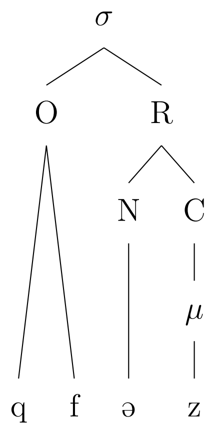
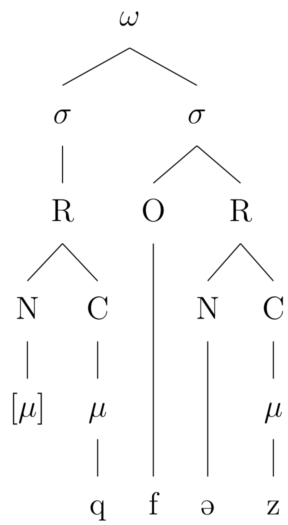


Figure 3.171: Catalexis Analysis of Apparently Subminimal Words

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, $[\mu]$ = Catalectic Mora



That an initial cluster contributes a mora in some way is even clearer in the following alternations from Libyan Tripoli, where (320-a) and (320-b) alternate, as do (320-c) and (320-d), between final geminate and initial geminate.

(320) Minimal/Subminimal Alternations (Yoda 2005, p. 163)

a. fəmː -∅
understand.PFV -SBJ.3SG.M
'he understood'

b. fːəm -∅
understand.PFV -SBJ.3SG.M
'he understood'

c. dənː -∅
anoint.PFV -SBJ.3SG.M
'he anointed'

d. dːən -∅
anoint.PFV -SBJ.3SG.M
'he anointed'

However, it is not just in North African varieties that apparently subminimal words are seen. Wadi Hadˤrami, a Yemeni variety not discussed here, has [bɣa], 'he wanted'.

It is on the basis of words such as those in (320) that Davis (2011) and others have argued for the initial geminate carrying a mora. However, as is clear above this phenomenon is also seen with non-geminate clusters so this is not merely a matter of whether initial geminates carry a mora. Rather, all the initial clusters appear to be moraic. If we wish to maintain the amoraicity of the onset position in syllable representations, then word initial catalexis can solve this problem. The advantage of this proposal is it suggests that these words are underlyingly [μ]CCəC as shown in Figure 3.171, and as such fit the bimoraic minimum established elsewhere in Arabic.

Evidence of subminimal words is harder to find in the Eastern varieties of Arabic, particularly those with syllable extrametricality. Even in Cairene with consonant extrametricality, the subminimal Classical Arabic words /ʔab/ 'father' and /ʔax/ 'brother' surface with gemination in isolation in Cairene: [ʔabː] and [ʔaxː] (Watson 2002, pp. 88–9). This also occurs in Maʕani Arabic, a variety spoken in Jordan (Rakhieh 2009, p. 95). But is this a problem? If catalexis is predominately a matter of repairing otherwise degenerate feet, then yes, the absence of degenerate feet does imply the absence of catalexis. However, in this proposal catalexis is repairing not degenerate feet but degenerate syllables, through the inclusion of a catalectic mora to syllabify otherwise extrasyllabic material. Therefore, the absence of degenerate feet should not be a problem — rather, we require evidence of degenerate syllables or otherwise extrasyllabic material, for which there is abundant evidence in all varieties of Arabic through the behaviour of final CVXC as shown above.

In this way, catalexis is viewed not just as a permanent 'on/off' switch, but as a repair mechanism for degenerate syllables and feet. Where it repairs degenerate feet, a key diagnostic should be the presence of degenerate feet, as argued by Kiparsky (1991) and Kager (1995a). However, where it repairs degenerate syllables, as is the case with final CVXC here, the key diagnostic should instead be the presence of degenerate syllables.

3.4.2 Catalectic X-slot

In Kiparsky (1991) and subsequent work on catalexis, it has been assumed that there are only catalectic syllables and catalectic moras, and the presence of one does not entail the presence of the other. However, what has not been considered is whether a catalectic X-slot is possible.

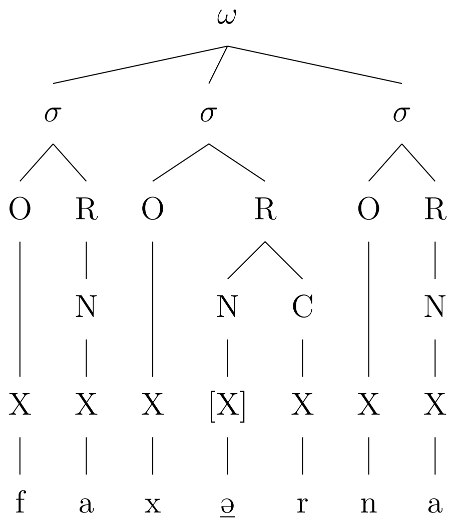
In the previous chapter I argued that Moroccan Casablanca and Libyan Tripoli require a TOLERANCE analysis in terms of X-slots rather than segments or moras. This is because there is an asymmetry between length and weight in these varieties. Both contain short vowels and schwa, where the short vowel carries one mora and the schwa none, but to account for the medial syllable structure restrictions the short vowel must be associated to two timing slots and the schwa one. In similar cases of asymmetry, including X-slots in the representation has been required, such as in Kraehenmann (2003).

Some setting for REPAIR is required for these varieties. However, is it necessary to mark this element as a catalectic *mora*, given that moras are not used for the syllable structure in these varieties otherwise? It is difficult to exclude the mora entirely from these dialects as it is useful in the construction of feet, but outside of stress it has no role in the phonology of either dialect.

Given that catalexis is a matter of a structural slot unfilled by overt segmental material, there seems to be no a priori reason why these varieties cannot be repaired using a catalectic X-slot as shown in Figure 3.172. The added advantage of using the catalectic X-slot here is it facilitates the use of an epenthetic schwa, which does not carry any mora in these varieties.

Figure 3.172: Catalectic Mora Repair of [faxr-na], 'our coal'

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, X = X-slot, [X] = Catalectic X-slot



3.5 Correlation with Syllable Typology

In Section 3.3.5 I concluded that while there is some variation in foot type and parsing directionality, the most consistent axis of variation in stress was the unit of extrametricality. This unit of extrametricality is correlated not with Repair but with Tolerance, as summarised in Table 3.2 below.

Recall that Tolerance posits that varieties differ at what level of syllable structure they impose restrictions on syllable size. Segmental varieties impose restrictions on the segmental level, where CVC can be associated to multiple moras, permitting surface long segments with CVːCː at the extreme. Moraic varieties impose restrictions on the moraic level, only permitting a particular number of moras thus not allowing CVːC or CVCː. X-slot varieties have an asymmetry between length and weight, and as such have restrictions on the number of X-slots permitted in a syllable. Since short vowels are affiliated to two X-slots, this permits CVCC and CVCː and accounts for the lack of CVːC syllables. The distribution of extrametrical unit, Repair and Tolerance parameters across these sixteen varieties is summarised in Table 3.2 below.

Table 3.2: Covariance between Stress and TOLERANCE across Arabic Varieties

Variety	Extrametricality	REPAIR	TOLERANCE
Palestinian	Syllable	Coda	Segmental
Lebanese	Syllable	Coda	Segmental
Muscat	Syllable	Coda	Segmental
Algerian	Syllable	Coda	Segmental
Iraqi	Syllable	Coda	Segmental
Wadi Ramː	Syllable	Coda	Segmental
Sˤanʕaːniː	Syllable	Onset	Segmental
Rufaidah	Syllable	Onset	Segmental
Rwaili	Syllable	Onset	Segmental
Haːyil	Syllable	Onset	Segmental
Makːan	Syllable	Onset	Moraic
Qatari	Unclear	Coda	Moraic
Cairene	Consonant	Onset	Moraic
Tunisian	Consonant	Coda	Moraic
Moroccan Casablanca	None	Coda	X-slot
Libyan Tripoli	None	Coda	X-slot

Here we see that syllable extrametricality is associated with segmental restrictions on syllable structure, consonant extrametricality is associated with moraic restrictions on syllable structure, and no extrametricality is associated with X-slots. Whether the variety has ONSET or CODA REPAIR strategies varies across these groups.

The sole exception in this data set is Makːan, which has syllable extrametricality with moraic, not the expected segmental, restrictions. However, remember from Section 2.6.2.8 that while Makːan has moraic restrictions it also permits branching nuclei in syllables created by syncope. Therefore, it is plausible that Makːan is transitioning from moraic to segmental restrictions in line with the neighbouring varieties (Rufaidah, Rwaili, Haːyil), and as part of this process the extrametricality has moved first. Note that the other variety that permits branching nuclei is Qatari, which as discussed above lacks unambiguous data as to whether it requires syllable or consonant extrametricality, so is plausibly undergoing a similar process of change to Makːan, and is also a peninsular variety of Arabic.

There are some geographic patterns across this distribution. Neighbouring varieties Palestinian, Lebanese, Wadi Ramː, Iraqi share both coda repair and segmental tolerance patterns, with the onset repair/segmental tolerance set of Sˤanʕaːniː, Rufaidah, Rwaili, and Haːyil also close geographically. The distributions of moraic and x-slot tolerance restrictions are less geographically clear, and this is in part due to the wide variation found in North Africa. While both Moroccan Casablanca and Libyan Tripoli are North African varieties, they are at opposite ends of the region with segmental and moraic varieties in between. The variation in North Africa is greater than in the Levantine and Peninsular varieties, and likely reflects waves of invasion. Investigation of a larger set of North African varieties with close attention paid to the historical origins of the varieties is needed to clarify the origins of this distribution.

3.6 Furthering the Syllable Structure Typology

Previous approaches to the phonological typology of Arabic focussed on the position and behaviour of epenthesis. I propose a more complex typology that also takes into account the tolerance of medial CVXC syllables and the distribution of extrametricality. This approach not only has broader empirical coverage, but also provides mechanisms to account for the behaviour of non-canonical varieties.

In order to account for the limited correlation between domain edge clusters and medial CVXC behaviours, I propose a bifurcated typology, where varieties have values for both Tolerance and Repair. Varieties differ in their Tolerance of medial CVXC syllables at either the segmental, moraic, or X-slot tiers, as summarised in Table 3.3. This accounts for the patterning of medial CVCː with medial CVːC syllables. These Tolerance types are correlated with extrametricality. Syllable extrametricality is associated with segmental restrictions on syllable structure, consonant extrametricality is associated with moraic restrictions on syllable structure, and no extrametricality is associated with X-slots.

Table 3.3: Overview of TOLERANCE

Tier of Syllable Restriction	Moraic	Segmental	X-Slot
No. of Elements		2-3	
Extra Elements	variety specific including coronality, sonority and speech tempo		
Extrametricality	Consonant	Syllable	None

Varieties also differ in REPAIR of illicit syllables in terms of epenthesis direction (CODA or ONSET) and the licensing of word edge extrasyllabic segments.

Table 3.4: Overview of REPAIR

	ONSET	Extreme ONSET	CODA	Extreme CODA
Epenthesis Position	CV		VC	
Syncope Type	Non-Promiscuous		Promiscuous	
Initial #CC-	NO	YES	YES	YES
Final -CC#	YES	YES	NO	YES

The behaviour of REPAIR is a matter of where catalectic moras or X-slots are permitted. I propose that catalectic moras are required to incorporate extrasyllabic material across the word. The ONSET/CODA distinction can therefore be expressed as whether the extrasyllabic material is incorporated as the onset or coda of a syllable containing a catalectic mora. Note this can occur at the right or left edge of the word. Epenthesis occurs where the catalectic mora is not permitted to surface without overt segmental material. The difference between domain edge behaviour therefore follows from this — where the catalectic mora in the edge syllable is the peripheral element, the surface cluster is permitted; but where it is non-peripheral, epenthesis occurs. 'Extreme' ONSET/CODA varieties are those that permit catalectic moras word-internally. I propose that varieties with X-slot TOLERANCE restrictions also use X-slots in REPAIR rather than requiring further machinery to capture their distribution.

Both TOLERANCE and REPAIR are needed to account for data found across Arabic varieties. As shown in the Table 3.5 below, they do not covary. Note that Wadi Ramː, Rufaidah, Rwaili, Haːyil, Moroccan Casablanca, Libyan Tripoli, Tunisian, and Algerian are marked as 'extreme' CODA or ONSET varieties. This is because they permit clusters at both edges, that is, they permit word-internal catalectic moras in peripheral syllables.

Table 3.5: TOLERANCE and REPAIR of Arabic Varieties

Variety	Extrametricality	REPAIR	TOLERANCE
Palestinian	Syllable	Coda	Segmental
Lebanese	Syllable	Coda	Segmental
Muscat	Syllable	Coda	Segmental
Algerian	Syllable	Coda	Segmental
Iraqi	Syllable	Coda	Segmental
Wadi Ramː	Syllable	Coda	Segmental
Sˤanʕaːniː	Syllable	Onset	Segmental
Rufaidah	Syllable	Onset	Segmental
Rwaili	Syllable	Onset	Segmental
Haːyil	Syllable	Onset	Segmental
Makːan	Syllable	Onset	Moraic
Qatari	Unclear	Coda	Moraic
Cairene	Consonant	Onset	Moraic
Tunisian	Consonant	Coda	Moraic
Moroccan Casablanca	None	Coda	X-slot
Libyan Tripoli	None	Coda	X-slot

Note that for X-slot varieties (here, Moroccan Casablanca and Libyan Tripoli), a catalectic *X-slot* rather than mora is used during REPAIR as there is no need to posit the mora as extra level of representation for syllable structure here, as argued above in Section 3.4.2.

This analysis has consequences outside of Arabic, in particular through revisiting catalexis. Here I construe catalexis not merely as a tool to ensure a universal bimoraic minimal word, but rather as a tool for repair of both degenerate feet and degenerate syllables.

3.7 Next Steps

The work so far has focused purely on the phonological behaviour of Arabic varieties. In the next chapter I will consider the phonological/morphology interface.

Specifically, I will consider the linear concatenation of formatives, and whether these are affixes or clitics. This is an important question, as it allows us to interrogate the prosodic word structure of Arabic. I demonstrate that they are indeed affixes, and that all the suffixes in question have a preference for attaching to a stressed or stressable syllable, and that this preference accounts for apparently idiosyncratic lengthening behaviours as well as the stress exceptions found across Arabic varieties.

4
The Classification of Clitics and Affixes

Contents

4.1 Introduction

In many languages, function words including pronouns, auxiliaries, and prepositions can be reduced under certain conditions and behave as phonological clitics (see Lahiri and Plank (2010) for Germanic, Lahiri and Fitzpatrick-Cole (1999) for Bengali, Zwicky (1977) for Spanish and Turkish, Nespor and Vogel (1986) for Greek defined as clitic groups, and Ngonyani (2006) for Bantu).

The difference between clitic and affix status for similar elements can be controversial, and there can be a mismatch between the syntactic and phonological status of particular elements. Indeed, for Swedish, the syntax literature would like to label the affix-like determiners as suffixes while phonologically they are clitics (Lahiri, Wetterlin, et al. 2005). In Norwegian compounds, [-e] is a suffix as tonal assignment occurs after suffixation and compounding, whereas [=s] is a clitic because tonal assignment occurs before cliticisation and compounding. Thus the set in (321) is found. The bare singular in (321-a) has accent 1 as it is monosyllabic, and the compounds in (321-b) and (321-c), regardless of whether the linking [-e] is present, have accent 2 as they are disyllabic. However, (321-d) retains the accent 1 of the bare singular as the clitic [=s] and subsequent material are attached after tonal assignment.

(321) Norwegian Compounds (Lahiri, Wetterlin, et al. 2005, p. 2)

a. land_1
   'land'
b. land-mann_2
   'farmer'
c. land-emerke_2
   'landmark'
d. land=smann_1
   'compatriot'

While the syntactic literature agrees that [=s] is a clitic, it views the definite marker as a suffix (Börjars 1994). However, Lahiri, Wetterlin, et al. (2005) demonstrate these are clitics as they display the same attachment behaviours that do not affect tone as the [=s] marker, as shown in (322) where all forms retain the accent 1 of the monosyllabic singular.

(322) Definite Markers (Lahiri, Wetterlin, et al. 2005, p. 8)

a. hus_1
   'house'
b. hus=et_1
   'the house'

c. hus=en$_1$
'the houses'

Semitic languages are full of elements which have caused diverse views on whether they are clitics or affixes. While these elements do include function words such as prepositions, this chapter will focus on the morphological elements attached to some degree to a host that have resisted clear diagnosis as clitic or affix. While 'phonological clitic' is a loose term that covers a wide range of behaviours and functions, phonological cliticisation is a much more defined process. Therefore, this chapter is concerned with function words that are not necessarily syntactic clitics but may or may not form feet on their own. Understanding the process of cliticisation or affixation that they undergo is vital for our understanding of Arabic prosodic word structure, and the implications of this structure for phonological process that occur across modern Arabic varieties.

The templatic, non-linear concatenation of formatives of Arabic has been extensively discussed in the literature, but this only refers to the shape of word stems. Linear formative concatenation still occurs to mark the subject, pronominal objects, the dative with indirect pronominal object, negation, as well as tense/aspect markers and the possessive.

In the Cairene example in (323), the direct object, dative, indirect object and negation have exponents linearly concatenated to the verb. Note here and throughout epenthetic vowels are underlined.

(323) ma katab -∅ -hum -l<u>u</u> -kum -ʃ
NEG write.PFV -SBJ.3SG.M -OBJ.3PL.M -DAT -OBJ.2PL.M -NEG
'He didn't write them for you' (Cairene, (Broselow 1976, p. 186))

In the Sˤanʕaːniː example in (324), the future tense is indicated by a proclitic (ʔa-) attached to the imperfective stem. However in Muscat in example (325) a different future proclitic is used (ba-).

(324) maː ʔa -t -ɟirːa -haː -l -iː -ʃ
NEG FUT -SBJ.2SG.M -take.IMPF -OBJ.3SG.F -DAT -OBJ.1SG -NEG
'You (m.s.) will not take her for me' (Sˤanʕaːniː, (Watson 2002, p. 177)

(325) ba -ji -tʕalːam -uː -ha
FUT -SBJ.3SG.M-learn.IMPF -SBJ.PL -OBJ.3SG.F
'they will learn it' (Muscat, (Glover 1988, p. 88))

This linear concatenation is not limited to verbs, but also occurs with nouns. In the Muscat example in (326), the possessive [-ak] is attached to the noun.

(326) qandiːl -ak
lantern -POSS.2SG.M
'your (m.s.) lantern' (Muscat, (Glover 1988, p. 60))

The formatives of interest in this chapter are the subject marker, the pronominal non-subject formative, the dative /-l-/, and negation /-ʃ/[1]. The status of these linearly concatenated formatives is in question. To date, a distinction has been drawn between the subject markers and the others, and this has been represented as the subject is an affix whereas the rest are clitics[2].

This distinction is drawn on the basis that the subject marker is obligatory regardless of the presence of a full NP subject, whereas the object marker only appears on the verb when it is a pronominal object (Brame 1973; Kenstowicz and Kisseberth 1979), as illustrated in the following examples. In (327-a), (327-b) and (327-c) the subject marker /-it/ surfaces in conjunction with the full NP subject 'Ghadeer', and in (327-d) and (327-e) the subject marker surfaces without the presence of the full NP subject. However, the object marker /-ha/ only surfaces in (327-b) and (327-e) where the full NP object 'Leila' is not present; for both to appear is ungrammatical, as in shown in (327-c).

(327) Pronominal Subject and Object Distribution

a. ɣadeːr ʃaːf -it Leila
Ghadeer see.PFV -SBJ.3SG.F Leila
'Ghadeer saw Leila'

[1]In some varieties negation has cumulative exponence with both final /ʃ/ and a proclitic /ma/; in others only one or the other is used.

[2]The Arabic syntax literature debates whether the subject formatives are incorporated pronouns (Fassi Fehri 1990; Fassi Fehri 1993; Jamary 1993; Ahmed 2015) or agreement markers (Mohammad 1990; Aoun et al. 2010; Benmamoun 2011).

b. ɣadeːr ʃaːf -it -ha
Ghadeer see.PFV -SBJ.3SG.F -OBJ.3SG.F
'Ghadeer saw her'

c. *ɣadeːr ʃaːf -it -ha Leila
Ghadeer see.PFV -SBJ.3SG.F -OBJ.3SG.F Leila
*'Ghadeer saw her Leila'

d. ʃaːf -it Leila
see.PFV -SBJ.3SG.F Leila
'She saw Leila'

e. ʃaːf -it -ha
see.PFV -SBJ.3SG.F -OBJ.3SG.F
'She saw her'

Furthermore, the form of the subject markers differs between past and present tense verbs, whereas the form of the other formatives does not change. In the following example, the stems and exponent for third person plural subject differ between the past tense suffix in (328-a) where it is [uː] and the present tense circumfix in (328-b) where it is [yu...uː][3], whereas neither object ([haː] and [kum]) nor dative ([l])[4] change. The exponents for subject have been emboldened here for ease of reading.

(328) a. katab **-uː** -haː -lu̲ -kum
write.PFV -SBJ.3PL.M -OBJ.3PL.M -DAT -OBJ.2PL.M
'They wrote it for you'

b. **yu** -ktub **-uː** -haː -lu̲ -kum
3SG -write.IMPF -SBJ.PL.M -OBJ.3PL.M -DAT -OBJ.2PL.M
'They are writing it for you'

That there is a distinction between these formatives and the subject marker is not disputed here; but rather the nature of the distinction. For the sixteen dialects explored here, I argue that these endings are in fact *suffixes* not enclitics and that the differences between them are a matter of affixation at Level 1 rather than Level 2, though I recognise that a prefix/clitic difference still remains synchronically.

[3]The prefix is the exponent of person and the suffix is the exponent of number.
[4]The vowel after the dative /l/ is epenthetic.

4.1.1 Structure of the Chapter

In Section 4.2, I will review the literature on how to diagnose clitic versus affix status and their typical behaviours. In Section 4.3, I will explore the properties of these endings to demonstrate that they are in fact affixes. In Section 4.4, I will investigate the phenomenon known as 'pre-clitic' lengthening and show that it is motivated by a preference to attach after a stressed or stressable foot, and that rather than a clitic phenomenon this is also seen in subject affixes, particularly in weak and biliteral verbs. In Section 4.5, I will demonstrate that apparent exceptions to stress parameters mentioned in Chapter 3 are not exceptional at all, but come from this same motivation to attach affixes after a stressed or stressable foot. In Section 4.6, I will demonstrate how they are incorporated into the prosodic word at Level 2 unlike the subject marker at Level 1. Finally, I will show that a clitic/affix distinction is still needed for Arabic in Section 4.7, before summarising my findings in the conclusion in Section 4.8 and showing the extent to which these correlate with the syllable structure restrictions discussed in Chapters 2 and 3.

4.2 Literature Review

Classifying different utterances to allow for a unified analysis of their behaviour is important across linguistics – for theoretical understandings of morphological structure, for analyses of language change, and for theories of language acquisition, among others. However, there are many examples of phenomena that do not fit neatly into theoretically clear categories. The concept 'clitic' attempts to cover a set of forms that do not neatly fit into 'word' or 'affix', though there is significant debate across the literature as to which utterances constitute clitics, what phonological, morphological, and syntactic behaviours are characteristic of this group, and indeed whether clitics exist at all.

The approach to this phenomenon in the generative grammar of the twentieth century began with a focus on function words. These 'little words' are independent words in terms of their syntax, but are phonologically dependent on a host. Two main issues emerged: what is the domain of affixation versus cliticisation, and is there a mismatch between phonological and syntactic phrasing?

Initial work considered how to account for the mismatch in phonological and syntactic phrasing. This mismatch can be seen in the following example from Kwak'wala[5] where the definite article /-da/ forms a syntactic constituent with the following word, but is phonologically part of the preceding word (Black 2011, p. 2):

(329) tɛnxalˈoɣda tsɪˈdaq

denxala -ox -da tsedak
sing -2.LOC -DEF woman

'The woman is singing'

Selkirk (1972; 1984; 1986; 1995), Nespor and Vogel (1986), and later scholars argue that the directionality is determined by the syntax, that is, that phonological and syntactic phrases should line up. As such, Selkirk proposes that there are only three ways function words can attach to prosodic words, as shown in (330).

(330) Attachment of Function Words to Phonological Words (Selkirk 1995)

a. $(\text{fnc})(\text{Lex})_{\text{Pwd}})_{\text{PPh}}$

   Free cliticc

   to Toronto

b. $((\text{Lex})_{\text{Pwd}}(\text{fnc})_{\text{Pwd}})_{\text{PPh}}$

   Prosodic word

   look at

c. $(((\text{verb})_{\text{Pwd}}\ \text{pro})_{\text{Pwd}})_{\text{PPh}}$

   Affixal clitic

   need him

This mapping matches the syntactic phrasing. In (330-a), the /t/ in *to* is not aspirated, whereas in *Toronto* it is. Selkirk argues that this is because *to* is phrase-initial rather than a phonological word onset. However, this is contested by those

[5] A Northern Wakashan language spoken in the Queen Charlotte Strait Canada, previously known as Kwakiutl

who view it as a matter of rhythm, with Lahiri and Plank (2010) arguing for trochaic phrasing and a universal preference for encliticisation. Lahiri and Plank (to appear) argue that *to* is encliticised to a preceding word, so *Toronto* is the start of the phonological word and aspiration then occurs as predicted. This is supported by experimental evidence in Wheeldon and Lahiri (1997). They studied the production of sets such as that in (331), where I have marked the phrases they predict.

(331) Wheeldon and Lahiri (1997) Data Set

a. ik drink het | water

b. ik drink | water

c. ik drink | vers | water

When subjects are given time to plan the production of these sentences, the number of prosodic units matters, so (331-a) and (331-b) take the same amount of time, whereas (332-c) takes longer. When subjects are not given planning time, the length of the first unit matters, so they find that (331-a) takes longer. This suggests leftwards encliticisation of the determiner [het], because if it was rightwards there should be no difference in the initiation time.

The domain of affixation versus cliticisation has also received much attention, as it predicts what phonological phenomena occur. Wackernagel (1892) focused on unstressed pronouns and particles, and according to Spencer and Luis (2012, p. 84) stress neutrality is the most frequent prosodic property found, but this is not an adequate diagnostic —- especially in light of languages such as Japanese that lack a stress or prominence system but do seem to have clitics (Zwicky 1985). Furthermore, there are clitics that do interact with the stress system, such as the Latin /-que/ that attracts stress to the preceding syllable, or Modern Greek enclitics that receive stress if there is sufficient material to build a trochaic foot at the right edge of the word without a stress clash as shown in example (332) below.

(332) Modern Greek Stressable Enclitics

a. triaˈðafilo
'rose'

b. triaˌðafiˈlo=mu
'my rose'

c. ˈðose=mu
'give me!'

d. ˌðose=ˈmu=to
'give it to me!'

Rather, Anderson (2005) proposes they should be considered 'prosodically deficient' in that they lack prosodic structure at the level of the prosodic word or below, so must be incorporated into the prosodic structure of another word. The same can be said of affixes, so some method to distinguish between clitics and affixes is necessary.

Clitics have resisted clear diagnostic properties. This is, in part, because they often emerge due to a process of grammaticalisation, so vary in where they are on this grammaticalisation cline. Clitics can exhibit a broad range of functions, including tense and aspect, case, possession, definiteness, agreement, and evidentiality (Spencer and Luis 2012, p. 37). Furthermore, there is no particular correspondence between agreement markers and affixes on the one hand, and arguments and words or clitics on the other (Spencer and Luis 2012, p. 24). Thus, clitics and affixes must be distinguished based on morphosyntactic and phonological behaviours rather than particular functions.

Zwicky and Pullum (1983) suggested that there were tendencies in the diagnosis of clitics and affixes that could be used to identify them. In exploring the behaviour of the English *n't*, which they proposed was a modal inflectional affix compared to clitic reduced auxiliary verbs (*'ll, 's*), they proposed the following set of criteria.

Host Selectivity: Clitics can select a host from a wide range of lexical categories; affixes have much higher selectivity, and can attach to stems from one or a few lexical categories

Arbitrary Gaps: Affixes have more arbitrary gaps than clitics

Morphophonological Idiosyncrasies Affixed words have more idiosyncrasies, such as the irregular plurals *dice, oxen* and *feet*; however clitics tend to be very regular

Semantic Idiosyncrasies: Affixed words have more semantic idiosyncrasies than clitics

Lexical Integrity: Syntactic rules can affect words with affixes rather than clitic groups

Ordering: Clitics can attach to material already containing clitics, but affixes cannot follow clitics

Whilst Zwicky and Pullum (1983) portrayed these as descriptive observations, for Anderson (2011) they are categorical diagnostics that logically follow from clitics being prosodically deficient elements not listed in the lexicon. However, there are elements across languages that do not seem to be definitively clitics or affixes (see Spencer and Luis (2012) Chapter 7). Diagnosing these can be challenging. Auger and Janda (1994) argues that the criteria of host selectivity and lexical integrity would lead to French weak pronouns (such as those in the example below) being identified as affixes; but the other criteria fit with analysis as clitics.

(333) Je le lis
SBJ.1SG OBJ.3SG.M read.PST
'I read it'

From a canonical typology perspective, it does not matter that these criteria cannot account for the behaviour of French weak pronouns. Rather, the aim is to define a canonical space of theoretical possibilities and to observe how this space is populated by real-world examples. Therefore, these criteria are a useful starting point for diagnosing the formatives in question in this chapter.

4.2.1 Conclusion

It is not necessarily an easy matter to identify whether a particular formative is a clitic or an affix, and certainly many researchers believe that a principled division is impossible (Börjars 1998; Embick and Noyer 2001; Luis 2004). However, for Arabic we do have a set of formatives that are regularly identified as clitics in the literature, so engagement with this question is necessary. These formatives have a deep impact on the surface phonology of the varieties in question, so understanding how they are incorporated into the prosodic word is vital for our understanding of Arabic prosodic and metrical structure.

Such a distinction is only necessary if there is a clear need to distinguish different phonological behaviours. In this chapter, I will begin with the Zwicky and Pullum criteria, before looking in more depth at the phonological behaviour of these formatives. It is on the basis of morphophonological idiosyncracies, host selectivity and ordering that the endings in question here are often identified as clitics in the literature — though in the next section I will dispute these conclusions. Note the focus here is on the domains of affixation and cliticisation rather than possible mismatches between phonological and syntactic phrasing.

4.3 Diagnosing the Clitic versus Affix Distinction

4.3.1 Host Selectivity and Ordering

The canonical affix displays high host selectivity in that it only affixes to hosts from a restricted set of lexical categories; whereas the canonical clitic displays low host selectivity in that it can select a host from a broad range of lexical categories.

The Arab grammarians and most subsequent western scholars identify two sets of pronominal forms: a set of free formatives for the grammatical subject and a set of bound endings for non-subject roles (Holes 2004, p. 177).

The paradigm for this non-nominative formative in Table 4.1 is representative of most varieties, where V stands for a vowel, which can vary depending on and even within the variety:

Table 4.1: Non-Nominative Formative Paradigm

SG	1	i ~ni
	2	(V)k
	3m	V(h)
	3f	ha
PL	1	na
	2	kum
	3	hum

The allomorphy for the first person singular ending is distributed such that [i] is used for the possessive and [ni] is used elsewhere. This formative attaches to verbs, prepositions, nouns, modals, auxiliaries and complementisers, as shown in (334) from Cairene where (334-a) shows attachment to verb, (334-b) attachment to preposition and (334-c) attachment to complementiser (Broselow 1976, p. 81).

(334) Attachment to Different Lexical Classes (Broselow 1976, p. 81)

a. dˤarab -∅ -u
hit.PFV -SBJ.3SG.M -OBJ.3SG.M
'he beat him'

b. il -kitaːb taħt -u
DET -book under -OBJ.3SG.M
'the book is under it'

c. ʔaːl -it inː -u ʕajːaːn
say.PFV -SBJ.3SG.F COMP -OBJ.3SG.M sick.M
'she said that he is sick'

It also attaches to participles in Lebanese as shown in (335), and in Muscat as shown in (336) with the infixation of /-in-/ between the participle and the object, where /-in-/ is glossed as ∅ as it does not contribute any meaning.

(335) Attachment to Participles (Haddad 1984, p. 165)

a. kaːtib -l -i
write.PTCP.M -DAT -OBJ.1SG
'his writing to me' (Lebanese)

(336) Attachment to Participles (Glover 1988, p. 212)

(336) mraːbʕ -in -niː
accompany.PTCP.M -∅ -OBJ.1SG
'his accompanying me' (Muscat)

It is on the basis of this apparently low selectivity for phonological host that this is identified as a clitic in the literature.

However, I dispute that this is displaying low selectivity for phonological host at all. Rather, it shows low selectivity with regards to the lexical category of its syntactic head, but very high phonological selectivity in that it always attaches to its syntactic head — a classic affix behaviour.

This is the opposite of English, where in the English examples in (337) below, the enclitic possessive *'s* attaches to a range of lexical categories as phonological host but solely selects a noun ('the dragon') as syntactic head.[6]

(337) English *'s* Low Phonological Selectivity

a. The dragon I'm training's teeth are sharp

b. The dragon over there's baby will kill us all

c. The dragon I told you about's tail is long

d. The dragon whose scales are green's home is full of gold

In Arabic, the formatives in question are always directly attached to the syntactic unit they form a relationship with. Although verbs can be a domain of cliticisation, as Klavans (1985) notes it is a property of affixes rather than clitics to be associated with a syntactic head rather than phrase position.

The dative and negation markers show high selectivity as well, in that they also attach to their syntactic head, but these heads are a much more restricted group of lexical categories. The dative can only incorporate into the verb in most varieties, but also the participle in Makːan as shown in (338).

(338) Makːan Dative Attachment to Participle and Verb (Kabrah 2004, p. 80)

[6] For discussion of the clitic versus affix status of the English possessive see Lowe (2016) and references therein.

a. magsuːm -la -naː
destine.PTCP -DAT -OBJ.1PL
'it was destined for you'

b. raħ -∅ -la -ha
go.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F
'he went to her'.

The negation [ʃ] only attaches to verbs or non-verbal predicates in most varieties at the end of the string of endings. It highly selects for this syntactic head. For example, in Sˤanʕaːniː the dative and indirect object are optionally incorporated into the word. When they are, the negation marker follows them, as in (339-a). However, if the dative and indirect object are not incorporated into the prosodic word but form a separate word, the negation marker still attaches to the verb rather than continuing to follow the dative, as in (339-b).

(339) Sˤanʕaːniː Negation Attachment to Verb (Watson 2002, p. 62)

a. maː katab -∅ -la -naː -ʃ
NEG write.PFV -SBJ.3SG.M -DAT -OBJ.1PL -NEG
'he did not write to us'

b. maː katab -∅ -ʃ la -naː
NEG write.PFV -SBJ.3SG.M -NEG DAT -OBJ.1PL
'he did not write to us'

This suggests high selectivity on the part of the negation marker, a classic affix trait; and furthermore, if the tendency for affixes to not follow clitics holds, this would also suggest that the dative and object markers are also affixes.

Therefore, whilst the non-subject pronominal formative does select heads from a wide range of lexical catgories, all of these formatives highly select for their syntactic head as phonological host. This is a canonical affix behaviour, and would be atypical for clitics.

4.3.2 Arbitrary Gaps

The canonical affix can display arbitrary gaps, whereas clitics tend to not do this. The presence of arbitrary gaps suggests affix status; the absence of arbitrary gaps is not conclusive either way.

The dative and negation do not display arbitrary gaps; however, there are co-occurence restrictions with multiple instantiations of the non-nominative formative as direct and indirect object that appear to be arbitrary gaps.

The pronominal direct and indirect objects can co-occur, but only if there is intervening material. In most dialects, this is achieved by ordering OBJ -DAT -OBJ (as in ‘give it to me’) with the second object marker functioning as indirect object. Only in Libyan Tripoli are there co-occurence restrictions in this ordering. They only co-occur for the third person indirect object as shown in (340), and for the rest of the paradigm the dative and indirect object form a separate phonological word as in (341), thus forms where they are incorporated are ungrammatical as in (341-b).

(340) Libyan Tripoli Third Person Indirect Object Pronominal (Yoda 2005, p. 126)

a.

ʒab	-∅	-u	-l	-u
bring.PFV	-SBJ.3SG.M	-OBJ.3SG.M	-DAT	-OBJ.3SG.M

‘he brought it to him’

b.

ʒab	-∅	-u	-l	-ək
bring.PFV	-SBJ.3SG.M	-OBJ.3SG.M	-DAT	-OBJ.2SG.M

‘he brought it to you (sg)’

(341) Libyan Tripoli Non-Third Person Indirect Object Periphrasis (Yoda 2005, p. 126)

a.

bəʕ	-tʃ	-ək	li	-lu
send.PFV	-SBJ.1SG	-OBJ.2SG.M	DAT	-OBJ.3SG.M

‘I sent you to him’

b. *bə-ʕtʃ-ək-l-u

However, in varieties that put the indirect object before the direct object (as in ‘give me it’), the empty formative [ijːa] is included between them, as shown in (342) for Iraqi[7] and Muscat. This formative has no meaning, so I am glossing it as ∅.

[7] This order is found in Erwin (1963, p. 144) but in Majdi (1988, p. 60) the opposite ordering is found:

(i)

kitab	-naː	-haː	-l	-ak
write.PFV	-SBJV.1PL	-OBJ.3SG.F	-DAT	-OBJ.2SG.M

‘we wrote it for you’

(342) Empty Formative [-ijːa-] Between Direct and Indirect Objects

a. ʤiːbi -l -hum -ijːaː -ha
bring.IMP.F -DAT -OBJ.3PL.M -∅ -OBJ.3SG.F
'bring it (f.) to them' (Iraqi, (Erwin 1963, p. 144))

b. kisfiː -l -iː -jːaː -haː
hem.IMP.F -DAT -OBJ.1SG -∅ -OBJ.3SG.F
'hem (fs) it (fs) for me!' (Muscat, (Glover 1988, p. 212))

c. naːdj -uː -l -iː -jːaː -h
call.PFV -SBJ.3PL.M -DAT -OBJ.1SG -∅ -OBJ.3SG.M
'they called him for me' (Muscat, (Glover 1988, p. 212))

This is the formative that is used across Arabic when producing an utterance solely of an object formative. These are always bound, so bind to -ijːaː-, as shown in (343-b).

(343) a. man tu -hibː
who SBJ.2SG -love.IMPF
'Who do you love?'

b. ijːaː -ha
∅ -OBJ.3SG.F
'Her'

The gaps in Libyan Tripoli would suggest these formatives are closer to affixes than clitics; however, otherwise the data is inconclusive for this criteria.

4.3.3 Morphophonological Idiosyncracies

Canonical affixes display morphophonological idiosyncracies, whereas canonical clitics do not.

Regular verbs in Arabic can be analysed as containing three root consonants, so in the verb [katab] meaning 'he wrote' the three root consonants are k-t-b. There is a set of verbs where the second root consonant is a semivowel, and this surfaces as a long vowel, so in the verb [qaːl] meaning 'he said' the root is q-w-l. These verbs are known as 'hollow verbs'. In the 1st and 2nd person, the long vowel is shortened and often undergoes vowel raising. A selection of these verbs are included in Table 4.2 below, where on the left is the third person form of the verb and on the right is a non-third person form of the same verb.

Table 4.2: Arabic Hollow Verbs

Verb	Translation	Verb	Translation
ʃaːf	'he saw'	ʃuf-t	'I saw'
qaːl	'he said'	qul-ti	'you (sg) said'
naːm-at	'she slept'	nim-tum	'you (pl) slept'
saːr-u	'they walked'	sir-na	'we walked'

In many dialects, this shortening and vowel raising is triggered by the subject marker but **not** the object formative. Therefore, minimal pairs are found such as the following in (344) from Palestinian, where the first person plural subject marker [-na] does trigger the vowel shortening and raising in (344-a), but the homophonous first person plural object marker [-na] does not trigger the same process in (344-b).

(344) Palestinian Vowel Alternation triggered by Subject not Object (Abu-Salim 1982a, p. 150-156)

a. ʃuf -na
see.PFV -SBJ.1PL
'we saw'

b. ʃaːf -∅ -na
see.PFV -SBJ.3SG.M -OBJ.1PL
'he saw us'

That the object doesn't trigger this morphophonological alternation is used as evidence that the object is a clitic by Kenstowicz and Kisseberth (1979) and Abu-Salim (1982a) among others.

This process is usually called Closed Syllable Shortening (CSS), and is often conflated with a similar process that occurs when a CVːC syllable that is not permitted by the variety's syllable restrictions is created through morphological means, and this is shortened to CVC, but without any vowel quality alternations (Abu-Salim 1982a; Abu-Mansour 1992; Younes 1995). For Abu-Mansour (1992) the two types are the same shortening process, with morpholexical vowel quality alternations; however Broselow (2001) recognises the oddness of closed syllable shortening occurring in varieties that permit other CVVC syllables. As I demonstrated in Chapter 2, there are far more CVːC syllables permitted in Arabic varieties than previously recognised, but there are still restrictions. The distribution of closed syllable shortening across the sixteen varieties I explored here is summarised in Table 4.3 below.

Table 4.3: Closed Syllable Shortening Across Arabic Varieties and Endings

Key: Y = Yes, N = No, O = Optional, P = Prefix/Proclitic, - = No Data

Variety	Subj	DirObj	Dat+Obj	Neg	Poss
Cairene	**Y**	**Y**	**Y**	**Y**	**Y**
Palestinian	**Y**	N	**Y**	**Y**	N
Sˤanʕaːniː	**Y**	N	**Y**	O	N
Algerian	**Y**	N	-	**Y**	N
Makːan	**Y**	N	**Y**	P	N
Tunisian	**Y**	O	O	-	O
Muscat	**Y**	N	N	P	N
Qatari	**Y**	N	N	P	N
Iraqi	**Y**	N	N	P	N
Wadi Ramː	**Y**	N	N	P	N
Libyan Tripoli	**Y**	N	N	N	N
Moroccan Casablanca	**Y**	N	-	-	N
Lebanese	**Y**	N	-	-	N
Rufaidah	**Y**	N	N	-	N
Rwaili	**Y**	N	N	-	N
Haːyil	**Y**	N	N	-	N

CSS occurs in all varieties with vowel raising with the 1st and 2nd person subject. It occurs regardless of the syllable restrictions in the variety, as illustrated by the following pair from Wadi Ramː where the alternation occurs with the subject in (345-a), but CVːCC syllables are permitted elsewhere as in (345-b).

(345) Wadi Ramː CSS Occurs Although CVːCC Syllables Permitted (Al Mashaqba 2015, p. 208, 119)

a. ʤib -t
bring.PFV -SBJ.1SG
'I brought'

b. ʤaːb -∅ -l. -hum
bring.PFV -SBJ.3SG.M -DAT -OBJ.3PL
'he brought for them'

A similar set is found in Tunisian, where the long vowel in (346-a) is shortened and raised with the 1st person plural subject marker (346-b), however a long version of this raised vowel can exist with the homophonous possessive marker in (346-c).

(346) Tunisian CSS Occurs Despite CVːC Syllables Permitted (Maamouri 1967, p. 51, 5)

a. raːħ -∅
go.PFV -SBJ.3SG.M
'he went'

b. ruħ -na
go.PFV -SBJ.1PL
'we went'

c. ruːħ -na
soul -POSS.1PL
'our soul'

CSS only occurs after the direct object in Cairene, and in Tunisian if would shorten an unstressed long vowel. Cairene does not permit any CVːC syllables, but when shortening occurs with the direct object, no vowel raising occurs as shown in (347-b), unlike with the 3rd person subject as in (347-a).

(347) Cairene Long Vowel Shortening without Raising in Non-Third Person Subject Cases (Watson 2002, p. 149)

a. ʃuf -na
see.PFV -SBJ.1PL
'we saw'

b. ʃaf -∅ -na
see.PFV -SBJ.3SG.M -OBJ.1PL
'he saw us'

Tunisian does not permit unstressed long vowels, so long vowels can persist after adding the object as in (348-a), but if the addition leaves the long vowel unstressed this shortens as in (348-c) compared to (348-b) — though notably without any vowel raising.

(348) Tunisian Open Syllable Shortening (Maamouri 1967, p. 159)

a. tˤaˈhunt -ik
mill -POSS.2SG
'your (sg.) mill'

b. ˈʃriː -tu
buy.PFV -SBJ.2PL
'you (pl.) bought'

c. ʃri -ˈtuː -hum
buy.PFV -SBJ.2PL -OBJ.3PL
'you (pl.) bought them'

After the dative and indirect object, shortening occurs in four dialects of the set: Palestinian, Cairene, Sˤanʕaːniː, and Makːan.

That the dative and indirect object can trigger shortening in Palestinian but not the direct object has been a matter of confusion in the literature. Pairs with the same syllable structure appear, but only shorten if the dative is included. In the following example, we only find shortening in (349-a) not (349-b) despite the same number of segments and syllable structure.

(349) Palestinian Shortening with Dative not Direct Object (Abu-Salim 1982a, p. 149)

a. /ʤaːb-l-i/
ʤab -∅ -l -i
bring.PFV -SBJ.3SG.M -DAT -OBJ.1SG
'he brought to me'

b. /ʤaːb-ha/
ʤaːb -∅ -ha
bring.PFV -SBJ.3SG.M -OBJ.3SG.F
'he brought her'

Abu-Salim (1982a) suggests this is due to having extra formative boundaries, however a simpler solution exists. The epenthetic vowel in Palestinian is inserted postlexically, so at the lexical level illicit CVːCC can be created by addition of the dative. This illicit CVːCC is shortened at the lexical level to create CVCC. This syllable is still not permissible medially in Palestinian unless C2 is coronal, so the cluster is broken up by epenthesis on the postlexical level. This occurs for all non-first person singular cases for the dative, as shown in Table 4.4 below. Note the variation in epenthetic vowel quality does not matter.

Table 4.4: Palestinian Dative Behaviour with Epenthesis

Underlying Form	Gloss	Surface Form	Meaning
ʤaːb-∅-l-i	bring.PFV -SBJ.3SG.M -DAT -OBJ.1SG	ʤabli	‘he brought to me’
ʤaːb-∅-l-k	bring.PFV -SBJ.3SG.M -DAT -OBJ.2SG	ʤabli̱k	‘he brought to you (sg.)’
ʤaːb-∅-l-h	bring.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M	ʤablu̱h	‘he brought to him’
ʤaːb-∅-l-ha	bring.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M	ʤabli̱ha	‘he brought to her’
ʤaːb-∅-l-na	bring.PFV -SBJ.3SG.M -DAT -OBJ.1PL	ʤabli̱na	‘he brought to us’
ʤaːb-∅-l-kum	bring.PFV -SBJ.3SG.M -DAT -OBJ.2PL	ʤabli̱kum	‘he brought to you (pl.)’
ʤaːb-∅-l-hum	bring.PFV -SBJ.3SG.M -DAT -OBJ.3PL	ʤabli̱hum	‘he brought to them’

I propose that the apparently aberrant behaviour of the 1st person singular suffix case is a matter of analogy with the rest of the paradigm. In all forms other than the first person singular indirect object, the underlying long vowel in the verb is shortened according to a restriction on CVːCC syllables. On analogy with this, the first person case also undergo shortening, even though phonologically the underlying CVːC syllable is acceptable. It is unusual for analogy to occur with the first person singular case changing to fit the rest of the paradigm, rather than other members of the set changing to fit the first person singular. However, there is no clear phonological reason for this alternation, so it appears to be a morphologised paradigm matter.

Cairene does not permit any internal CVːC syllables, so the shortening with the dative is a logical consequence of this restriction, as shown in (350).

(350) Cairene Dative Long Vowel Shortening (Broselow 1976, p. 69)

/gaːb-haː-l-ak/

gab -∅ -haː -l -ak
bring.PFV -SBJ.3SG.M -OBJ.3SG.F -DAT -OBJ.2SG

'he gave it to you'

Sˤanʕaːniː shortens long vowels with the dative, as shown in (351).

(351) Sˤanʕaːniː Dative Long Vowel Shortening (Watson 2002, p. 67)

a. /raːħ-l-ih/
raħ -∅ -l -ih
go.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M

'he went off'

b. /gaːl-l-iː/
gal -∅ -l -iː
say.PFV -SBJ.3SG.M -DAT -OBJ.1SG

'he said to me'

c. /ji-guːl-l-naː/
ji -gul -l<u>a</u> -naː
SBJ.3SG -say.IMPF -DAT -OBJ.1PL

'he says to us'

d. /ji-biːʕ-l-ak/
ji -biʕ -l -ak
SBJ.3SG -sell.IMPF -DAT -OBJ.2SG.M

'he sells to you (m.s.)'

Makːan does shorten long vowels with the dative, whether in verbs or participles, as shown in (352-a) and (352-b) respectively.

(352) Makːan Dative Long Vowel Shortening in Verbs and Participles (Kabrah 2004, p. 80)

a. /raːħ-l-ha/
raħ -∅ -l<u>a</u> -ha
go.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F

'he went to her'

b. /magsuːm-l-na/
magsum -l<u>a</u> -na
destine.PTCP.M -DAT -OBJ.1PL

'it was destined for us'

As discussed in Chapter 2, Makːan reduces internal CVːC when created by affixation, though internal CVːC created by syncope are accepted.

Shortening occurs optionally with the dative for Tunisian because of the aforementioned restriction on unstressed long vowels, as shown in (353).

(353) Tunisian Unstressed Long Vowels Shortened (Maamouri 1967)

a. ˈʤiːb -ha
bring.IMP -OBJ.3SG.F
'bring it'

b. ʤib -ˈhaː -l -i
bring.IMP -OBJ.3SG.F -DAT -OBJ.1SG
'bring it to me'

It is rarer after negation, occurring in only Palestinian and Cairene, and optionally in Sˤanʕaːniː. In Palestinian, it follows from a restriction on final CVːCC syllables. After shortening occurs, an epenthetic vowel is inserted due to the restrictions on word final consonantal clusters, as shown in (354).

(354) Palestinian Negation Long Vowel Shortening and Epenthesis (Abu-Salim 1982a, p. 150)

a. ʃaːf -∅
see.PFV -SBJ.3SG.M
'he saw'

b. ʃaf -∅ -i̲ʃ
see.PFV -SBJ.3SG.M -NEG
'he didn't see'

In Cairene, the shortening with the negation follows from the restriction on word internal CVːC syllables, as shown in (355). Note epenthesis does not occur in (355-b) as it did in (354-b) for Palestinian as final consonantal clusters are permitted in Cairene.

(355) Cairene Negation Long Vowel Shortening (Younes 1995, p. 158)

a. ʃaːf -∅
see.PFV -SBJ.3SG.M
'he saw'

b. ʃaf -∅ -ʃ
see.PFV -SBJ.3SG.M -NEG
'he didn't see'

In Sˤanʕaːniː, the shortening with the negation marker is optional, as shown in (356) where the optional long vowel is marked with brackets. This suggests that the negation marker is not fully subject to the syllable restriction of this variety. However, final coronal sibilants often escape syllable restrictions, as seen in English.

(356) Sˤanʕaːniː Optional Negation Long Vowel Shortening (Watson 2002, p. 67)

(maː) ka(ː)n -∅ -ʃ
NEG be.PFV -SBJ.3SG.M -NEG

'he was not'

Long vowel shortening only occurs in Cairene and Tunisian if the vowel is unstressed after the possessive, due to the Cairene restriction on non-final CVːC syllables and Tunisian avoidance of unstressed long vowels. This is shown in (357) and (358) respectively.

(357) Cairene Possessive Long Vowel Shortening (Broselow 1976, p. 68)

a. beːt ħala
house Hala
'Hala's house'

b. bet -ha
house -POSS.3SG.F
'her house'

(358) Tunisian Possessive Long Vowel Shortening if Unstressed Long Vowel. (Maamouri 1967, p. 159, 161)

a. ˈmsaːmir

'nails'

b. msaˈmir -kum
nail.PL -POSS.2PL
'your nails'

c. ˈxaːl
'uncle'

d. ˈxaːl -ik
uncle -POSS.2SG
'your uncle'

4.3.4 Conclusion: Clitic versus Affix

Therefore, it is clear we have two processes happening here: a morphophonological alternation including stem allomorphy/vowel raising triggered by the 1st and 2nd person subject marker across all varieties regardless of syllable structure restrictions that I have previously labelled CSS-MORPH; and a closed syllable shortening process without vowel changes triggered by variety-specific syllable structure restrictions that I have previously labelled CSS-PHON. These processes are exemplified in (359) and (360), where a CVːC stem is found in (359-a) and (360-a), which undergoes shortening and raising in the 1st person plural subject forms in (359-b) and (360-b), but only undergoes shortening (without raising) in (359-c) not (360-d) as Cairene does not permit medial CVːC syllables whereas Palestinian does.

(359) Cairene Morphophonological Alternations and Phonological Shortening (Watson 2002, p. 149)

a. ʃaːf -∅
see.PFV -SBJ.3SG.M
'he saw'

b. ʃuf -na
see.PFV -SBJ.1PL
'we saw'

c. ʃaf -∅ -na
see.PFV -SBJ.3SG.M -OBJ.1PL
'he saw us'

(360) Palestinian Morphophonological Alternations and no Phonological Shortening (Abu-Salim 1982a, p. 129)

a. ʃaːf -∅
see.PFV -SBJ.3SG.M
'he saw'

b. ʃuf -na
see.PFV -SBJ.1PL
'we saw'

c. ʃaːf -∅ -na
see.PFV -SBJ.3SG.M -OBJ.1PL
'he saw us'

Whilst CSS-Morph, the subject-triggered morphophonological alternation, may have been motivated by syllable restrictions in earlier forms of Arabic, it is not now, but rather is a morphological matter.

That the object does not trigger a morphophonological alternation found in the subject paradigm is not sufficient evidence to claim it is a clitic. Zwicky and Pullum (1983) argue that affixes should trigger morphophonological alternations, and if the element in question does not, it is not an affix but a clitic. However, this argument does not hold. This is recognised in patterns in other languages. It is on a similar basis that Pierrehumbert (1980) and Nevis (1986) claimed Finnish possessives are clitics, as they fail to trigger consonant gradation. However, Kanerva (1987) argues that the possessive triggers different allomorphic alternations, just not consonant gradation. That affixes can have different phonological impacts is also seen in English, where Class I past tense triggers allomorphy in verbs, whereas Class II past tense is more regular d/t affixation.

(361) English Class I Past Tense:

a. see → saw

b. sing → sang

(362) English Class II Past Tense:

a. shout → shouted

b. walk → walked

The insights of Lexical Phonology are particularly important here, as they provide a mechanism for affixes to trigger different processes while remaining affixes. If the subject marker is attached on Level 1 but the others on Level 2, CSS-Morph can occur on Level 1 and CSS-Phon can occur on Level 2. This then accounts for the difference in affix behaviour without having to use the clitic distinction that as argued above does not fit the data here. In particular, the subject affix and the plural affix attach on Level 1, while the Dative, Non-subject pronominal (functioning as direct object, indirect object, or possessive depending on context) and negation attach on Level 2.

Above, I have explored how these endings behave with regards to the Zwicky and Pullum criteria. The key criteria for these endings are host selectivity, arbitrary gaps and morphophonological idiosyncracies. I have argued that these endings have high host selectivity in that they select their syntactic head, which is much more of an affix property than clitic. While they do not trigger the vowel shortening and stem allomorphy in hollow verbs that subject markers do, I have argued that this is not sufficient to suggest these are clitics. In some varieties, there are some arbitrary co-occurence gaps, which weakly suggests affix.

Therefore, based on these criteria, there does not seem to be strong evidence that they are clitics, even if they display different behaviours to the subject markers.

4.4 Pre-'clitic' lengthening

The discussion of morphophonological idiosyncracies thus far has focused on the differences between the subject marker and other endings. In this, I have aimed to demonstrate that whilst there are differences between these two groups, the difference does not entail that these endings are clitics.

I will now turn to a phenomenon that seems to be triggered by these endings, and is often called 'pre-clitic lengthening'. In the Tunisian example in (363), the subject marker ending [-tu] surfaces with a long vowel as [-tuː] when followed by the direct object bound formative.

(363) Tunisian'Pre-Clitic Lengthening' (Maamouri 1967, p. 159)

a. ˈʃriː -tu
buy.PFV -SBJ.2PL
'you (pl.) bought'

b. ʃri -ˈt**uː** -hum
buy.PFV -SBJ.2PL -OBJ.3PL
'you (pl.) bought them'

Likewise in Qatari, where the final vowel of the verb lengthens before both the pronominal direct object and the dative, as shown in (364).

(364) Qatari 'Pre-Clitic Lengthening' (Al-Sulaiti 1993, p. 242)

a. ʃara -∅
buy.PFV -SBJ.3SG.M
'he bought'

b. ʃar**aː** -∅ -hum
buy.PFV -SBJ.3SG.M -OBJ.3SG.M
'he bought them'

c. ʃar**aː** -∅ -l -kum
buy.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M
'he bought for you'

In individual dialects, it appears to be idiosyncratic, and debates abound as to whether these vowels are underlyingly short but lengthened before endings or underlyingly long but lengthened word-finally.

Those advocating for underlyingly long vowels that are shortened word finally include Abdo (1969) and Abu-Salim (1982b, pp. 134–147) for Lebanese and Palestinian, Bosisio (2003, p. 203) for Ammani, Abu-Mansour (1987, p. 134) for Makkan, Abumdas (1985, p. 46) for Zliten Libyan, Glover (1988, p. 204) for Muscat, and McCarthy (2005) and Faust and Ulfsbjorninn (2018) for Cairene. The primary arguments for this position are that firstly, final vowel shortening has extensive typological precedents, unlike a requirement for long vowels before a suffix (McCarthy 2005, pp. 7–9); secondly, that it is more economical to have one shortening process across a variety rather than a Duke of York derivation with lengthening and subsequent shortening of unstressed vowels (Abu-Salim 1982b; McCarthy 2005); and thirdly, that the non-application of apocope follows from word-final vowels being underlyingly long (Abu-Mansour 1987, p. 133).

Those advocating for underlying short vowels that are lengthened before suffixes include Abdel-Massih et al. (1979, p. 323), Broselow (1976, pp. 106–118), Watson (2002, pp. 201–203), Ali (2013) and Abboud-Haggar (2015) for Cairene, Benyagoub (2017, p. 44) for Bechar Algerian, Prochazka Jr. (1988) for a broad set of Saudi Arabian varieties, Bohas (1978, p. 98) for Damascene, Abdul-Karim (1980, pp. 101–105), Brame (1971, p. 584) and Haddad (1984) for Lebanese and Palestinian, Hwaidi (2016) for Al'ain Libyan and Hamid (1984) for Sudanese. The primary arguments for this position are that underlyingly long vowel analysis creates a rule ordering paradox with stress placement (Broselow 1976); that homophony avoidance (Broselow 1976) or redefining syncope context (Watson 2011a) can account for the lack of apocope; and that the pre-suffix vowel lengthening is a negative constraint on monomoraic-final morphemes pre-suffixally (Watson 2002, p. 202).

I argue that the stem-final vowels are underlyingly short in varieties that undergo these alternations and lengthened as part of a prosodic preference for affixes to attach after a heavy syllable, that is, a stressable foot. I demonstrate that not only can stem and affix vowels be lengthened pre-suffixally, but that consonants can also be lengthened pre-suffixally. Similar to the vowel quantity alternations in (364) and (363), in Lebanese when followed by the direct object bound pronominal the /t/ from the third person subject marker [-it] geminates to become [itː], as shown in (365)[8].

(365) Lebanese 'Pre-Clitic Lengthening' (Haddad 1984, p. 158)

a. sakar -it
intoxicate.PFV -SBJ.3SG.F
'she intoxicated'

b. sakar -**itː** -ak
intoxicate.PFV -SBJ.3SG.F -OBJ.2SG
'she intoxicated you'

[8]The keen Arabist may query whether the verb in (365) should have a medial geminate as the causative meaning suggests a different verb form. According to Haddad (1984), the transitivity is reflected in the vowel quality in some Form I verbs, thus we find the minimal pair [sikir], 'he became drunk', versus [sakar], 'he intoxicated'. For suggestions as to the historical origins of this pattern, see Haddad (1984, pp. 90-95).

That is, while geminates are found extensively in the morphology of Arabic, consonants that undergo this pre-suffixal length alternation are not underlyingly geminates but rather singleton consonants that are lengthened to meet this prosodic attachment preference. Furthermore, this prosodic preference underlies a set of disparate exceptions to metrical stress across Arabic varieties that have hitherto belied satisfactory analyses. Understanding how these exceptions work is key to understanding the nature of this prosodic attachment preference.

If we assume that this length alternation in consonants pre-suffixally and word-finally is the same process as the vowel length alternations, we must ask whether the singleton or geminate consonant is underlying. The options here are that underlying singletons are geminated before suffixes, or that underlying geminates are degeminated word-finally. Let us first take the analysis that these are underlyingly geminates but degeminated word finally. If word-final degemination affects these consonants, we would expect to find word-final degemination across the varieties. However, this is not what is found. Rather, evidence from stress suggests that final geminates do persist. Recall that across Arabic varieties, stress falls on the rightmost available foot, though the unit of extrametricality does differ. Davis and Ragheb (2014) give the following minimal pair in (366) where the form without a final geminate in (366-a) receives initial stress whereas the form in (366-b) with a final geminate receives final stress, thus escaping the final consonant extrametricality.

(366) Cairene Final Geminate Stress (Davis and Ragheb 2014, p. 4)

a. ˈʔamal
   'hope'
b. ʔaˈmalː
   'more/most boring'

Similar cases of final geminates escaping extrametricality in other varieties were discussed in Section 3.4. If word-final degemination occurred, we would not expect this difference in whether the final syllable received stress in the examples above. Furthermore, the geminate form is predominately found before vowel initial suffixes.

Consider the following paradigm in Table 4.5 from a differential Lebanese variety discussed in Abdul-Karim (1980, p.98). Here, pre-suffixal gemination is optional, but only occurs before vowel initial suffixes.

ˈdˤarab-it		she hit
dˤarab-ˈit-ni		she hit me
dˤarab-ˈit-ak	∼ dˤarab-ˈ**itː**-ak	she hit you (m.s.)
dˤarab.-ˈit-ik	∼ dˤarab-ˈ**itː**-ik	she hit you (f.s.)
dˤarab-ˈit-u	∼ dˤarab-ˈ**itː**-u	she hit him
dˤarab-ˈit-(h)a	∼ dˤarab-**itː**-a	she hit her
dˤarab-ˈit-na		she hit us
dˤarab-ˈit-kun		she hit you (pl.)
dˤarab-ˈit-(h)-un	∼ dˤarab-ˈ**itː**-un	she hit them

Table 4.5: Lebanese Pre-Suffix Gemination (Abdul-Karim 1980)

If these consonants are underlyingly geminate, we would need a rule for degemination syllable finally for tautosyllabic geminates – not to mention also optionally between vowels, regardless of whether the following syllable is light or heavy. However, a key concern of Chapter 2 of this thesis is the distribution of geminates, where I have shown that these either are preserved as tautosyllabic geminates as in (367), or undergo epenthesis as in (368), but there is no evidence for medial degemination elsewhere.

(367) Algerian Geminate Singleton Minimal Pairs without Degemination (Bouhadiba 1988, p. 323)

a. kadːb -u
accuse.of.lying.PFV -SBJ.3PL
'they accused (s.o.) of lying'

b. kadb -u
liePFV -SBJ.3PL
'they lied'

(368) Makːan Geminate vs Singletons with Epenthesis (Kabrah 2004, p. 90, 14)

a. madː -∅ -<u>a</u>ha
extend. -SBJ.3SG.M -OBJ.3SG.F
'he extended it (f.)'

b. katab -∅ -ha
write.PFV -SBJ.3SG.M -OBJ.3SG.F
‘he wrote it (f.)’

Therefore, it is difficult to argue that they are underlyingly geminate but undergo final degemination where the stress data suggests otherwise; and similarly, it is difficult to argue they are underlyingly geminate but degeminate before a consonant word medially when in such cases epenthesis should occur. Rather, it appears that singleton consonants geminate pre-suffixally. It is unlikely that when such an alternation exists that the consonants geminate but the vowels shorten pre-suffixally. As Table 4.6 demonstrates, some form of lengthening occurs in all varieties except Sˤanʕaːniː and Moroccan Casablanca.

Table 4.6: Pre-Clitic Lengthening

Variety	Affix Vowel Lengthening	Affix Consonant Lengthening	Dative Gemination	Stem Consonant Gemination
Tunisian	Y	Y		
Lebanese	Y	Y		
Iraqi	Y	Y		
Muscat	Y	Y		
Rwaili	Y	Y		
Haːyil	Y	Y		
Qatari	Y		Y	
Makːan	Y		Y	
Libyan Tripoli		Y	Y	Y
Algerian	Y			
Cairene	Y			
Palestinian	Y			
Wadi Ramː	Y			
Rufaidah	Y			
Moroccan Casablanca		Y		
Sˤanʕaːniː				

I will now explore each variety in the same order as this table, and have put the lengthening in bold for ease of reading in the following examples.

Tunisian lengthens vowels and affix consonants, as in (369) where the final vowel [u] of the subject marker in (369-a) is [uː] before the pronominal object in (369-b), and the final consonant in the feminine singular 3rd person subject marker [it] in (369-c) is geminated to [itː] in (369-d) before the pronominal object.

(369) Tunisian Vowel Lengthening and Gemination (Maamouri 1967, p. 159, 24)

a. ʃriː -tu
buy.PFV -SBJ.2PL
'you pl. bought'

b. ʃri -t**uː** -hum
buy.PFV -SBJ.2PL -OBJ.3PL
'you bought them.

c. baːs -it
kiss.PFV -SBJ.3SG.F
'she kissed'

d. bas -i**tː** -ik
kiss.PFV -SBJ.3SG.F -OBJ.2SG
'she kissed you'

Lebanese lengthens vowels and affix consonants, as shown in (370) where the first person subject marker [-na] in (370-a) is [-naː] before the object pronominal in (370-b), while the feminine singular third person subject marker [-it] in (370-c) is geminated to [-itː] before the object pronominal in (370-d)

(370) Lebanese Suffix Vowel Lengthening and Gemination (Haddad 1984, p. 11, 157)

a. ħamal -na
neglect.PFV -SBJ.1PL
'we neglected'

b. ħamal -n**aː** -hum
neglect.PFV -SBJ.1PL -OBJ.3PL
'we neglected them'

c. sakar -it
intoxicate.PFV -SBJ.3SG.F
'she intoxicated'

d. sakar -i**tː** -ak
intoxicate.PFV -SBJ.3SG.F -OBJ.2SG

'she intoxicated you'

Iraqi lengths vowels and affix consonants. The first person plural subject [-na] in (371-a) is [-naː] before the object pronominal in (371-b) and (371-c). In (371-d)–(371-e), the affix ends in a consonant and this is lengthened from [-at] to [-atː] before the vowel-initial object suffixes.

(371) Iraqi Vowel and Affix Consonant Lengthening (Majdi 1988, p. 60, 176)

a. kitab -na
write.PFV -SBJ.1PL
'we wrote'

b. kitab -n**aː** -ha
write.PFV -SBJ.1PL -OBJ.3SG.F
'we wrote it'

c. kitab -n**aː** -h**aː** -l -ak
write.PFV -SBJ.1PL -OBJ.3SG.F -DAT -OBJ.2SG
'we wrote it for you'

d. nis -a**tː** -ak
forget.PFV -SBJ.3SG.F -OBJ.2SG.M
'she forgot you (m.)'

e. qir -a**tː** -a
read.PFV -SBJ.3SG.F -OBJ.2SG.F
'she read it'

Muscat lengthens vowels, affix consonants, and the dative. In (372-a), the affix vowel for the third person plural marker is long before the object pronominal; in (372-b) and (372-c) the final consonant of the affixes are geminated before the object pronominals; while in (372-d) and (372-e) the dative is geminated, though the geminated and epenthesised (372-e) does alternate with (372-f) without gemination or epenthesis.

(372) Muscat Affix Vowel Lengthening, Affix Consonant and Dative Gemination (Glover 1988, p. 75, 203, 211)

a. ba -ji -tʕalːam -**uː** -ha
FUT -SBJ.3 -learn.IMPF -SBJ.PL -OBJ.3SG.F
'they will learn it'

b. kal -**it:** -oh
eat.PFV -SBJ.3SG.F -OBJ.3SG.M
'she ate it'

c. na:de: -ta**n:** -oh
call.PFV -SBJ.3PL.F -OBJ.3SG.M
'you (f.p.) called him'

d. katab -t -**i̱l:** -oh
write.PFV -SBJ.1SG -DAT -OBJ.3SG.M
'I/you wrote to him'

e. katb -∅ -**i̱l:** -oh
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M
'he wrote to him'

f. katab -∅ -l -oh
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M
'he wrote to him'

Rwaili lengthens vowels, and geminates the 2nd feminine plural ending consonant. Without an object suffix, the final subject affix vowel is short in (373-a) and (373-c). However, with object suffix affixation, it becomes long in (373-b) and (373-d). If the subject affix ends in a consonant as in (374), this consonant lengthens before the affix.

(373) Rwaili Affix Vowel Lengthening (Prochazka Jr. 1988, p. 108, 184, 78, 188)

a. ti -ji
SBJ.2SG -come.IMPF
'you come'

b. ta -r**u:** -ni
SBJ.2SG -see.IMPF -OBJ.1SG
'you see me'

c. ʃif -ti
see.PFV -SBJ.2SG
'you saw'

d. ʃif -t**i:** -ni
see.PFV -SBJ.2SG -OBJ.1SG
'you saw me'

(374) Rwaili Consonant Lengthening (Prochazka Jr. 1988, p. 78, 188)

a. ʃif -tin
see.PFV -SBJ.2PL.F
'you (f.pl.) saw'

b. ʃif -ti**nː** -əhum
see.PFV -SBJ.2PL.F -OBJ.3PL
‘you (f.pl.) saw them’

Haːyil lengthens stem final vowels, as shown in (375) where (375-a)–(375-b) show the short final vowel in non-affixed forms, but in (375-c)–(375-d) the final stem vowel is long. Haːyil also lengthens the final affix vowel in (376-a)–(376-c) and final affix consonant in (376-d)–(376-e).

(375) Haːyil Stem Final Vowel Lengthening (Prochazka Jr. 1988, p. 91-92, 171–172)

a. miʃa -∅
go.PFV -SBJ.3SG.M
‘he went’

b. nisi -∅
forget.PFV -SBJ.3SG.M
‘he forgot’

c. lig**aː** -∅ -na
find.PFV -SBJ.3SG.M -OBJ.1PL
‘he found us’

d. kal**aː** -∅ -h
eat.PFV -SBJ.3SG.M -OBJ.3SG.M
‘he ate it (m.)’

(376) Haːyil Suffix Final Vowel and Consonant Lengthening (Prochazka Jr. 1988, p. 78, 188)

a. ʃuf -ti
see.PFV -SBJ.2SG
‘you saw’

b. ʃuf -t**iː** -ham
see.PFV -SBJ.2SG -OBJ.3PL
‘you saw them’

c. ʃuf -n**aː** -ham
see.PFV -SBJ.1PL -OBJ.3PL
‘we saw them’

d. ʃuf -ti**nː** -an
see.PFV -SBJ.2SG.F -OBJ.1SG
‘you saw me’

e.

ʃuf	-ti**nː**	-a̲na
‘see.PFV	-SBJ.2SG.F	-OBJ.1PL

‘you saw us’

Qatari lengthens final stem vowels, as shown in (377-b-c) compared to the short final stem vowel in (377-a). However, if the stem ends in a consonant so there is no stem vowel to lengthen, the dative affix consonant is optionally lengthened instead, as shown in (377-d).

(377) Qatari Vowel and Dative Consonant Lengthening (Al-Sulaiti 1993, p. 242)

a.

ʃara	-∅
buy.PFV	-SBJ.3SG.M

‘he bought’

b.

ʃar**aː**	-∅	-hum
buy.PFV	-SBJ.3SG.M	-OBJ.3PL

‘he bought them’

c.

ʃar**aː**	-∅	-l	-kum
buy.PFV	-SBJ.3SG.M	-DAT	-OBJ.2PL

‘he bought for you’

d. katb -∅ -i̲**lː** -oh ~

katab	-∅	-l	-oh
write.PFV	-SBJ.3SG.M	-DAT	-OBJ.3SG.M

‘he wrote to him’

Makːan lengthens the affix vowel and geminates the dative in (378). In (378-b), the subject suffix [-na] is lengthened before the object affix. (378-c)–(378-f) demonstrate the variation in lengthening a final affix vowel or dative consonant. In (378-c) and (378-d), we see that if there is an underlying vowel, Makːan will lengthen that; however in (378-e) and (378-f) there is no underlying vowel so an epenthetic one is inserted, thus Makːan will geminate the dative instead.

(378) Makːan Lengthens Affix Vowel and Dative Consonant (Kabrah 2004, p. 27, 90, 107)

a.

katab	-na
write.PFV	-SBJ.1PL

‘we wrote’

b. katab -n**aː** -ha
write.PFV -SBJ.1PL -OBJ.3SG.F
‘we wrote it’

c. /madː-u-l-ha/
madː -**uː** -la -ha
extend.PFV -SBJ.3PL -DAT -OBJ.3SG.F
‘they extended for/to her’

d. /madː-∅-l-ha/
madː -∅ -a**lː**a -ha
extend.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F
‘he extended for/to her’

e. katab -t -a**lː**a -ha
write.PFV -SBJ.1SG -DAT -OBJ.3SG.F
‘I wrote to her’

f. rameː -t -a**lː** -u
throw.PFV -SBJ.1SG -DAT -OBJ.3SG.M
‘I threw to him’

Libyan Tripoli lengthens the dative or verb stem, and sometimes the affix consonant (though this alternates with syncope). In (379), the final stem consonant is optionally lengthened, as is the dative affix consonant. (379-a)–(379-c) and (379-d)–(379-f) show the variants for two different items. In (380), the final affix consonant is lengthened instead.

(379) Libyan Tripoli Stem Consonant and Dative Lengthening (Yoda 2005, pp. 124-125)

a. ʒab -∅ -l -əm

b. ʒa**bː** -∅ -əl -əm

c. ʒa**bː** -∅ -ə**lː** -əm
bring.PFV -SBJ.3SG.M -DAT -OBJ.3PL
‘he brought to them’

d. ktʃəb -∅ -l -a

e. ktʃə**bː** -∅ -əl -a

f. ktʃə**bː** -∅ -ə**lː** -a
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.F
‘he wrote to her’

(380) Libyan Tripoli Final Affix Consonant Lengthening (Yoda 2005, p. 107)
dˤərb -ətʃː -u
hit.PFV -SBJ.3SG.F -OBJ.3SG.M

‘she hit him’

Algerian lengthens affix vowels, as in the following where the first person plural subject marker [-na] in (381-a) is [-naː] before the object pronominal in (381-b).

(381) Algerian Affix Vowel Lengthening (Bouhadiba 1988, p. 221, 228)

a. sabag -u
overtake.PFV -SBJ.3PL
‘they overtook’

b. sabag -uː -h
overtake.PFV -SBJ.3PL -OBJ.3SG.M
‘they overtook him’

As part of this Algerian unusually[9] lengthens the vowel in the feminine ending when followed by a C-only suffix. Thus, in (382-a) there is vowel lengthening when followed by the C-only 3rd person masculine suffix /h/, though epenthesis later follows this with a vowel; whereas when followed by the 3rd person feminine suffix /ha/ there is no need to lengthen the vowel — and in fact there is vowel reduction of /a/ to [ə], as seen in (382-b).

(382) Algerian Affix Internal Vowel Lengthening (Bouhadiba 1988, p. 288)

a. ħækm -**a**ːt -ə̲k
catch.PFV -SBJ.3SG.F -OBJ.2SG
‘she caught you’

b. ħækm -ət -hæ -l -kum
catch.PFV -SBJ.3SG.F -OBJ.3SG.F -DAT -OBJ.2PL
‘she caught her for you’

Cairene lengthens final short vowels before the ending, as in (383) where the final vowel [-a] of the verb is [aː] before the negation marker.

(383) Cairene Stem Vowel Lengthening (Broselow 1976, p. 70)

[9] A similar occurrence is also seen in the Libyan dialect discussed by Farwaneh (1997, p. 6) which lengthens the affix vowel of the 3rd feminine subject if followed by a vowel initial suffix: [ˈktib-at] ‘she wrote’, [ktiˈb-aːt-a] ‘she wrote it(m)’, but [ktiˈb-at-ha] ‘she wrote it (f)’

a. bana -∅
build.PFV -SBJ.3SG.M
‘he built’

b. ma ban**aː** -∅ -ʃ
NEG build.PFV -SBJ.3SG.M -NEG
‘he didn’t build’

Palestinian lengthens final short vowels before the ending, as in (384) where the [-u] of the subject marker is [-uː] before the direct object ending.

(384) Palestinian Affix Vowel Lengthening (Abu-Salim 1982a, p. 134)

a. ˈsaʔal -u
ask.PFV -SBJ.3PL
‘they asked’

b. saʔaˈl **-uː** -ha
ask.PFV -SBJ.3PL -OBJ.3SG.F
‘they asked her’

Wadi Ramː lengthens final vowels. In (385-a) the subject ending without further affixes has a short vowel, [-u], but when followed by object suffixes in (385-b)–(385-c) it is lengthened to [-uː]. Lengthening also happens with stem final vowels, as shown in (385-d)–(385-e).

(385) Wadi Ramː Lengthens Final Vowels (Al Mashaqba 2015, p. 136, 109, 243, 107)

a. ja -tˤulb -u
SBJ.3.M -ask.IMPF -SBJ.PL
‘they m. ask’

b. ji -tigaːsam **-uː** -hin
SBJ.3.M -share.IMPF -SBJ.PL -OBJ.3PL
‘they (m.) share them’

c. ji -tinaːwal **-uː** -ha
SBJ.3.M -cooperate.to.catch.IMPF -SBJ.PL -OBJ.3SG.M
they m. cooperate to catch it f.’

d. ʔabu

‘father’

e. ʔab**uː** -ha
father -POSS.3SG.F
'her father'

Rufaidah lengthens stem vowels, as shown in (386-c)–(386-d) where before suffixes the stem vowel is lengthened compared to the final short vowel in the non-suffixed (386-a)–(386-b), as well as affix vowels, as shown in the long pre-object vowels in (386-b)–(386-c) compared to the short final vowel in (387-a).

(386) Rufaidah Lengthens Stem Vowels (Prochazka Jr. 1988, pp. 90-91, 170)

a. maʃa -∅
go.PFV -SBJ.3SG.M
'he went'

b. rama -∅
shoot.PFV -SBJ.3SG.M
'he shot'

c. xaðaː -∅ -h
take.PFV -SBJ.3SG.M -OBJ.3SG.M
'he took him'

d. kalaː -∅ -ha
eat.PFV -SBJ.3SG.M -OBJ.3SG.F
'he ate it (f.)'

(387) Rufaidah Lengthens Affix Vowels (Prochazka Jr. 1988, p. 187)

a. ðarab -ti
hit.PFV -SBJ.2SG
'you hit'

b. ðarab -t**iː** -hum
hit.PFV -SBJ.2SG -OBJ.3PL
'you hit them'

c. ðarab -n**aː** -h
hit.PFV -SBJ.1PL -OBJ.3SG.M
'we hit him'

Moroccan Casablanca optionally geminates the feminine affix consonant, as shown in (388) where the feminine 3rd person singular subject marker [-at] is optionally [-ətː] before an object pronominal.

(388) Moroccan Casablanca Optional Affix Consonant Gemination (Boudlal 2001, p. 58)

a. kətb -at -u ~kətb -ə**tː** -u
write.PFV -SBJ.3SG.F -OBJ.3SG.M
‘she wrote it’

b. sərq -at -u ~sərq -ə**tː** -u
steal.PFV -SBJ.3SG.F -OBJ.3SG.M
‘she stole it’

c. qətl -at -u ~qətl -ə**tː** -u
kill.PFV -SBJ.3SG.F -OBJ.3SG.M
‘she killed him’

Otherwise, Moroccan Casablanca does not lengthen but there is an interaction with stress as shown in (389). There is no extrametricality in this variety, so it appears that the suffixes are rejecting stress. This will be discussed further below.

(389) Moroccan Casablanca Suffix Stress (Boudlal 2001, p. 122, 116, 339)

a. lawˈy -in
wilted -PL
‘wilted pl’

b. lawˈy -in -hum
twist.PFV -SBJ.3PL.M -OBJ.3PL
‘they are twisting them’

c. ˈkal -∅ -kum
eat.PFV -SBJ.3SG.M -OBJ.2PL
‘he ate you’

d. mərməd -ˈna -k
trail.PFV -SBJ.1PL -OBJ.2SG
‘we trailed you (in dust’)

e. daˈr -u -h
do.PFV -SBJ.3PL -OBJ.3SG.M
‘they did it’

f. gəlb -u -ha -ˈli -kum
reverse.PFV -SBJ.3PL -OBJ.3SG.F -DAT -OBJ.2PL
‘they reversed it for you’

Unlike the rest of the set explored here, Sˤanʕaːniː has long vowels underlyingly so no further vowel lengthening can occur, as shown in (390). However, consonant lengthening does occur in prepositions as shown in (391).

(390) Sˤanʕaːniː No Lengthening as Underlying Long Vowels (Watson 2002, p. 62, 144)

a. katab -∅ -la -**naː**
write.PFV -SBJ.3SG.M -DAT -OBJ.1PL
'he wrote to us'

b. maː katab -∅ -la -**naː** -ʃ
NEG write.PFV -SBJ.3SG.M -DAT -OBJ.1PL -NEG
'he didn't write to us'

c. ram**aː** -∅
throw.PFV -SBJ.3SG.M
'he threw'

(391) Sˤanʕaːniː Preposition Lengthening before Pronominal Affixes (Watson 2002, p. 205)

a. /min+ha/ → minːahaː
'from her'

b. /ʕan + iʃ/ → ʕanːiʃ
'about you (f.s.)'

In almost all of the varieties discussed here, prepositions undergo lengthening before pronominal affixes, as seen in the preposition final vowel in (392). This is often assumed to be lengthening to meet minimal word requirements

(392) Preposition Lengthening before Pronominal Affixes

a. fi
'in'

b. f**iː** -h
in -OBJ.3SG.M
'in it'

However this does not occur in Rufaidah, Rwaili or Haːyil, as shown in (393), (394) and 395 respectively.

(393) Rufaidah No Prepositional Lengthening before Pronominal Affix (Prochazka Jr. 1988, p. 219)

a. min-h
 ‘from him’

b. min-ha
 ‘from her’

c. min-ni
 ‘from me’

d. min-kum
 ‘from you’

(394) Rwaili No Prepositional Lengthening before Pronominal Affix (Prochazka Jr. 1988, p. 219)

a. mni-h
 ‘from him’

b. min-ha
 ‘from her’

c. min-hum
 ‘from me’

(395) Haːyil No Prepositional Lengthening before Pronominal Affix (Prochazka Jr. 1988, p. 219)

a. mnu-h
 ‘from him’

b. mn-ah
 ‘from her’

c. mn-ak
 ‘from you’

d. mn-iy
 ‘from me’

The extensive pattern of lengthening discussed in this section suggests that these endings prefer to attach after a stressed or stressable foot — often ideally a heavy monosyllabic foot — though varieties differ in terms of what element they are willing to lengthen to achieve this heavy foot.

4.4.1 Pre-subject lengthening

This preference for attachment after a stressed or stressable heavy foot could appear to be a clitic behaviour, or that of an aberrant affix.

However, this is not a preference solely found in these endings. It is also seen in non-3rd person subject markers in modern varieties in weak, biliteral and sound verbs. In traditional analyses of Arabic, verbs are classified in terms of the structure of their roots. The canonical root is triliteral with three consonants.

4.4.1.1 Weak Verbs

In so-called 'weak' verbs with the root C1-C2-Gl where Gl means Glide, the final glide surfaces as a long vowel or a vowel plus glide combination before the non-third person subject markers, as seen in the paradigms from Sˤanʕaːniː and Cairene in Table 4.7 below.

Table 4.7: Weak Verb Paradigms (Watson 2002, p. 144)

	Sˤanʕaːniː	Cairene
he threw	ramaː	rama
she threw	ramit	ramit
they (m.) threw	ramaw	ramu
they (f.) threw	ramayn	
you (m.s.) threw	ramayt	rameːt
you (f.s.) threw	ramaytiː	rameːti
you (m.pl.) threw	ramaytuː	rameːtu
you (f.pl.) threw	ramaytayn	
I threw	ramayt	rameːt
we threw	ramaynaː	rameːna

This length is likely the result of coalesence of the vowel and glide. This gives a heavy stressed foot before subject marker, and occurs in all the varieties explored here as shown in the Table 4.8[10]. Gaps indicate not that lengthening does not occur but rather that data was not available.

Table 4.8: Weak Verbs

Variety	3SG M Verb	Non-3SG M Verb	Non-3SG Verb Meaning
Palestinian	-	-	-
Cairene	ˈrama	raˈmeː-na	‘we threw’
Sˤanʕaːniː	ˈrama	raˈmay-na	‘we threw’
Tunisian	ˈqraː	ˈqray-t	‘I read’
Moroccan Casablanca	-	-	-
Lebanese	ˈdaʕa	daˈʕaw-tum	‘you (pl.) summoned’
Muscat	ˈbakaː	baˈkeː-na	‘we cried’
Algerian	-	-	-
Qatari	ˈdifaj	diˈfeː-t	‘you got warm’
Iraqi	ˈnisa	niˈseː-na	‘we forgot’
Wadi Ramː	-	-	-
Libyan Tripoli	-	-	-
Rufaidah	ˈnasi	niˈsiː-na	‘we forgot’
Rwaili	ˈnisi	niˈsiː-na	‘we forgot’
Haːyil	ˈnisi	niˈsiː-na	‘we forgot’
Makːan	ˈrama	raˈmeː-na	‘we threw’

4.4.1.2 Biliteral Verbs

In Classical Arabic, biliteral verbs of the root form C1-C2-C2 underwent degemination with epenthesis between the two homorganic root consonants when followed by a consonant initial suffix. However, in modern varieties this degemination no longer occurs. Instead, perhaps on analogy with weak verbs (Ferguson 1959a; Blau 1988; Glover 1988; Watson 2002), they have gained a long vowel before the subject marker where previously the geminate split into two singleton consonants separated by a vowel. Table 4.9 below compares the Classical paradigm with the modern Sˤanʕaːniː and Cairene.

[10] Data Sources: Cairene and Sˤanʕaːniː from Watson (2002, p. 144); Tunisian from Maamouri (1967, p. 83, 16); Lebanese from Haddad (1984, p. 107); Muscat from Glover (1988, p. 169); Qatari from Al-Sulaiti (1993, p. 180); Iraqi from Majdi (1988, pp. . 164); Rufaidah, Rwaili and Haːyil from Prochazka Jr. (1988, pp. 90-92); Makːan from Kabrah (2004, p. 31)

Table 4.9: Classical Arabic versus Sˤanʕaːniː and Cairene Biliteral Verb Paradigm (Watson 2002, p. 146)

	Classical	Sˤanʕaːniː	Cairene
he loved	ħabːa	ħabː	ħabː
she loved	ħabːat	ħabːat	ħabːit
they (m.) loved	ħabːuː	ħabːuː	ħabːu
they (f.) loved	ħababna	ħabːayn	
you (m.s.) loved	ħababta	ħabːayt	ħabːeːt
you (f.s.) loved	ħababti	ħabːaytiː	ħabːeːti
you (m.pl.) loved	ħababtum	ħabːaytuː	ħabːeːtu
you (f.pl.) loved	ħababtunːa	ħabːatayn	
I loved	ħababtu	ħabːayt	ħabːeːt
we loved	ħababnaː	ħabːaynaː	ħabːeːna

This is found across the sixteen varieties explored here where data is available, except for Lebanese, as shown in Table 4.10[11].

Table 4.10: Biliteral Verbs

Variety	3SG M Verb	Non-3SG M Verb	Meaning
Palestinian	-	-	-
Cairene	ħabː	ħabːeː-na	'we loved'
Sˤanʕaːniː	ħabː	ħabːay-na	'we loved'
Tunisian	-	-	-
Moroccan Casablanca	-	-	-
Lebanese	-	-	-
Muscat	ʃalː	ʃalːeː-na	'we carried'
Algerian	ʃədː	ʃədːiː-na	'we held'
Qatari	katː	katːeː-t	'you (m.s.) spilt'
Iraqi	dazː	dazːeː-tu	'you (p.) sent'
Wadi Ramː	astaʕadː	astaʕadːeː-t	'I was ready'
Libyan Tripoli	ʃədː	ʃədːi-na	'we seized'
Rufaidah	madː	madːeː-na	'we spread'
Rwaili	madː	madːeː-na	'we spread'
Haːyil	madː	madːeː-na	'we spread'
Makːan	madː	madːeː-t	'I spread'

[11] Data Sources: Cairene and Sˤanʕaːniː from Watson (2002, p. 146); Muscat from Glover (1988, p. 169); Algerian from Bouhadiba (1988, p. 280); Qatari from Al-Sulaiti (1993, p. 180); Iraqi from Erwin (1963, p. 40); Wadi Ramː from Al Mashaqba (2015, p. 168); Libyan Tripoli from Yoda (2005, p. 55); Rufaidah, Rwaili and Haːyil from Prochazka Jr. (1988, pp. 55-56); and Makːan from Kabrah (2004, p. 90)

While in weak verbs, the long vowel is plausibly the result of coalescence, there is no underlying source for this long vowel in the biliteral verbs. Notably, a short vowel is all that is necessarily required by the syllable structure, and such is seen after geminates in other situations, as shown in (396)–(397) below.

(396) Cairene Epenthesis after Geminates (Watson 2002, p. 64)

/kulː-hum/

kulːu̲ -hum
all -OBJ.3PL

'all of them'

(397) Qatari Epenthesis after Geminates (Al-Sulaiti 1993, p. 241)

/radː-ha/

radːi̲ -∅ -ha
return.PFV -SBJ.3SG.M -OBJ.3SG.F

'he returned it'

Furthermore, this insertion of a long vowel before the subject markers occurs even in varieties that permit medial tautosyllabic CVCː, therefore it is not that this vowel is an epenthetic vowel inserted to repair syllable structure violations, as shown in Examples (398)- 400.

(398) Muscat CVCː Permitted but Long Vowel Insertion in Biliteral Verbs (Glover 1988, p. 60, 176)

a. maʃaːbː -ha
fan.PL -POSS.3SG.F
'her fans'

b. ʃaːlː -ha
carry.PTCP -OBJ.3SG.F
'carrying it'

c. dalːeː -na
guide.PFV -SBJ.1PL
'we guided'

(399) Algerian CVCː Permitted but Long Vowel Insertion in Biliteral Verbs (Bouhadiba 1988, p. 296, 280)

a. ʃafː -∅ -na
inspire.pity.PFV -SBJ.3SG.M -OBJ.1PL
'he inspired pity in us'

b. ʃədːiː -na
hold.PFV -SBJ.1PL
'we held'

(400) Wadi Ram: CVCː Permitted but Long Vowel Insertion in Biliteral Verbs

(Al Mashaqba 2015, p. 109, 168)

a. mitʕalːm -iːn
educate.PTCP.PASS -PL.M
'educated (mpl.)'

b. aʃtadːeː -t
warm.PFV -SBJ.1SG
'I warmed'

Rather than a short epenthetic vowel, it is a long stressable vowel that is inserted before the endings. This ensures that there is a heavy stressed or stressable foot before subject marker affixation — just as I have shown with the other endings.

4.4.1.3 Sound Verbs

In sound/regular verbs of the root C1-C2-C3, the non 3rd person subject markers are all consonant initial. Therefore, when they attach to the CV.CVC stem, they affix to a heavy stressed foot without requiring any metrical structure changes, as shown in the paradigms for Classical Arabic, Sˤanʕaːniː, and Cairene in Table 4.11 below.

Table 4.11: Classical Arabic versus Sˤanʕaːniː and Cairene Sound Verb Paradigm

	Classical	Sˤanʕaːniː	Cairene
he wrote	kataba	katab	katab
she wrote	katabat	katabat	katabit
they (m.) wrote	katabuː	katabuː	katabu
they (f.) wrote	katab	katabayn	
you (m.s.) wrote	katabta	katabt	katabt
you (f.s.) wrote	katabti	katabtiː	katabti
you (m.pl.) wrote	katabtum	katabtuː	katabtu
you (f.pl.) wrote	katabtunːa	katabtayn	
I wrote	katabtu	katabt	katabt
we wrote	katabnaː	katabnaː	katabna

Notably, in Muscat the long vowel addition is optionally emerging in strong verbs as well, as shown in (401).

(401) Muscat Long Vowel In Strong Vowels (Glover 1988, p. 175) atab -na

katb**eː** -na
write.PFV -SBJ.1PL

‘we wrote’

(401) labːaq -ti

labːaq**eː** -ti
turn.on.PFV -SBJ.2SG.F

‘you (f.s.) turned on’

(401) qaːbal -t

qaːbl**eː** -t
see.PFV -SBJ.1SG

‘I saw’

(401) ftakar -na

ftakr**eː** -na
think.PFV -Sbj.1PL

‘we thought’

4.4.2 Conclusion: Pre-affix Lengthening

Given that all these endings either require a heavy stressed foot to be created phonologically before affixation, or move stress to this foot without altering the metrical structure as shall be discussed in the next section, we cannot claim that this is merely ‘pre-clitic lengthening’ or some aberrant exception. Rather, it is a fundamental requirement ‘baked in’ to the phonology of modern Arabic varieties. Thus, a distinction between the subject marker and others on this basis is unfounded.

4.5 Exceptions to Stress Parameters

The preference for affixes to attach to a stressed foot accounts for otherwise exceptional behaviours of stress across Arabic varieties. In this section, I will discuss cases that assign stress morphemically before the affix without changing the underlying metrical structure; feminine past tense verbs with object suffixes that appear exceptional; heteromorphemic long vowels that escape syllable extrametricality; and the unexpectedly stressed epenthetic vowel in the Palestinian [katab'ti̲lha].

4.5.1 Pre-Affix Stress Assignment

Moroccan Casablanca stresses the syllable before the non-subject affix as shown in (402)[12]. Recall from Chapter 3 that Moroccan Casablanca does not have extrametricality and the schwa does not contribute any mora. In (402-a–b), stress falls where we would expect for moraic trochees built right-to-left with End Rule Right and no extrametricality. However, in (402-c–d) we would expect stress on the final syllable, but it is attested on the penultimate syllable instead. What all four examples share is stress on the syllable before the final affix.

(402) Moroccan Casablanca Stress Assigned to Before the Final Affix (Boudlal 2001, p. 122)

[12]Note this pattern is clearest in the instrumental data that Boudlal (2001) provides. However, in the native speaker tests, there is variation on which syllable is stressed, with increasing disagreement in longer words

a. məl.ˈml -u. -na
shake.Pfv -SBJ.3PL -OBJ.1PL
'they shook us'

b. mqulˈb -in -ək
deceive.PTCP -PL -OBJ.2SG
'they are deceiving you (s.)'

c. lawˈy -in -hum
twist.PTCP -PL.M -OBJ.3PL
'they are twisting them'

d. ˈkal -kum
all -OBJ.2PL
'all of you (pl.)'

An alternative analysis is that these affixes are neutral for stress. However, the stress in affixed words falls on the same syllable as non-affixed words of the same shape, as shown in (403). This suggests they do impact stress.

(403) Moroccan Casablanca Non-Stress-Neutral Affixes (Boudlal 2001, p. 117)

a. minˈzara
'sharpener'

b. rubˈl -u -ni
disturb.PFV -SBJ.3PL -OBJ.1SG
'they disturbed me'

Furthermore, if the formative is part of a heteromorphemic heavy syllable, that syllable is stressed instead as you cannot stress only part of a syllable, as shown in (404).

(404) Moroccan Casablanca Final Heteromorphemic Heavy Syllable Stress (Boudlal 2001, p. 116-117)

a. mərməd -ˈna -k
trail.in.dust.PFV -SBJ.1PL -OBJ.2SG
'we trailed you (s.) (in dust)'

b. daˈr -u -h
do.PFV -SBJ.3PL -OBJ.3SG.M
'they did it'

Therefore, Moroccan Casablanca is displaying the same behaviour as the other dialects discussed here, but by moving stress rather than amending the underlying metrical structure.

Rwaili does lengthen vowels and geminate the 2nd person plural subject affix consonant, as shown above but repeated here as (405) and (405) for convenience.

(405) Rwaili Affix Vowel Lengthening (Prochazka Jr. 1988, p. 108, 184, 78, 188)

a. ti -ji
SBJ.2SG -come.IMPF
'you come'

b. ta -r**uː** -ni
SBJ.2SG -see.IMPF -OBJ.1SG
'you see me'

c. ʃif -ti
see.PFV -SBJ.2SG
'you saw'

d. ʃif -t**iː** -ni
see.PFV -SBJ.2SG -OBJ.1SG
'you saw me'

(406) Rwaili Consonant Lengthening (Prochazka Jr. 1988, p. 78, 188)

a. ʃif -tin
see.PFV -SBJ.2PL.F
'you (f.pl.) saw'

b. ʃif -ti**nː** -əhum
see.PFV -SBJ.2PL.F -OBJ.3PL
'you (f.pl.) saw them'

However, in other cases it does not change the metrical structure. Rwaili assigns the stress to the syllable before the final affix as shown in (407) — contrary to the syllable extrametricality exhibited elsewhere in this variety. Therefore, in the following (C)CV.CV.CVC words we would expect initial stress based on the metrical structure, but rather medial stress is attested.

(407) Rwaili Stress Move before Final Affix (Prochazka Jr. 1988, pp. 135-140)

a. ħziˈm -it -ih
tie.PFV -SBJ.3SG.F -OBJ.3SG.M
'she tied him'

b. iðˈbiħ -∅ -ih
kill.PFV -SBJ.3SG.M -OBJ.3SG.M
'he killed it'

c. ðruˈb -it -ih
hit.PFV -SBJ.3SG.F -OBJ.3SG.M
'she hit him'

d. samˈʕ -it -ih
hear.PFV -SBJ.3SG.F -OBJ.3SG.M
'she heard him'

These cases of stress assignment contrary to our metrical expectations can be compared to neighbouring varieties Haːyil in (408) and Rufaidah in (409) that do not have this behaviour on the same word as (407-d).

(408) Haːyil No Pre-Affix Stress Movement (Prochazka Jr. 1988, p. 140)
ˈsamʕ -at -uh
hear.PFV -SBJ.3SG.F -OBJ.3SG.M

'she heard him'

(409) Rufaidah No Pre-Affix Stress Movement (Prochazka Jr. 1988, p. 140)
ˈsamʕ -at -ih
hear.PFV -SBJ.3SG.F -OBJ.3SG.M

'she heard him'

This is also seen in Wadi Hadˤrami Arabic, where syllable extrametricality is found unless there is an affix as shown in (410).[13]

(410) Wadi Hadˤrami Arabic Pre-Affix Stress Movement (Al-Saqqaf 1999, p. 100-101)

a. ˈʔajnabi

'foreign'

b. ʔamˈmat -i
aunt -POSS.1SG
'my aunt'

[13] Wadi Hadˤrami Arabic was not considered in depth here as according to Bamakhramah (2010) CCC is syllabified as C.CC, unlike in the rest of the varieties explored here, as well as inconsistent epenthesis data with medial CVCː syllables in Al-Saqqaf (1999).

4.5.2 Feminine Past Tense Verbs with Pronominal Objects

A common exception to stress parameters across Arabic varieties is third person singular feminine past tense verbs with an object suffix. These tend to receive stress on the feminine ending. As such, they escape syncope affecting forms with the same CV structure (411-a)–(411-b). It is not that they are underlyingly stressed, as in the form in (411-c) without an object suffix there is initial stress rather than final stress on the [-it]. However, the [-it] feminine subject marker is stressed in both (411-d) and (411-e) when followed by an object marker. In (411-d) this is what would be expected by the metrical structure in Cairene (moraic trochees built left-to-right with End Rule Right and consonant extrametricality), but in (411-e) we would expect initial stress as depicted in Figure 4.1 rather than the attested medial stress.

(411) Cairene Stress Feminine Past Tense Affix before Pronominal Object (Watson 2011a, p. 15-16)

a. /katabit-u/

kataˈb	-it	-u
write.PFV	-SBJ.3SG.F	-OBJ.3SG.M

‘she wrote it (m)’

b. /kanakit-u/

kaˈnakt	-u
coffee.pot	-POSS.3SG.M

‘his coffee pot’

c.

ˈdˤarab	-it
beat.PFV	-SBJ.3SG.F

‘she beat’

d.

dˤaraˈb	-it	-u
beat.PFV	-SBJ.3SG.F	-OBJ.3SG.M

‘she beat him’

e.

raˈm	-it	-u
throw.PFV	-SBJ.3SG.F	-OBJ.3SG.M

‘she threw it (m.)’

Figure 4.1: Cairene Expected but Unattested Stress on [ramitu]

X		
(x	.)	
ra	mi	tu

Similarly, in Lebanese we see that while we expect initial stress in CV.CV.CV words such as (412-a), where the feminine subject marker ending [-it] is followed by an affix in (412-b) medial stress is attested.

(412) Lebanese Stress Feminine Past Tense Affix before Pronominal Object (Haddad 1984, p. 123)

a. ˈsimʕ -it -u
reputation -POSS.3SG.M
'his reputation'

b. ʕimˈl -it -u
make.PFV -SBJ.3SG.F -OBJ.3SG.M
'she made him'

Libyan Tripoli does syncopate the feminine subject vowel unlike other dialects as shown in (413), though this alternates with gemination of the affix consonant as shown in (414). Both processes have the impact of ensuring there is a heavy stressable syllable before the non-nominative affix.

(413) Libyan Tripoli Syncopation of Feminine Subject Marker Vowel (Yoda 2005, p. 107)

a. /dˤərb-əʧ-u/
ˈdˤərb -ʧ -u
hit.PFV -SBJ.3SG.F -OBJ.3SG.M
'she hit him'

b. /zar-eʧ-u/
ˈzar -ʧ -u
visit.PFV -SBJ.3SG.F -OBJ.3SG.M
'she visited him'

(414) Libyan Tripoli Gemination of Feminine Subject Marker Final Consonant (Yoda 2005, p. 107)

a. /dˁərb-ətʃ-u/
dˁ ərˈb -ətʃː -u
hit.PFV -SBJ.3SG.F -OBJ.3SG.M
'she hit him'

b. /fərʃk-ətʃ-u/
fərʃˈk -ətʃː -u
relax.PFV -SBJ.3SG.F -OBJ.3SG.M
'she relaxed him'

This final gemination is optional in nouns before the possessive, as shown in (415).

(415) Libyan Tripoli Optional Syncopation or Final Gemination (Yoda 2005, p. 108)

zar -ətʃ əmː -a ~

zar -ətʃː əmː -a ~

zar -tʃ əmː -a
visit.PFV -SBJ.3SG.F mother -POSS.3SG.F

'she visited her mother'

However, syncope in Libyan Tripoli is blocked before the dative, as shown in (416).

(416) Libyan Tripoli Syncope Blocked before the Dative Affix (Yoda 2005, p. 103)

kətʃb -ətʃ -ˈəl -kəm
write.PFV -SBJ.3SG.F -DAT -OBJ.2PL

'she wrote to you (pl.)'

Makːan assigns stress regardless of metrical structure in nouns and verbs before the non-nominative affix whether functioning as an object or possessive. That this is different to cases without this affix is clear in (417-c) and (417-d) where the near minimal pair has different stress because of the non-nominative affix.

(417) Makːan Stress Assignment before Final Affix (Kabrah 2004, p. 123, 60–61)

a. /ʃiriˈb-at-u/
ʃirˈb -at -u
drink.PFV -SBJ.3SG.F -OBJ.3SG.M
'she drank it'

b. /sˁaːħiˈbat-u/

ʃˤaːħˈbat -u
friend.F -Poss.3SG.M
'his girl friend'

c. ˈgatal -u
kill.PFV -SBJ.3PL
'they killed'

d. gaˈl -at -u
fry.PFV -SBJ.3SG.F -OBJ.M.SG
'she fried it'

Note this also happens in Benghazi Arabic (Benkato 2014) and Andalusi Arabic (Corriente 2008, p. 104).

McCarthy (1979) proposes a branching node over the feminine ending /-it/ and following material associated with primary stress; Angoujard (1981) suggests /-it/ has an indestructible rhyme; Haddad (1984) views it as a lexical rule. As Broselow (1976) notes, it is not that /-it/ is stressed in the lexicon as without the pronominal object it is not stressed. However, she does not think it is triggered by the object affix due to cases like [ˈdarabu], 'he beat him', where stress is not moved to the syllable before the affix. Furthermore, as noted by Majdi (1988), a branching node over /-it/ doesn't work for longer strings, as shown in (418).

(418) Iraqi Stress Not on [-it] with Longer Strings (Majdi 1988, p. 68)
kitab -it -ˈhaː -l -ak
write.PFV -SBJ.1SG -OBJ.3SG.F -DAT -OBJ.2SG
'I wrote it for you (s.)'

Watson (2002, p. 97) claims there is an exceptional reversal of parsing directionality in these cases. Thus for [raˈmitu], while we normally expect left-to-right building of moraic trochees, in these cases she argues that we get right-to-left building of moraic trochees: [ra(ˈmitu)].

However, I argue that such exceptional architecture is not needed to account for this phenomenon. Rather, it follows from the same preference to have a stressed or stressable foot before the object suffix. That [ˈdarabu] is not stressed medially as [*daˈrabu] does not pose a problem, as there is still a stressed foot before the object — and furthermore, 3rd person forms are more frequent than 1st and 2nd person forms so likely to interact with requirements slightly differently.

The range of approaches taken in different varieties is summarised in Table 4.12 below. In some of the other varieties I explore here, the feminine ending is geminated, which has the effect of creating a heavy syllable that receives stress and removing the environment for syncope — though there are some varieties that do not do either of these, such as Palestinianin (419).

(419) Palestinian No Gemination or Stress Shift (Abu-Salim 1982a, p. 93, 60)

a. ˈsaʔl -at
ask.PFV -SBJ.3SG.F
'she asked'

b. ˈsaʔl -at -o
ask.PFV -SBJ.3SG.F -OBJ.3SG.M
'she asked him'

c. raːˈsal -at -o
correspond.PFV -SBJ.3SG.F -OBJ.3SG.M
'she corresponded with him'

Table 4.12: Strategies to Create Stressed Syllables Before Affixes

Variety	Affix Vowel/Consonant Lengthening	Itu Stress	Pre-Affix Stress
Algerian	Y		
Tunisian	Y		
Libyan Tripoli	Y		
Palestinian	Y		
Iraqi	Y		
Wadi Ramː	Y		
Muscat	Y		
Qatari	Y		
Rufaidah	Y		
Haːyil	Y		
Cairene	Y	Y	
Makːan	Y	Y	
Lebanese	Y	Y	
Moroccan Casablanca	Y		Y
Rwaili	Y		Y
Sˁanʕaːniː			Y

Table 4.12 demonstrates that while there is a range of strategies taken, all of the varieties discussed here undertake at least one process to affix next to a stressed foot. There is not a particularly clear geographical distribution of these strategies. Three of the four North African varieties (Algerian, Tunisian, Libyan Tripoli but not Moroccan Casablanca) only do affix vowel/consonant lengthening and as such appear to have metrically expected stress. Cairene, Makːan, and Lebanese also stress the feminine suffix in addition to this lengthening, but do not have a particularly clear geographical motivation for this grouping given that neighbouring varieites such as Palestinian, Wadi Ramː, and Rufaidah only do pre-affix lengthening. Similarly, the grouping of Moroccan Casablanca, Rwaili and Sˁanʕaːniː for pre-affix stress assignment regardless of metrical structure does not have clear synchronic geographical motivations. It may be that there are historical patterns of invasion and colonisation that can account for these, but this is beyond the scope of this synchronic thesis.

4.5.3 Final Stressed Long Vowels

Another common exception to stress parameters in Arabic is to find verbs with stress on a final long vowel that is not impacted by final syllable extrametricality.

As demonstrated in the previous chapter, Palestinian displays final syllable extrametricality. However contrary to this there is final stress when the final long vowel is heteromorphemic, as it is in (420-b).

(420) Palestinian Stress on Final Heteromorphemic Long Vowel (Abu-Salim 1982a, p. 136)

a. ˈkursi
   ‘chair’

b. kurˈsi -i
   chair -POSS.3SG.M
   ’his chair’

Similarly, in Makːan final syllable extrametricality fails to occur when the final long vowel is heteromorphemic, as it is in (421-b), (421-d) and (421-f).

(421) Makːan Stress on Final Heteromorphemic Long Vowel (Kabrah 2004, p. 103, 27)

a. ˈnaːdi
   ‘club’

b. naːˈdi -i
   club -POSS.3SG.M
   ‘his club’

c. ˈwaːdi
   ‘valley’

d. waːˈdi -i
   valley -POSS.3SG.M
   ‘his valley’

e. kaˈtab -na
   write.PFV -SBJ.1PL
   ‘we wrote’

f. katab -ˈna -a
   write.PFV -SBJ.1PL -OBJ.3SG.M
   ‘we wrote it’

Iraqi also has final syllable extrametricality that fails to make final long syllables extrametricality when they are heteromorphemic as in (422).

(422) Iraqi Stress on Final Heteromorphemic Long Vowel (Kabrah 2004, p. 27)

a. kaˈtab -na
   write.PFV -SBJ.1PL
   ‘we wrote’

b. katab -ˈna -a
   write.PFV -SBJ.1PL -OBJ.3SG.M
   ‘we wrote it (m.)’

This length alternation occurs with both nouns and verbs in Iraqi as shown in (423).

(423) Iraqi Final Heteromorphemic Long Vowels (Erwin 1963, p. 273)

a. sira
   ‘turn’

b. sira -a
   turn -POSS.3SG.M

‘his turn’

c. yi -nsa
SBJ.3 -forget.IMPF
‘he forgets’

d. yi -nsa -a
SBJ.3 -forget.IMPF -OBJ.3SG.M
‘he forgets him’

Cairene has consonant extrametricality, so it is not surprising that final long vowels receive stress in (424-b). While in Cairene this is not ‘exceptional’ behaviour, it reflects the same metrical attachment preference even though it is also ratified by the regular phonology.

(424) Cairene Stress on Final Heteromorphemic Long Vowel (Watson 2002, p. 80)

a. ˈram -u
throw.PFV -SBJ.3PL
‘they threw’

b. raˈm -u -u
throw.PFV -SBJ.3PL -OBJ.3SG.M
‘they threw it’

There are no long vowels in Libyan Tripoli, so this pattern appears as stress shift rather than vowel lengthening per se, as shown in (425).

(425) Libyan Tripoli Stress on Final Heteromorphemic Vowel (Yoda 2005, p. 116)

a. ˈstənːa

‘wait(m.)!’

b. stənˈna

‘wait (m.) for him!’

c. ˈkərsi

‘chair’

d. kərˈsi

‘his chair’

e. ˈdərbu
‘they hit’

f. dərˈbu
‘they hit him’

Brame (1971) and McCarthy (1979) account for this as that all final stressed CVː are underlying CVː(h) where the /h/ is deleted word finally. This /h/ can be seen in the Cairene case [ramaˈhulha], ‘he threw it to her’. In many modern Arabic varieties, there is still an /h/ that surfaces word finally for the 3rd person masculine singular object, as shown in Table 4.13[14].

Table 4.13: Final /h/ in 3rd Person Masculine Singular Object

Variety	Word	Gloss	Translation
Sˤanʕaːniː	gul-t-a̲l-ih	say.PFV -SBJ.1SG -DAT -OBJ.3SG.M	‘I said to him’
Tunisian	smaː -h	sky -POSS.3SG.M	‘his sky’
Algerian	sabag -uː -h	overtake.PFV -SBJ.3PL -OBJ.3SG.M	‘they overtook him’
Rufaidah	liħig -tiː -h	catch.up.PFV -SBJ.2SG -OBJ.3SG.M	‘you (s.) caught up with him’
Rwaili	ʃif -tiː -h	see.PFV -SBJ.2SG -OBJ.3SG.M	‘you (s.) saw him’
Haːyil	ʃuf -tiː -h	see.PFV -SBJ.2SG -OBJ.3SG.M	‘you (s.) saw him’

Muscat varies whether an /h/ surfaces, as shown in (426) where it is found with the third person singular masculine object in (426-a) but not in (426-b).

(426) Muscat /h/ Variation (Glover 1988, p. 211, 30)

a. katab -∅ -l -oh
write.PFV -SBJ.3SG.M -DAT -OBJ.3SG.M
‘he wrote to him’

b. ðrib -∅ -a
hit.PFV -SBJ.3SG.M -OBJ.3SG.M
‘he hit him’

[14] Data Sources: Sˤanʕaːniː from Watson (2002, p. 103); Tunisian from Maamouri (1967, p. 157); Algerian from Bouhadiba (1988, p. 227); Rufaidah, Rwaili and Haːyil from Prochazka Jr. (1988, p. 185)

However, even if a final /h/ is part of the historical formative, and seen in some modern varieties, it does not entail that it is present underlyingly in the synchronic grammars of speakers where this doesn't surface — particularly where there is no evidence of this /h/ word medially, as pointed out by Majdi (1988, p. 65) for Iraqi.

Rather, it follows from the preference for attaching to a stressed or stressable foot that is shared by all the affixes in this chapter. Since stress must impact the entire syllable, not just part of it, when the affix attaches and is incorporated into the syllable it becomes part of that stressed syllable, rather than pushing stress back further. This is the same behaviour that I demonstrated earlier with Moroccan Casablanca where affixes appear to attract stress to the syllable before them as shown in (427-a)–(427-c), but if the final syllable is heteromorphemic this is stressed as in (427-d).

(427) Moroccan Casablanca Affix Stress Behaviour (Boudlal 2001, p. 115, 319, 117)

a. daˈr -u -h
do.PFV -SBJ.3PL -OBJ.3SG.M
'they did it'

b. ˈdar -u
house -POSS.3SG.M
'his house'

c. banˈya -ha
build.PTCP.F -OBJ.3SG.F
'she built it'

d. mərməd -ˈna -k
trail.in.dust.PFV -SBJ.1PL -OBJ.2SG
'we trailed you (in dust)'

4.5.4 Unexpectedly Stressed Epenthetic Vowels

Epenthetic vowels in Palestinian do not receive stress, which leads to cases of surface opacity in stress assignment, as seen in the following examples where the vowel [i] is underlying in (428-a) and (428-c) so interacts with the building of metrical structures, whereas in (428-b) and (428-d) it is epenthetic and does not impact stress or metrical structure. This is because epenthesis occurs postlexically after the assignment of stress. The derivation of (428-a–b) is shown in Figure 4.2 below.

(428) Palestinian Opaque Stress with Unstressed Epenthetic Vowels (Brame 1974, p. 41)

a. ˈkatab -it
write.PFV -SBJ.3SG.F
'she wrote'

b. kaˈtab -i̱t
write.PFV -SBJ.1SG
'I wrote'

c. ʔaˈbil -na
accept.PFV -SBJ.1PL
'we accepted'

d. ˈʔabi̱l -na
before -OBJ.1PL
'before us'

Figure 4.2: Palestinian Opaque Stress Derivation

UR Level 1	katabit	katabt
Stress	ˈkatabit	kaˈtabt
Postlexical Level		
Epentheis	ˈkatabit	kaˈtabi̱t
SR	ˈkatabit 'she wrote'	kaˈtabi̱t 'I wrote'

However, where the epenthetic vowel is inserted between four consonants rather than three, it does receive stress, as shown in (429).

(429) Palestinian CC_CC Epenthetic Vowel Stressed (Abu-Salim 1982a, p. 216)

ka.tab. -ˈt -i̱l. -ha
write.PFV -SBJ.1SG -DAT -OBJ.3SG.F

'I wrote to her'

This is not a lexical matter, as in the same verb but with the 3rd person subject rather than first, the epenthetic vowel is unstressed, as shown in (430).

(430) Palestinian C_CC Epenthetic Vowel Unstressed (Abu-Salim 1982a, p. 216)

kaˈtab	-∅	-i̲l	-na
write.PFV	-SBJ.3SG.M	-DAT	-OBJ.1PL

'he wrote to us'

Existing studies of Palestinian have struggled to account for this. Abu-Salim (1982a) suggests that this could be 'vowel insertion' occurring at the point of syllabification earlier in the derivation. However, this means that there would be two versions of epenthesis in the one language, and it is unclear why it would be inserted to syllabify with four consonants rather than three.

In light of the strong preference to attach after a stressed or stressable foot motivated in this chapter, I suggest it could reflect this preference. However, given that Palestinian does not include other 'exceptional' stress cases such as stressing the feminine subject before an object discussed above, this may be a weaker requirement in Palestinian than other varieties. An alternative possibility is that it is a remnant of historical morphological structure, where there was an underlyingly vowel present when the 1st person singular subject in the past tense was /-tu/. Which option is more likely requires further historical investigation.

4.5.5 Summary

Whilst non-subject formatives may not trigger vowel shortening and raising, it is clear that they do participate in other morphophonological alternations. This provides further evidence that they are in fact affixes not clitics.

4.6 Degree of Incorporation into the Prosodic Word

I have demonstrated above that these endings do not in their nature appear to be clitics, but rather are affixes. Further confirmation of this is the nature of prosodic output — that is, how these endings are incorporated into their phonological host.

Selkirk (1978) proposed a prosodic hierarchy, defining the relationship between different domains of phonological structure, which is depicted in Figure 4.3 below.

Figure 4.3: Prosodic Hierarchy of Selkirk (1978)

Utterance (U)
|
Intonational Phrase (I)
|
Phonological Phrase (ϕ)
|
Prosodic/Phonological Word (ω)
|
Foot (F)
|
Syllable (σ)

For Nespor and Vogel (1986), each level of the hierarchy immediately and fully dominates the constituent below it without recursion, which they expressed as the Strict Layering Hypothesis. As such, they argued that clitics were not incorporated into the phonological word, but formed a new domain between the phonological word and phrase: that of the clitic group. This is necessary as the domain for certain processes such as Greek Stress Adjustment (Nespor and Vogel 1986, p. 151) as seen in 332 on page 278 where there can be no more than two unstressed syllables after primary stress, so if this is violated by encliticisation, secondary stress must be added. This domain is further supported by Hayes (1989, pp. 209–10), who points to English, where s-palatalization and v-deletion occur between lexical words and function words. In (431), s-palatalisation occurs in (431-a) between the clitic possessor and the noun, but not in (431-b). Similarly, v-deletion occurs in (432-a) between the verb and the clitic pronominal object, but not in (432-b)[15].

(431) English s-palatalisation (Lahiri and Plank to appear, p. 35)

a. I could only see (his shadow)

[15]Note that this is likely to be a gradient rather than categorical effect (Shattuck-Hufnagel and Turk 1996).

b. I could only see John's shadow

(432) English v-deletion (Lahiri and Plank to appear, p. 35)

a. Please (lea~~ve~~ me) alone

b. Please leave Molly alone

However, this has been disputed on several grounds. Perhaps most crucially, many of the rules that appear to have the clitic group as domain can be better analysed within the Phonological Phrase or Prosodic Word plus an adjunct (Zec and Inkelas 1991; Booij 1995; Peperkamp 1997). Furthermore, the Strict Layering Hypothesis implies that clitics must be prosodic words in their own right. This does not hold for all cases, such as clitics in Dutch discussed in Booij (1996). In addition, many hold that prosodic words can be recursive prosodic constituents (Gussenhoven 1986b; Berendsen 1986; Booij 1995)[16].

Assuming that there is no need to posit the clitic group as an intermediate level of representation, Selkirk (1995) argued that there is a small set of possible levels of clitic attachment to independent phonological words.

[16] See Peperkamp (1997) for review of the clitic group

Figure 4.4: Selkirk (1995): Clitic Typology

Figure 4.5: PWord Clitic

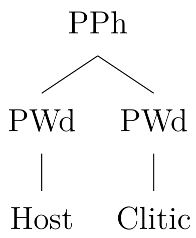


Figure 4.6: Free Clitic

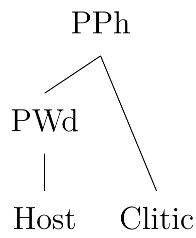


Figure 4.7: Affixal Clitic

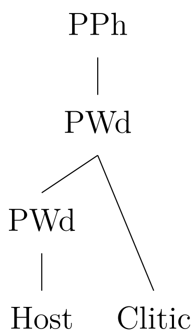
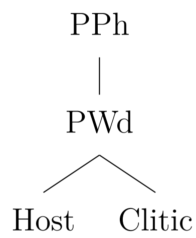


Figure 4.8: Internal Clitic



These options are illustrated by three Italian dialects (Peperkamp 1997, 176f.). In Standard Italian, clitics attach to the phrase, outside of the stress domain of the host so can be considered a free clitic as in Figure 4.6.

(433) Standard Italian

a. ˈporta
'bring'

b. ˈporta=mi
'bring me'

c. ˈporta=me=lo
'bring it to me'

In Neapolitan, they do receive stress but the host also retains stress so can be considered an affixal clitic as in Figure 4.7

(434) Neapolitan Italian

a. ˈconta
'tell'

b. ˈconta=lə
'tell it'

c. ˈconta=ˈti=lle
'tell you it'

In Lucanian, they receive stress but the host does not, so can be considered an internal clitic as in Figure 4.8. Note that the gemination is a result of radoppiamento fonosintattico, so it is not that the clitics are stressed because of the nature of their attachment here, but rather because because they constitute a large enough foot to receive stress through standard metrical processes.

(435) Lucanian Italian

a. ˈvinnə
'sell'

b. vənˈnil=lə
'sell it'

c. vinnə=ˈmi=lle
'sell me it'

Whether this is truly a matter of different types of attachment has been disputed. Kenstowicz (1991) and Bafile (1994) argue that the Neapolitan data in (434) is a matter of cyclical stress assignment with clitics attached postlexically after stress. Similarly, Loporcaro (2000) agrees that differences in stress assignment correspond to whether postlexical stress reassignment is possible. However, these accounts are not necessarily contradictory, and may correspond to a distinction between representation and process.

If we do take the output to be a prosodic word regardless of whether a clitic or affix is attached, there should be a fairly small array of differences that can distinguish between a clitic and an affix. Crucially, affixes can attach to stems to form prosodic words, whereas clitics can only attach after its host has undergone any processes needed to become a prosodic word, and affixes cannot be added after clitics. Recall the principle of Lexical Phonology that affixes can be attached cyclically, so there can be multiple levels of phonological processes occurring.

Affixes and clitics can function broadly the same way within a prosodic word. As such, if the formative in question appears fully incorporated into the prosodic word, behaving as expected in terms of syllable structure, segmental processes and stress, it does not allow us to diagnose affix or clitic status either way. However, if a formative is behaving differently, it does suggest it is a clitic. In the following, I will explore the behaviour of these formatives with regards to the prosodic word. I demonstrate that the behaviour is not definitive as to their status, rather than clearly clitic behaviours as has been suggested in the literature.

4.6.1 Arabic Prosodic Word Formation

A key piece of evidence for the differences between subject markers and the non–subject formatives is the different application of syncope and epenthesis in Levantine varieties. This has been analysed by Brame (1974) and later scholars as the non-subject endings attaching on Level 2, and that these are enclitics. However, I dispute that attachment on Level 2 entails clitic status. Furthermore, I demonstrate that these Level 2 processes are part of prosodic word formation processes, rather than 'clitic group' level processes.

The following Palestinian set appear quite opaque on the surface. The forms in (436-a) and (436-b) have the same segmental structure but different stress, and the forms in (436-c) have the same structure but different meaning, and the forms in (436-d) have different segmental structures but the same stressed syllable.

(436) Palestinian Opaque Stress Assignment (Brame 1974, p. 41)

a. ˈkatabit 'She wrote'
ka'tabit 'I wrote'

b. ʔˈbilna 'we accepted'
ˈʔabilna 'before us'

c. tˤaˈlabna ‘we ordered’

tˤaˈlabna ‘he ordered us’

d. fˈhimna ‘we understood’

fiˈhimna ‘ he understood us’

From Brame (1974, pp. 41–43)

This can be accounted for as matter of lexical phonology rule ordering (Brame 1974). If we assume that some affixes are ordered before others and undergo phonological processes before others are attached, this pattern becomes clear. In the original Brame analysis, he orders stress assignment before syncope on Level 2, and requires some form of tier conflation to remove the secondary stress as this is not attested in Palestinian Arabic. However, following Kiparsky (2000), if stress is ordered after syncope on Level 2 this is no longer needed, as shown in the figure below.

Figure 4.9: Rule Ordering Solves Opacity

/fihim+na/	/fihim/	
fiˈhim+na	ˈfihim	Level 1 STRESS
		Level 2
fˈhim+na	—	SYNCOPE
fˈhim+na	fiˈhim+na	STRESS
fˈhimna	fiˈhimna	
‘we understood’	‘he understood us’	

Both CSS and these cases of opacity demonstrate that there are differences between the subject marker and other endings. However, this does not mean that the other endings are clitics. In the rule ordering above, the subject is attached on Level 1, so triggers any paradigm effects at this level, whereas the rest are attached on Level 2.

It is not that Level 2 is the domain of ‘clitic group’ processes, as might be argued in the spirit of Nespor and Vogel (1986). Rather, words with only the subject marker also undergo processes on Level 2, which means it must be part of phonological word formation. As we see here, it is vital that the words with only the subject marker undergo Level 2 processes in order for this opacity to occur. Furthermore, Level 2 in other languages is not necessarily the domain of clitics. In English, other affixes do attach on Level 2, and tend to be Germanic rather than Romance in origin, and stress neutral.

While stress and syncope opacity is mostly found in Levantine varieties, there is evidence for these endings being incorporated into the prosodic word across Arabic varieties. I will now review the evidence for their impact of stress, syllable structure, and assimilation.

4.6.2 Impact on Stress

These formatives are counted for stress assignment and syllable structure. In the examples below, the addition of these endings moves stress in accordance with the stress parameters discussed in the previous chapter. While apparent ‘exceptions’ to stress have been discussed above, here I explore that cases that do not display exceptional behaviour. Particularly in Levantine varieties there is evidence of lexical levels influencing patterns of stress assignment (cf. Brame (1974) inter alia). However, the details of different processes occurring on which lexical level and impact on cyclical stress assignment is beyond the scope of this particular discussion.

In Palestinian the addition of the object suffix alone in (437-b) is insufficient to move stress as it does not contribute a further rightward foot. However, once the negation suffix is added as well in (437-c), there is enough material for a second foot, so stress moves in accordance with End Rule Right.

(437) Palestinian Suffixes Affect Stress (Abu-Salim 1982a, p. 10, 150)

a. ˈʃaːf -∅
see.PFV -SBJ.3SG.M
‘he saw’

b. ˈʃaːf -∅ -ni
see.PFV -SBJ.3SG.M -OBJ.1SG
‘ he saw me’

c. ʃaːf -∅ -ˈniː -ʃ
see.PFV -SBJ.3SG.M -OBJ.1SG -NEG
'he didn't see me'

The same process can be seen in Cairene in (438), where both the object and negation suffixes are required to form a foot to attract stress.

(438) Cairene Suffixes Affect Stress (Watson 2002, p. 102)

a. ˈkatab -∅
write.PFV -SBJ.3SG.M
'he wrote'

b. kaˈtab -∅ -ha
write.PFV -SBJ.3SG.M -OBJ.3SG.F
'he wrote it'

c. ma katab -∅ -ˈhaː -ʃ
NEG write.PFV -SBJ.3SG.M -OBJ.3SG.F -NEG
'he didn't write it'

Likewise in Algerian the object suffix forms a foot with the subject suffix so can receive stress in (439-b).

(439) Algerian Suffixes Affect Stress (Bouhadiba 1988, p. 227)

a. saˈbag -na
overtake.PFV -SBJ.1PL
'we overtook'

b. sabag -ˈnaː -h
overtake.PFV -SBJ.1PL -OBJ.3SG.M
'we overtook him'

In Sˤanʕaːniː, the addition of the object suffix in (440-b) allows a CVC foot in the stem to be visible to stress, so this receives stress while the object suffix is extrametrical through syllable extrametricality.

(440) Sˤanʕaːniː Suffixes Affect Stress (Watson 2002, p. 103)

a. ˈdˤarab -∅
hit.PFV -SBJ.3SG.M
'he hit'

b. dˤaˈrab -∅ -kum
hit.PFV -SBJ.3SG.M -OBJ.2PL

‘he hit you’

The same process occurs in Lebanese with the object suffix and Qatari with the possessive suffix in (441) and (442) respectively,

(441) Lebanese Suffixes Affect Stress (Haddad 1984, p. 56)

a. ˈji -ktib
SBJ.3 -write.IMPF
‘he writes’

b. ji -kˈtib -ha
SBJ.3 -write.IMPF -OBJ.3SG.M
‘he writes it’

(442) Qatari Suffixes Affect Stress (Al-Sulaiti 1993, p. 233)

a. ˈgalam

‘pen’

b. gaˈlam -ha
pen -POSS.3SG.F
‘her pen’

In Iraqi, not only does the addition of the object suffix in (443-b) combined with pre-suffix lengthening allow the stress to move rightwards, the further addition of the dative and object in (443-c) allows this process to move stress further again.

(443) Iraqi Suffixes Affect Stress (Majdi 1988, p. 60)

a. kiˈtab -na
write.PFV -SBJ.1PL
‘we wrote’

b. kitab -ˈnaː -ha
write.PFV -SBJ.1PL -OBJ.3SG.F
‘we wrote it’

c. kitab -naː -ˈhaː -l -ak
write.PFV -SBJ.1PL -OBJ.3SG.F -DAT -OBJ.2SG
‘we wrote it for you’

Similarly, pre-suffix lengthening allows stress to move in Makːan in (444-b) after addition of the object suffix.

(444) Makːan Suffixes Affect Stress (Majdi 1988, p. 27)

a. ˈkatab -u
write.PFV -SBJ.3PL
'they wrote'

b. kataˈb -uː -ha
write.PFV -SBJ.3PL -OBJ.3SG.F
'they wrote it'

This stress behaviour is the same as the Modern Greek clitic pattern given in 332 on page 278. Thus, it cannot tell us definitively whether these formatives are certainly affixes or certainly clitics, but it does tell us that they are incorporated into the prosodic word, which is consistent with the arguments for affixhood so far.

4.6.3 Impact of Syllable Structure

That the attachment of these formatives triggers epenthesis is central to previous syllable structure typologies of Arabic that focus on where an epenthetic vowel is inserted into a tri-consonantal cluster.

In Qatari, the object triggers syncope and epenthesis in the stem, as shown in Table 4.14 below. All bar the third person singular feminine, second person plural and third person masculine plural trigger syncope in the initial syllable, with epenthesis of [i] occurring for the first person objects.

Table 4.14: Qatari Arabic Objects Trigger Syncope in the Stem (Al-Sulaiti 1993, p. 30)

ðirab-ni	'he hit me'
ðrib-ak	'he hit you (s.)'
ðrib-a	'he hit him'
ðarab-ha	'he hit her'
ðirab-na	'he hit us'
ðarab-kum	'he hit you (pl.)'
ðarab-hum	'he hit them (m.)'
ðrib-atʃ	'he hit them (f.)'

This interaction with syncope and epenthesis is also seen with possessives in Moroccan Casablanca as shown in (445), where the addition of a consonant initial suffix does not trigger changes in the stem in (445-b), but a vowel initial suffix does in (445-c) as it changes the syllabification.

(445) Moroccan Casablanca Stem Syncope and Epenthesis (Boudlal 2001, p. 164)

a. ktəf
‘shoulder’

b. ktəf -na
shoulder -POSS.1PL
‘our shoulder’

c. kətf -u
shoulder -POSS.3SG.M
‘his shoulder’

The derivation of (445-c) is as follows: the affixation of vowel initial suffix in (445-c) triggers resyllabification as [ktə.fu]. This produces a schwa in an open syllable, which is not permitted so it is deleted, creating [ktfu]. This initial cluster is not permitted, so epenthesis occurs to produce the observed surface form [kətfu]

More interesting is that negation displays some variation in incorporation into syllable structure, but always affects stress. In Cairene in (446), negation requires epenthesis after a geminate in (446-b).

(446) Cairene Negation Incorporation into Syllable Structure (Watson 2002, p. 123, 183)

a. ma ʕamal -∅ -ʃ
NEG make.PFV -SBJ.3SG.M -NEG
‘he didn’t do/make’

b. ji -ħibː -i̲ʃ
SBJ.3 -love.IMPF -NEG
‘he doesn’t love’

In Sˤanʕaːniː, there is variation in whether the negation affix triggers Closed Syllable Shortening, as shown in (447) where the optionality is marked with brackets. Whether this reflects variation between speakers or within speakers is unclear based on the available information in Watson (2002).

(447) Sˤanʕaːniː Variation in CSS with Negation (Watson 2002, p. 67)

a. ma ka(ː)n -∅ -ʃ
NEG be.PFV -.SBJ.3SG.M -NEG
'he was not'

b. ma diri(ː) -t -ʃ
NEG know.PFV -SBJ.1SG -NEG
'I didn't know'

In Libyan Tripoli there is variation in whether the negation triggers epenthesis, as shown in (448), where the optional epenthesis is marked with brackets.

(448) Libyan Tripoli Variation in Epenthesis with Negation (Yoda 2005, p. 96)

a. ma qəl -t -(ə̲)ʃ
NEG say.PFV -SBJ.1SG -NEG
'I did not say'

b. ma qal -∅ -(ə̲)ʃ
NEG say.PFV -SBJ.3SG.M -NEG
'he did not say'

c. ma ʒab -∅ -(ə̲)ʃ
NEG bring.PFV -SBJ.3SG.M -NEG
'he did not bring'

That the negation displays this variation in whether it triggers repairs to syllable structure violations might suggest that it is not fully incorporated into the prosodic word — indicating perhaps clitic status. However, that this final coronal strident varies in whether it triggers violates syllable structure or triggers epenthesis to repair this is not unusual cross-linguistically — it is rather similar to final English sibilants that are permitted to exceed usual syllable restrictions. Therefore, this does not preclude affix status.

4.6.4 Qatari /h/ Assimilation

In Qatari the /h/ in these endings is often assimilated when following a voiceless obstruent. This doesn't apply in slow/careful speech, and only affects clitics. This assimilation is shown in (449).

(449) Qatari H Assimilation (Al-Sulaiti 1993, p. 155, 237)

a. /tiras-ha/ → tirasːa

'he filled it up'

b. /gmaːʃ-hum/ → gmaːʃːum

'their pearls'

It doesn't apply word-internally, as shown in (450).

(450) Qatari No Word-Internal H Assimilation (Al-Sulaiti 1993, p. 237)

a. /jifham/ → *jifːam

'he understands'

b. /mathuːm/ → *matːuːm

'accused'

It doesn't apply across words, as shown in (451).

(451) Qatari no H Assimilation Across Words (Al-Sulaiti 1993, p. 238)

a. /riːʃ hudhud/ → * riːʃ ʃudhud

'feathers of a kingfisher'

b. /zixː haːðiːhi/ → *zixː xaːðiːhi

'touch this'

This forms part of Al-Sulaiti's (1993, pp. 237-239) evidence that these form clitics. However, this pattern does not mean that there is a separate clitic group. Rather, we would not expect assimilation to occur stem internally, as it tends to occur at formative boundaries not within formatives. In Arabic in particular, it is important that such complete assimilation not occur inside verb stems, as each consonant contributes to meaning[17]; and furthermore, there are restrictions on homorganic consonants in verbs in Arabic and other Semitic languages (Greenberg 1950; McCarthy 1986; McCarthy 1994; Pierrehumbert 1993; Buckley 1997; Frisch, Broe, et al. 1997; Berent and Shimron 2003).

Note this assimilation also occurs in Jordanian Arabic as shown in (452).

[17]Though regressive voicing assimilation does occur in Palestinian such as ʔakbar → ʔagbar (Abu-Salim 1987)

(452) Jordanian Arabic (Haddad 1984, p. 161)

a. /madrasit-ha/ → madraˈsitːa
‘her school’

b. /madrasit-ak/ → madˈrastak
‘your school’

c. /xaːdm-it-ha/ → xaːdˈmiːtːa
‘she served her’

d. /xaːdm-it-na/ → xaːdˈmiːtna
‘she served us’

4.6.5 Conclusion: Incorporation into the Prosodic Word at Level 2

Throughout this chapter I have not disputed that there is a difference between the subject marker and the other formatives under discussion. However, I argue that it is not a matter of affix versus clitic; but rather, they are all affixes, with the non-subject formatives added at Level 2 rather than Level 1. The subject and plural affixes are attached at Level 1, with the dative, non-subject pronominal (functioning as direct object, indirect object, or possessive depending on context), and negation attaching at Level 2.

I have demonstrated that under the Zwicky-Pullum criteria, these formatives appear much closer to affixes than clitics. In this section, I have shown that the phonological differences can be accounted for by attachment at Level 2 — and that this level must still be the domain of phonological word processes rather than clitic group processes, as words without these non-subject formatives still go through the processes on this level. I show that these formatives are fully incorporated into the stress and syllable structure requirements of all the varieties; and the only cases that are less tightly incorporated — the negation marker, which varies in whether it triggers epenthesis — are not dissimilar to English sibilants that also permit violations of usual syllable restrictions. Furthermore, I argue that processes such as assimilation in Qatari that appear to be a clitic group process are still a prosodic word level process.

4.7 Do We Need a Clitic versus Affix Distinction in Arabic?

In the sections above, I have argued that this set of endings are not clitics as suggested in the literature but in fact affixes that are merely incorporated at Level 2. Thus, we must ask to what extent is a clitic/affix distinction even needed for Arabic?

Whilst synchronically the endings discussed here are all suffixes, there are clitics found in some varieties. Although stress cannot be in general the main diagnostic for clitic status cross-linguistically, if only because not all languages have stress, there is a difference in whether some formatives that attach to the start of the word are counted for stress, and varieties differ in this regard. This difference can best be accounted for as a clitic versus prefix distinction.

Prepositions cannot receive stress in many varieties, even when in combinations that we would expect to be stressed based on stress parameters. This can be seen in the following example from Cairene, where we would expect that in CV.CV.CVC the first two light syllables would form a foot that would be stressed as the final CVC undergoes consonant extrametricality, as shown in (453-c). However, when this first light syllable is formed by a preposition this does not happen, as seen in (453-b)

(453) Cairene Initial Prepositions Invisible to Stress (Welden 1980, p. 105)

a. ˈkutub

books

b. /fi+ˈkutub/ → *ˈfi+kutub

'in books'

c. ˈkatab -it
write.PFV -SBJ.3SG.F
'she wrote'

However, in Rwaili, prepositions can be stressed if syllable extrametricality has exhausted the noun, as shown in (454).

(454) Rwaili Initial Prepositions Visible to Stress (Prochazka Jr. 1988, p. 21)

ˈb -al -xala
in -DEF -desert

'in the desert'

Similarly, the definite article [ʔal] is unstressable in most dialects, but not in Haːyil, as shown in (455) where syllable extrametricality leaves final CVC invisible to stress so the remaining initial foot is stressed instead. As Rosenhouse (2011) notes, including the definite article in the domain for stress is typical of many Bedouin varieties, and is also found in Bedouin Hejazi Arabic (Kenstowicz 1986).

(455) Haːyil Definite Article Visible to Stress (Prochazka Jr. 1988, p. 21)

a. ˈʔal-bar
'the open (desert)'

b. ˈʔal-ɟimal
'the camel'

In these cases, it seems that the definite article is incorporated into the prosodic word as an affix within the domain for stress assignment, whereas in other varieties where it is not stressable it remains a clitic.

The negation [ma] is unstressable in many varieties including Cairene and Palestinian, but is a stressable prefix in Muscat (see (456)), Iraqi (see (457)), and many Bedouin varieties if it meets the requirements for stress parameters (Fischer and Jastrow 1980). Note it is not that [ma] is obligatorily stressed; rather that the negation [ma] is only stressed if it heads the stressed foot; in (457-b), the word ends in a CVːC which is not affected by syllable extrametricality, so this is stressed under end rule right.

(456) Muscat Negation Prefix Visible to Stress (Glover 1988, p. 229)

a. ˈn -abɣi
SBJ.1PL -want.IMPF
‘we want’

b. ˈmaː n -bɣi
NEG SBJ.1PL -want.IMPF
‘we don’t want’

(457) Iraqi Negation Prefix Visible to Stress (Majdi 1988, p. 59)

a. ˈmaː ʔija -∅
NEG come.PFV -SBJ.3SG.M
‘he didn’t come’

b. ma y -ˈriːd
NEG SBJ.3 -want.IMPF
‘he doesn’t want’

c. ˈlaː ʔil -ak
NEG DAT -OBJ.2SG
‘not for you’

Furthermore, there is also variation in syncope application that can be accounted for by viewing some formatives as clitics attached at the phonological phrase level as ‘free clitics’ in Selkirk’s typology. In the following example from Broselow (1976), the vowel of the habitual marker /bi/ is not syncopated unlike the vowel in the 3rd person marker /ji/ as it is incorporated outside of the phonological word.

(458) Cairene Variation in Syncope Application between Clitics and Affixes (Broselow 1976, p. 5)

/ma bi-ji-zaːkir-uː-ʃ/
ma bi -j -zakr -uː -ʃ
NEG HAB -SBJ.3 -study.IMPF -SBJ.PL -NEG

‘they are not studying’

Whereas in Palestinian, the habitual is a prefix, part of the phonological word, so can be syncopated, as shown in (459).

(459) Palestinian Habitual as Prefix (Abu-Salim 1982a, p. 136)
b -tu -kutb -uː -l -o
HAB -SBJ.2 -write.IMPF -SBJ.PL -DAT -OBJ.3SG.M

‘you (pl.) write to him’

This variation in whether initial formatives are incorporated into the phonological word and as such are within the domain for stress application or syncope can be accounted for by a prefix versus clitic distinction. Therefore, such a distinction is important and exists for Arabic — but the formatives explored in this chapter are not enclitics as has been suggested, but rather suffixes.

4.8 Conclusion

The non-subject bound pronominal formative, the dative /l/ and the negation /ʃ/ are not clitics, as the literature has suggested, but rather suffixes because they display high selectivity for their syntactic head, a canonical affix behaviour. Furthermore, that they do not trigger stem allomorphy in hollow verbs as subject markers do does not mean they must be clitics; rather, there is a distinction but this is a matter of Level 1 versus Level 2 affixation. They are incorporated into the phonological word, and trigger epenthesis, vowel shortening, syncope and stress within this domain. Whilst the non-subject pronominal cannot co-occur with a full NP non-subject, this at most suggests that these endings are slightly aberrant affixes and not agreement markers, rather than entailing that they are enclitics as their behaviours are more consistent with affixhood than clitichood.

They do trigger what is often called 'pre-clitic lengthening' in the literature, where a vowel lengthens before they are incorporated — which I rename 'pre-suffix lengthening'. This comes from a preference to attach after a stressed or stressable foot. As such, it surfaces as vowel lengthening, subject affix consonant gemination, dative gemination, and, in one rare case, final stem consonant lengthening in Libyan Tripoli. This strong preference also underlies apparently exceptional stress behaviour, where stress movement, stressing final long vowels and feminine subject markers before objects can be accounted for as this preference to attach after a stressed or stressable foot. This preference is not unique to this set of endings, but rather also occurs with subject affixes, which can be seen most clearly with weak and biliteral verbs.

Therefore, whilst a range of different strategies may be used either to adapt the underlying metrical structure to form a stressable foot, or moving stress independently of the underlying structure, the same motivation is used for all suffixes. The range of approaches are summarised in Table 4.15 on page 356. If we divide this table by REPAIR, no particularly clear pattern emerges, as shown in Table 4.16 on page 357. If instead we divide it by TOLERANCE as shown in Table 4.17 on page 358, there is little clarity either, beyond the fact that the stress movement without changes to the underlying metrical structure isn't found in varieties with segmental level restrictions within my dataset. However, there is no clear reason why this might be the case, and broader empirical coverage is needed to confirm whether this is merely an artifact of this particular dataset.

Therefore, we can conclude this preference for attachment after a stressed or stressable foot is not a quirk of any particular subtype of Arabic, but rather is a 'baked-in' requirement found across all the varieties explored here, with a wide range of strategies.

Table 4.15: Affixation After Stressable Foot Mechanisms

Variety	Affix Vowel Lengthening	Affix Consonant Lengthening	Dative Gemination	Stem Consonant Gemination	Itu Stress	Stress Movement	Final VV Stress
Palestinian	Y						Y
Cairene	Y				Y		Y
Sˤanʕaːniː						Y	
Tunisian	Y	Y					
Moroccan Casablanca		Y				Y	
Lebanese	Y	Y			Y		
Muscat	Y	Y					
Algerian	Y						
Qatari	Y		Y				
Iraqi	Y	Y					Y
Wadi Ramː	Y						
Libyan Tripoli		Y	Y	Y			
Rufaidah	Y						
Rwaili	Y	Y				Y	
Haːyil	Y	Y					
Makːan	Y		Y		Y		Y

Table 4.16: Affixation After Stressable Foot Mechanisms by Repair Position

Variety	Affix Vowel Lengthening	Affix Consonant Lengthening	Dative Gemination	Stem Consonant Gemination	Itu Stress	Stress Movement	Final VV Stress
Onset Varieties							
Cairene	Y				Y		Y
Rufaidah	Y						
Rwaili	Y	Y				Y	
Haːyil	Y	Y					
Sˤanʕaːniː						Y	
Makːan	Y		Y		Y		Y
Coda Varieties							
Palestinian	Y						Y
Tunisian	Y	Y					
Moroccan Casablanca		Y				Y	
Lebanese	Y	Y			Y		
Muscat	Y	Y					
Algerian	Y						
Qatari	Y		Y				
Iraqi	Y	Y					Y
Wadi Ramː	Y						
Libyan Tripoli		Y	Y	Y			

Table 4.17: Affixation After Stressable Foot Mechanisms by Syllable Restriction

Variety	Affix Vowel Lengthening	Affix Consonant Lengthening	Dative Gemination	Stem Consonant Gemination	Itu Stress	Stress Movement	Final VV Stress
Segmental Restrictions							
Makːan	Y		Y		Y		Y
Palestinian	Y						Y
Lebanese	Y	Y			Y		
Muscat	Y	Y					
Algerian	Y						
Iraqi	Y	Y					Y
Wadi Ramː	Y						
Moraic Restrictions							
Cairene	Y				Y		Y
Rufaidah	Y						
Rwaili	Y	Y				Y	
Haːyil	Y	Y					
Sˤanʕaːniː						Y	
Tunisian	Y	Y					
Qatari	Y		Y				
X-Slot Restrictions							
Moroccan Casablanca		Y				Y	
Libyan Tripoli		Y	Y	Y			

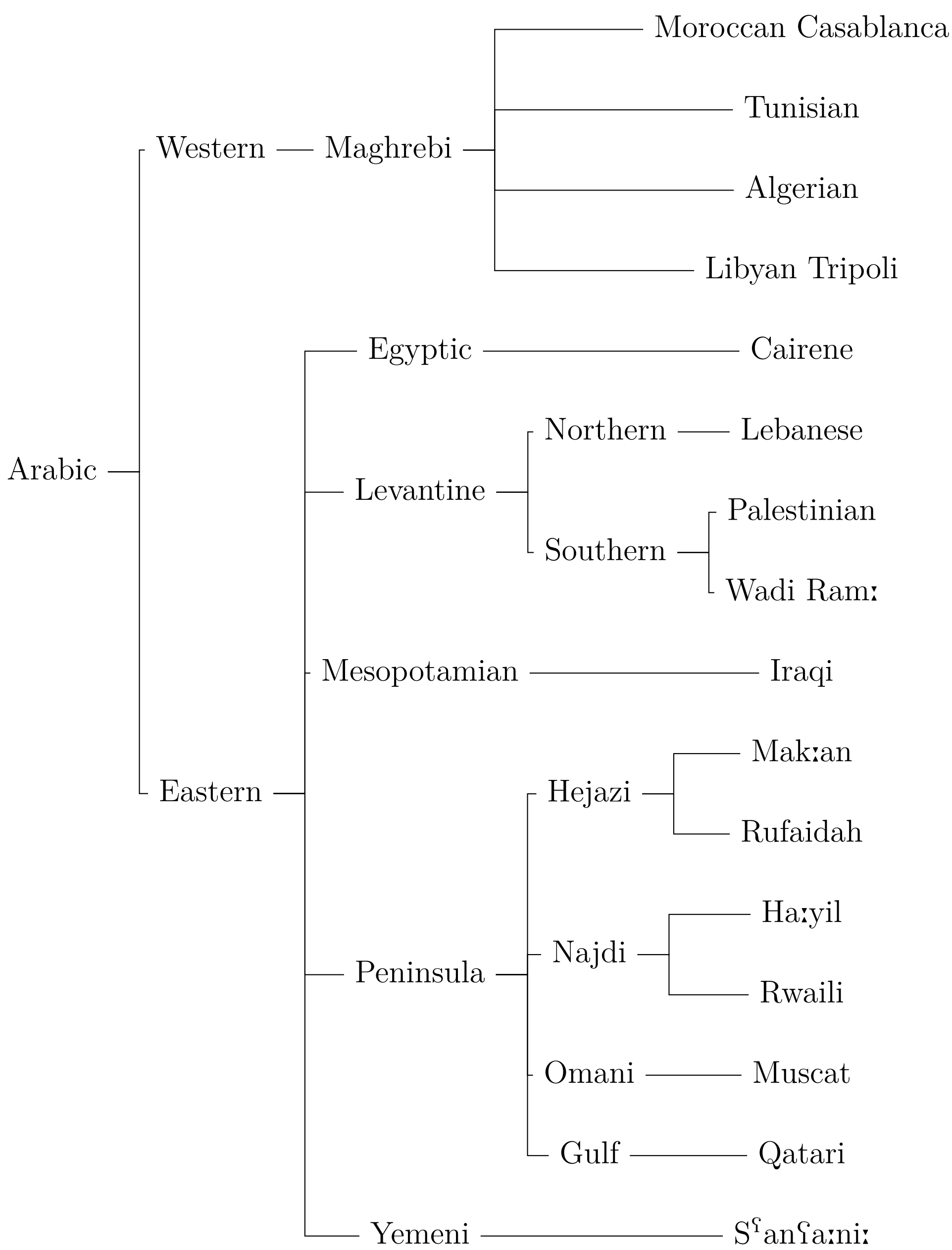
5
Conclusion

Contents

Modern Arabic colloquial varieties differ significantly in their phonology, grammar and lexicon. The phonological variation includes epenthesis position, syncope application, vowel lengthening processes, vowel shortening processes, permissibility of medial and final consonant clusters, permissibility of larger than CVC medial syllables, foot type, parsing directionality, extrametricality, which elements are incorporated into the prosodic word, and exceptions to stress. This thesis has developed a phonological typology for modern varieties, incorporating this wealth of variation into a small set of parameters. In doing so, I have explored not just syllable and foot structure, but also the lexical and postlexical nature of prosodic word structure.

The sixteen varieties I have explored are Moroccan Casablanca, Tunisian, Algerian, Libyan Tripoli, Cairene, Lebanese, Palestinian, Wadi Ramː, Iraqi, Makːan, Rufaidah, Haːyil, Rwaili, Muscat, Qatari, and Sˤanʕaːniː. The distribution of these varieties across the geographical subgroups discussed on page 8 is depicted in Figure 5.1 below.

Figure 5.1: Geographical Classification of Modern Arabic Varieties in the Thesis



In addition to reflecting the geographical spread of Arabic, these varieties also demonstrate the wide ranging phonological variation found, and as such provide a solid empirical basis for the contributions of this thesis. In this chapter, I summarise the findings of this thesis in Section 5.1, before discussing why the variation may be distributed in this way in Section 5.2. I explore the broader implications of my results for phonological theory more widely in Section 5.3 and finish with suggestions for further research in Section 5.4.

5.1 Outcomes and Consequences of this Thesis

I argue that there are two axes for phonological variation in Arabic: Tolerance and Repair, that is, varieties differ in terms of what type of syllables they tolerate and in how they repair violations to syllable structure, and these axes do not covary. Furthermore, I explore the prosodic word structure of Arabic, finding that there are preferences for affixes to attach after a stressed syllable that accounts for lengthening phenomena and apparent stress exceptions.

Tolerance is concerned with what type of syllables are tolerated by different varieties. Crucially, this is a matter of how quantity is tolerated. Some varieties permit medial syllables with long vowels or geminate consonants but not consonant clusters. This can be accounted for with restrictions at different levels of syllable structure, following the vital insight of earlier work that long vowels and geminates can be represented as a single segment associated to multiple mora or X-slot positions whereas there is a one-to-one association between singleton segments and higher levels of structure. Thus, varieties that permit long segments can have restrictions on the *segmental* level, whereas varieties that do not can have restrictions on the *moraic* level, and varieties that allow long consonants or clusters but not long vowels have restrictions on the timing, or *X-slot*, tier. In addition to the canonical two or three segments, moras or X-slots, individual varieties can also permit extra elements in their medial syllable rime according to variety–specific restrictions, thus allowing for some restricted consonantal clusters. These restrictions can include coronality, fast speech, or sonority. As such, this typology provides mechanisms to account for the data in each and every dialect, rather than merely providing a canonical set of features that not all dialects fit. The extrametrical unit is correlated with this parameter, thus syllable extrametricality is associated with segmental level restrictions; consonant extrametricality is associated with moraic level restrictions; and no extrametricality is associated with X-slot restrictions. This is summarised in Table 5.1 below.

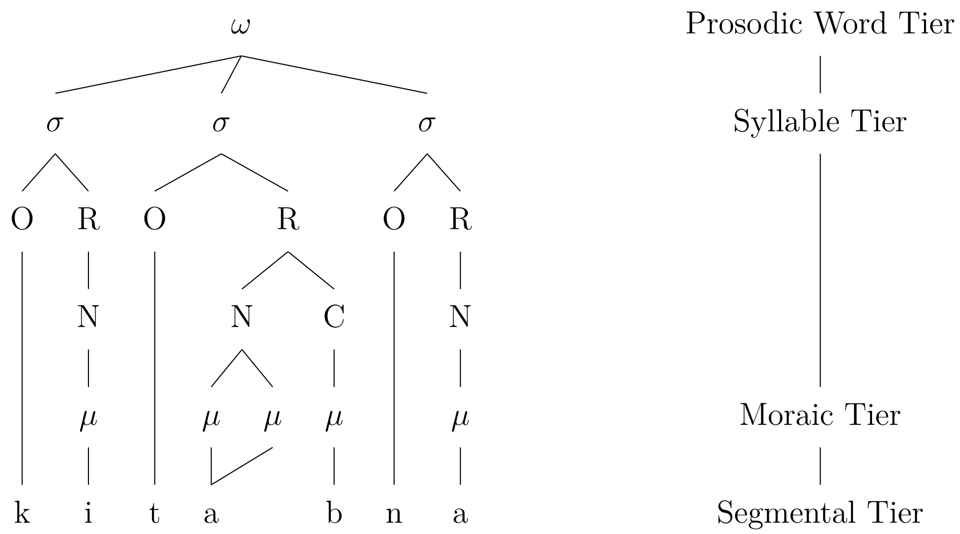
Table 5.1: Overview of TOLERANCE

Tier of Syllable Restriction	Moraic	Segmental	X-Slot
No. of Elements	2–3		
Extra Elements	Variety specific including coronality, sonority and speech tempo		
Extrametricality	Consonant	Syllable	None

For the moraic and segmental structure restrictions, I assume the representation in Figure 5.2.

Figure 5.2: Syllable Structure for [kitaːb-na] with Moraic and Segmental Tiers

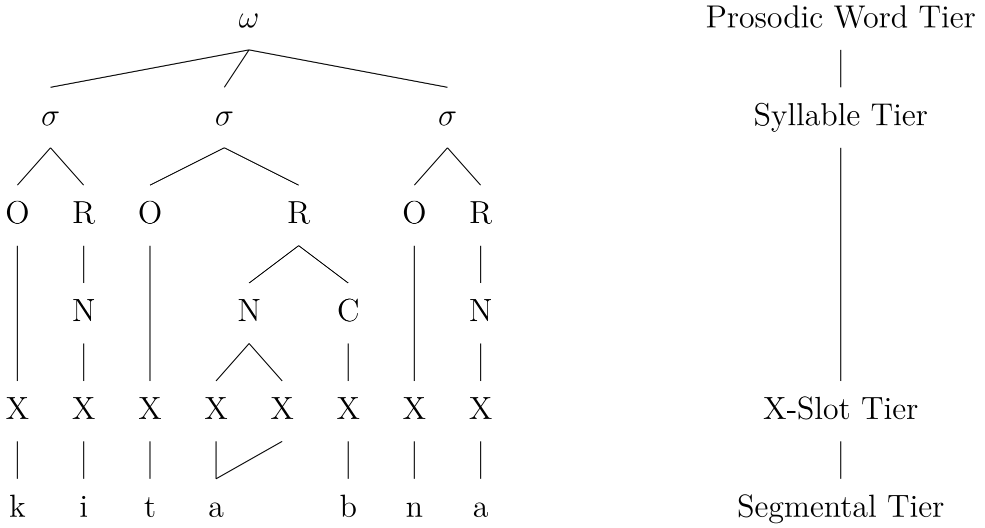
Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora



For the X-slot restrictions, I assume the representation in Figure 5.3.

Figure 5.3: Syllable Structure for [kitab-na] with X-slot and Segmental Tiers

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, X = X-slot



Repair is concerned with how violations of syllable structure are repaired in terms of epenthesis direction and the licensing of word edge extrasyllabic segments. Following Broselow (1992) I agree that varieties can be split into Onset and Coda by whether the epenthetic vowel makes an illicit segment into the onset or coda of a new syllable. I propose that catalectic moras are required to incorporate extrasyllabic material across the word. The Onset/Coda distinction can therefore be expressed as whether the extrasyllabic material is incorporated as the onset or coda of a syllable containing a catalectic mora. Note this can occur at the right or left edge of the word. Epenthesis occurs where the catalectic mora is not permitted to surface without overt segmental material. The difference between domain edge behaviour therefore follows from this — where the catalectic mora in the edge syllable is the peripheral element, the surface cluster is permitted; but where it is non-peripheral, epenthesis occurs as the catalectic mora is filled by an epenthetic vowel. This is illustrated in (460) and (461), repeated from Section 3.4.1.

(460) Initial Clusters

a. Onset Variety: *f[μ]red → fi̲red

b. Coda Variety: [μ]fred

Figure 5.4: Initial Cluster Onset Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora

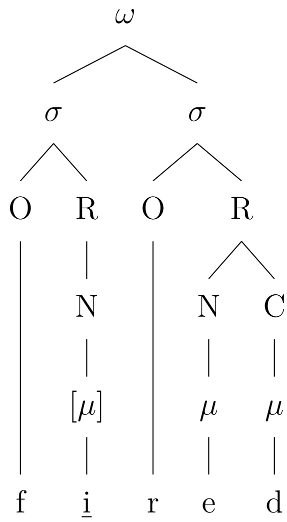
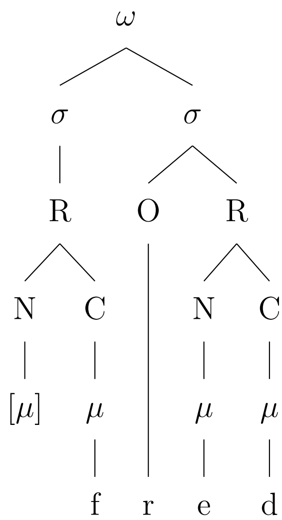


Figure 5.5: Initial Cluster Coda Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora



(461) Final Clusters

a. Onset Variety: katabt[μ]

b. Coda Variety: *katab[μ]t $\rightarrow$ katabi̱t

Figure 5.6: Final Cluster Onset Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora

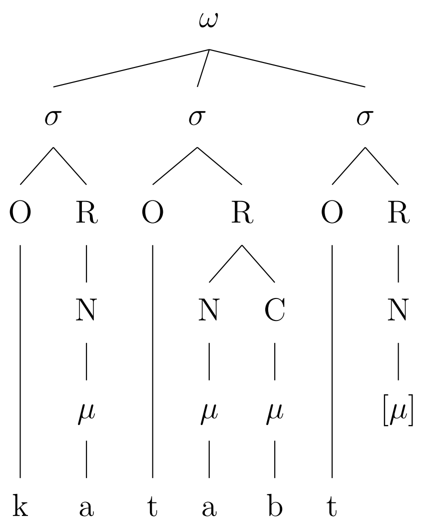
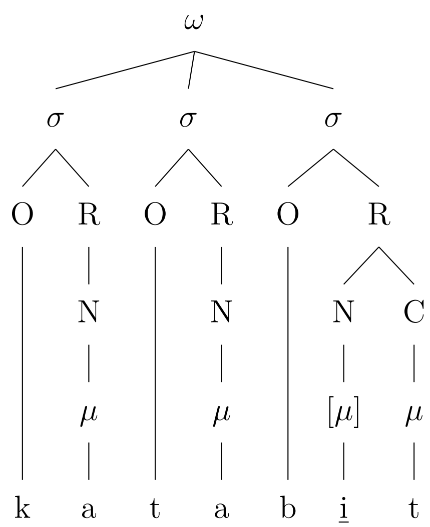


Figure 5.7: Final Cluster Coda Variety

Key: ω = Prosodic Word, σ = Syllable,
O = Onset, R = Rime, N = Nucleus, C = Coda, μ = Mora, [μ] = Catalectic Mora



'Extreme' Onset/Coda varieties are those that permit catalectic moras word-internally in domain edge syllables. I propose that varieties with X-slot Tolerance restrictions also use X-slots in Repair rather than requiring further machinery to capture their distribution. This summarised in Table 5.2 below.

Table 5.2: Overview of REPAIR

Key: Y = Yes, N = No

	ONSET	Extreme ONSET	CODA	Extreme CODA
Epenthesis Position		CV̲		V̲C
Initial #CC-	N	Y	Y	Y
Final -CC#	Y	Y	N	Y

By using catalectic moras (or X-slots) to account for domain-edge clusters, this proposal can also better account for the existence of apparently subminimal words, as well as gesture research in phonetics including Shaw et al. (2009; 2011) that suggests that initial clusters in Moroccan Casablanca are heterosyllabic — which this approach predicts.

Both TOLERANCE and REPAIR are needed to account for data found across Arabic varieties. As shown in the Table 5.3 below, they do not covary.

Table 5.3: TOLERANCE and REPAIR of Arabic Varieties

Variety	TOLERANCE Type	REPAIR Type
Algerian	Segmental	Extreme Coda
Wadi Ramː	Segmental	Extreme Coda
Palestinian	Segmental	Coda
Lebanese	Segmental	Coda
Iraqi	Segmental	Coda
Muscat	Segmental	Coda
Rufaidah	Segmental	Extreme Onset
Rwaili	Segmental	Extreme Onset
Haːyil	Segmental	Extreme Onset
Sˁanʕaːniː	Segmental	Onset
Tunisian	Moraic	Extreme Coda
Qatari	Moraic	Coda
Makːan	Moraic	Onset
Cairene	Moraic	Onset
Moroccan Casablanca	X-Slot	Extreme Coda
Libyan Tripoli	X-Slot	Extreme Coda

This approach has some affinities with constraint and repair theories of Piggott and Singh (1985), Singh (1987), Paradis (1988) and Calabrese (2005), in which the constraints/principles were distinct from the repair mechanisms.

A comprehensive understanding of Arabic phonology typology also requires exploration of prosodic word structure. I challenge the accepted clitic status of the bound pronominal, dative [l] and negation [ʃ], arguing that these are affixes attached on Level 2 whereas the subject marker is attached on Level 1 of the lexical phonology. This difference in lexical level can account for the pattern of closed syllable shortening, distinguishing between CSS-Morph that occurs on Level 1 for all varieties and CSS-Phon that occurs on Level 2 according to the syllable structure restrictions of each individual variety. I demonstrate that these affixes in addition to the subject affix share a prosodic preference to attach after a stressed syllable. This preference triggers pre-affix lengthening in many varieties, but also accounts for extensive patterns of apparent stress exceptions in others — as varieties either lengthen the syllable before the affix to make it stressable, or they move stress without amending the underlying metrical structure.

5.1.1 Advantages and Benefits

This bipartite typology of Arabic phonology has stronger empirical coverage of the data than previous work as it defines a more nuanced canonical space as well as providing mechanisms for the production of non-canonical forms.

In Chapter 2 on page 51, I made the following criticisms of existing work on Arabic phonological typology. Firstly, I demonstrated that there were more medial CVXC syllables than predicted in earlier work, and in particular that more CVCC and CVCː surface. Secondly, I noted the contradiction in the work of Kiparsky (2003) and Watson (2007) where they claim that final clusters surface in CV varieties, but at the same time CV varieties do not permit semisyllables, their mechanism for permitting such clusters. Thirdly, I questioned why C and Cv varieties should share so many features in common, including medial CVːC syllables, final CVCCC and CVːCC, postgeminate high vowel deletion, and initial clusters, but yet require completely different analyses in Watson (2007) — where C varieties permit semisyllables at both lexical and postlexical levels whereas Cv dialects do not permit semisyllables at either level, and mora sharing is permitted iff the syllable rhyme is CVːC or CVCC. This formalisation fails to express the similarities seen between Cv and C varieties. Finally, in Chapter 4, I also critique earlier work on Arabic clitics and affixes, as well as analyses of apparent stress exceptions that need extensive novel machinery to solve. In the face of such issues, my proposal here has several advantages over earlier work.

Firstly, I have demonstrated that indeed more medial CVCC and CVCː surface than expected, given that previous work expects no medial CVCC and CVCː syllables to surface outside of C varieties. I find that such syllables do not pattern together, and crucially there is variation in acceptance of quantity that has not been recognised in earlier work. This is the core insight of the TOLERANCE part of my typology, and as such accepts and predicts the existence of such medial syllables rather than ignoring them as inconvenient or exceptional.

Secondly, the REPAIR component of my typology can account for why the asymmetry in edge cluster permissibility correlates with epenthesis position and provides catalexis as a mechanism for repair across the word and across the different varieties. I have shown that ONSET varieties can permit final clusters not initial clusters, as in final clusters the catalexic mora is peripheral unlike in initial clusters, whereas the reverse is true of CODA varieties. Thus, edge clusters and medial epenthesis can be accounted for by the same mechanism: catalexic moras are inserted to make extrasyllabic material the onset or coda of a new syllable. Whether the catalectic mora becomes segmentally overt with an epenthetic vowel is a matter of whether the catalectic mora is permitted to remain covert. Not only does this analysis unify the behaviours of clusters across the word more comprehensively than previous accounts, but it also explains why apparently superheavy final CVXC syllables can escape extrametricality.

Thirdly, by defining a more nuanced canonical space through the use of two axes and several positions within these as well as providing mechanisms for variation within the canonical space, this proposal does not require the constant addition of extra dialect groups. Rather, the similarity of apparent 'Cv' varieties with others comes from the combination of ONSET repair with SEGMENTAL tolerance restrictions. Thus, I provide a more flexible approach that is resilient to the broad range of variation found across varieties identified and those yet to be analysed.

Finally, I find that existing solutions for stress exceptions including an exceptional reversal of parsing directionality, branching nodes, indestructible rhymes and more complex underlying forms (Watson 2002; McCarthy 1979; Angoujard 1981; Haddad 1984; Broselow 1976; Brame 1971) are not needed — rather the extensive pattern of stress exceptions follows from a prosodic preference to attach after a stressed syllable. This prosodic preference thus accounts for extensive variation in lengthening, shortening and stress exception behaviour, turning them from apparent idiosyncratic behaviours to a baked-in requirement of the phonological grammar.

In addition to the advantages explored here, a further benefit to this research is that the behaviour of stress is firmly incorporated into the phonological typology, whereas in previous research it was not.

5.2 Geographical Distribution of Patterns

In the preceding chapters I have motivated this phonological typology and commented on the geographical distributions of particular features but not discussed the implications for classifications of modern varieties, nor hypothesised as to the direction of change.

Table 5.4 below summarises the typology of the thesis with the varieties organised by geographical location. The map of where each variety is spoken from Figure 1.5 is repeated here as Figure 5.8

Table 5.4: Distribution of Phonological Typology Parameters Across Varieties

Maghrebi	Moroccan Casablanca	X-slot	Extreme Coda
	Algerian	Segmental	Extreme Coda
	Tunisian	Mora	Extreme Coda
	Libyan Tripoli	X-slot	Extreme Coda
Egyptic	Cairene	Mora	Onset
Levantine	Lebanese	Segmental	Coda
	Palestinian	Segmental	Coda
	Wadi Ramː	Segmental	Extreme Coda
Mesopotamian	Iraqi	Segmental	Coda
Peninsula	Muscat	Segmental	Coda
	Qatari	Mora	Coda
	Rwaili	Segmental	Extreme Onset
	Haːyil	Segmental	Extreme Onset
	Rufaidah	Segmental	Extreme Onset
	Makːan	Mora	Onset
Yemeni	Sˤanʕaːniː	Segmental	Onset

Figure 5.8: Map of the Locations of the Varieties from the thesis

As discussed in Section 1.1.4, in the dialectological literature modern varieties have been classified in terms of non-linguistic commonalities, including geography, historical development, and social factors. In this section, I will focus on the geographical distribution of these phonological behaviours.

A key distinction made in the literature is between Western and Eastern varieties; that is, between the Maghrebi varieties of North Africa and the rest. In previous phonological work, Farwaneh (2009) has marked these varieties as 'Extreme Coda' varieties, and predicts that these are more innovative, with other varieties moving in this direction. Here, I suggest that there is a greater degree of phonological variation amongst Maghrebi varieties than has previously been recognised, and that their features are not unique to North Africa but rather are found across the entire dataset. While all four share the same Repair strategy, there is a wide range of Tolerance types. The use of *X-slot* syllable restrictions is innovative, and reflects the extreme degree of vowel reduction in Moroccan Casablanca and Libyan Tripoli to the extent that it has affected the stress system. The extent to which this might reflect Berber influence is unclear, particularly in light of the mixed support in the literature for Berber influence in Moroccan Arabic more broadly[1] and requires further investigation. Contrary to the predictions of Farwaneh (2009), the Maghrebi coda varieties are not the only ones that can be described as 'extreme'. In this analysis, 'extreme' refers to the acceptance of non-peripheral catalectic moras, and this is seen not just in Moroccan Casablanca, Algerian, Tunisian, and Libyan Tripoli, but also in some Onset peninsula varieties: Rufaidah, Rwaili, Haːyil. In doing so, this questions whether the Peninsula varieties truly are the more 'archaic' varieties. Note that the 'extreme' varieties are also those that permit three elements in their medial rimes, whether X-slot, moras or segments. Thus, the Maghrebi North African varieties can be viewed as innovative in the use of X-slot tolerance in some varieties, and indeed in permitting larger syllables. However, these innovations are not uniquely Maghrebi features, but rather reflect ongoing change across the region.

[1] See Diem (1979, pp.52–55), El Aissati (2006), Taine-Cheikh (2007), Chtatou (2009), Bensoukas and Boudlal (2012) and Aguadé (2018, p. 35) inter alia for more on this debate.

Beyond the Maghrebi/Mashriqi division, there are some geographic patterns in the distribution of these phonological behaviours. Neighbouring varieties Palestinian, Lebanese, Wadi Ramː, Iraqi share both coda repair and segmental tolerance patterns, with the onset repair/segmental tolerance set of Sˤanʕaːniː, Rufaidah, Rwaili, and Haːyil also close geographically. However, despite these close patterns, there are nearby varieties that do not fit with the rest of the set, notably Qatari and Muscat behave differently to the other peninsula varieties, and Makːan has a moraic tolerance compared to the other Saudi Arabian varieties' segmental tolerance. The distributions of moraic and x-slot tolerance restrictions are less geographically clear, and this is in part due to the wide variation found in North Africa. While both Moroccan Casablanca and Libyan Tripoli are North African varieties, they are at opposite ends of the region with segmental and moraic varieties in between. The variation in North Africa is greater than in the Levantine and Peninsular varieties, and likely reflects waves of invasion. Investigation of a larger set of North African varieties with close attention paid to the historical origins of the varieties is needed to clarify the origins of this distribution.

In terms of the mechanisms used to ensure that affixes attach after stressed feet, the geographical distribution is not clear, as shown in Table 5.5 below. Thus, it seems there are variety-specific motivations for the particular mechanisms chosen, though they all share the same prosodic attachment preference.

Given the similarities found, there is perhaps motivation to group the mesopotamian and levantine varieties together, while also recognising greater variation within the maghrebi and peninsula sets. However, ultimately it does appear that the phonological behaviour of an individual variety is reflects a complex interactions between several extralinguistic features. For example, the Bedouin origins of Wadi Ramː, Rwaili and Haːyil leads to shared iambic feet, but these varieties do differ in their repair strategies, perhaps reflecting different geographical regions. They do share their tolerance restrictions, but so do a wide number of sedentary varieties in the region. Although the synchronic situation reflects a complex interaction of extralinguistic pressures, hypotheses can still be formed as to the direction of change occurring in the region.

I hypothesise that *moraic* syllable restrictions and ONSET repair is the older form, and that the use of *segmental* restrictions, *X-slot* restrictions, and CODA repair are innovations which I expect to spread further across the varieties over time, just as we see transition here in Makːan and Qatari from *moraic* to *segmental* restrictions. Although there is limited evidence that Classical Arabic is the direct ancestor of the modern varieties, certainly there is evidence that this variety used onset repair and moraic syllable restrictions, not permitting CVXC syllables except in prepausal positions. Ibn Yaʔiːʃ, in his ʃarħ al-Mufasˤːal (IX:127) presents the classical behaviour as follows:

> ‘the general rule, in all cases where two consonants without a vowel come together is to insert an *i* after the first, as in *baga:t i l-’ama* (the slave fornicated) and in *qa:mat i l-ga:riya* (the servant arose), and is not applied only when there is good reason for not doing so.’
>
> translated by Bohas et al. (1990, p.92)

This depiction of epenthesis follows the Onset behaviour discussed in this thesis, and is still shared by the more conservative peninsula varieties of Rufaidah, Rwaili, and Haːyil. However, permitting syllables larger than CVC in non-pausal positions is an innovation that has spread farther than the changes in epenthesis position, and perhaps was triggered by the loss of case endings and other final short vowels across modern varieties.

Ultimately, the complex modern synchronic pattern would benefit from historical investigation of its development, taking into account the impact of the close contact between varieties.

Table 5.5: Affixation After Stressable Foot Mechanisms by Syllable Restriction

Variety	Affix Vowel Lengthening	Affix Consonant Lengthening	Dative Gemination	Stem Consonant Gemination	Itu Stress	Stress Movement	Final VV Stress
Maghrebi							
Algerian	Y						
Tunisian	Y	Y					
Moroccan Casablanca		Y				Y	
Libyan Tripoli		Y	Y	Y			
Egyptic							
Cairene	Y				Y		Y
Levantine							
Palestinian	Y						Y
Lebanese	Y	Y			Y		
Wadi Ramː	Y						
Mesopotamian							
Iraqi	Y	Y					Y
Peninsula							
Makːan	Y		Y		Y		Y
Muscat	Y	Y					
Rufaidah	Y						
Rwaili	Y	Y				Y	
Haːyil	Y	Y					
Qatari	Y		Y				
Yemeni							
Sˤanʕaːniː						Y	

5.3 Broader Implications of these Results

This thesis contributes a new phonological typology of modern Arabic varieties, providing a comprehensive view of all matters related to word prosody. While these findings have clear relevance for our understanding of Arabic in particular, they also have implications for phonological theory more widely, in particular with regards to the nature of syllable structure restrictions, the repair of extrasyllabic material, the relationship between stress and syllable structure, and clitic diagnostics.

The debate on syllable structure restrictions has focused on X-slot versus moraic representations. While I do find that there are indeed varieties that place syllable size restrictions on these levels, I have argued here for the existence of varieties with segmental restrictions on syllable structure. This contradicts the prediction of Kenstowicz and Pyle (1973) that prosodic processes including syllabification will occur on a moraic tier. Regardless of the tier of syllable restrictions, I show that there are more trimoraic syllables word-medially than would be predicted either by the Bimoraicity Constraint of Broselow (1976) or by the broader literature that suggests trimoraic syllables are unlikely (Trubetzkoy 1939; Hayes 1995; Broselow, Chen, and Huffman 1997).

Repair of extrasyllabic material has often been construed as addition or deletion of segmental material. Here, I have reanalysed catalexis as a tool to repair extrasyllabic material as well as degenerate feet. I demonstrate that catalectic moras can occur symmetrically on both the left and the right edge, unlike extrametricality where the right edge is predominant. This mechanism also explains why final apparently superheavy syllables can escape extrametricality.

While syllable weight has been shown to directly affect stress placement, the correlation between extrametrical unit and syllable structure restrictions found here is more opaque and unexpected. This points to a more complex relationship between the two within the phonological system as a whole that requires further investigation.

I have also challenged approaches to diagnosing the clitic/affix distinction. I demonstrate that a sharper understanding of host selectivity is required, and crucially argue that failure to undergo morphophonological alternations is not necessarily sufficient to require clitic status. While there are clitics for which this failure is a logical consequence of clitic status, there are morphophonological alternations that only affect particular affixes, which can be accounted for a distinction between lexical levels rather than in types of bound formatives.

These findings lead to several predictions about what is possible and what is not possible in language. Firstly, I predict that there is no language that permits CVCC but not CVːC and CVCː where long vowels and geminates can exist. If syllable restrictions are on a segmental level, there is no difference between the syllables with long segments versus singletons, and if CVCC is permitted so should CVC. However, if the syllable restrictions are on a moraic level and there is weight by position, there is no difference between any of the CVXC syllables. Therefore, it is unlikely that a distinction will be made permitting CVCC but not CVːC or CVCː if these exist.

Secondly, I predict that where catalectic moras repair degenerate syllables, there will not be a directional mismatch between domain-edge clusters and domain-medial clusters unless sonority intervenes, as the epenthesis in all positions should reflect the insertion of a catalectic mora underlyingly and it is likely there is consistency in the insertion rules. That is, apparent epenthesis is unlikely to have different positions word medially and word finally unless there are sonority requirements on syllable structure that supersede this. However, there can be more edge clusters than expected if the language permits word-internal catalectic moras in peripheral syllables.

Thirdly, I predict that word edge clusters will be treated as heterosyllabic in production, whether in phonetic articulation or psycholinguistic planning, as these reflect two syllables with a catalectic mora in the edge syllable. This is confirmed by gesture research including Shaw et al. (2009; 2011) for Moroccan Casablanca.

Fourthly, I predict that catalexis can be present not just where degenerate feet are found but also degenerate syllables at either word edge, and that in languages where catalexis does repair extrasyllabic material, CVXC syllables cannot be made extrametrical as the catalectic mora makes these syllables non-peripheral and as such not available for extrametricality.

5.3.1 Derivational versus Constraint-based Models of Phonology

Constraint-based models are the koiné of the day in phonology, however this thesis has been written within the derivational tradition instead. While some aspects of the thesis, including the prosodic preferences of affixes, may appear suited to constraint–based approaches, I have not converted my analyses due to concerns of the limitations of these approaches. In particular, Arabic varieties not only appear to undergo input constraints, but also require an abstract underlying form that is not compatible with output-constraint approaches.

Optimality Theory contends that the phonological grammar involves language-specific ranking of violable constraints on outputs against which a universal set of inputs is tested to choose the most optimal candidate. Crucially, it is claimed that there are no constraints on the input, through the principle of Richness of the Base. Smolensky (1996) suggests that the set of inputs is universal, so that all cross-linguistic variation is found in constraint re-ranking. The original constraints are Faithfulness (to input, or underlying form) and Markedness, though these are enriched in later work. Faithfulness constraints entail that the output needs to be a similar as possible to the input, and as such have consequences for the structure of the lexicon (Yip 1996). Thus a key question is what is this input, that is, what is the shape of the lexicon. To answer this, Prince and Smolensky (1993) proposed Lexicon Optimisation:

> 'Suppose that several different inputs I_1, I_2, ..., I_n when parsed by a grammar G lead to corresponding outputs O_1, O_2, ..., O_n, all of which are realized as the same phonetic form ϕ - these inputs are all phonetically equivalent with respect to G. Now one of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labelled O_k,. Then the learner should choose, as the underlying form for ϕ, the input I_k. '

From Prince and Smolensky (1993, p. 209)

That is, in language acquisition the learner will choose the input that would make the observed output form the most optimal candidate. As a result, 'lexicon optimization entails that the constraints on surface syllable structure will be echoed in the lexicon as well' (Prince and Smolensky 1993, p. 210). I question how lexicon optimisation does not inherently involve constraints on the input, as the characterisation of language acquisition in choosing the input that would make the output most optimal does appear to require that the input is constrained by these constraints. As such, we see in Arabic that epenthesis occurs only at morpheme edges, suggesting that morphemes fit syllable structure constraints on input, and that there are homorganicity restrictions on root consonants. How this is not therefore inherently the 'carefully contrived limits of the inputs to the grammar' that McCarthy (2007, p. 18) eschews is unclear.

The consequences of lexicon optimisation lead to a further question: what about morphological alternations? Is a more abstract underlying form for all the surface alternations needed even though it doesn't surface faithfully, as is often chosen in derivational approaches, or a multiplicity of concrete underlying forms stored, each corresponding to a different surface form? Lexical economy is introduced to prioritise a smaller number of more abstract forms in the lexicon, though other approaches exist, including the stored allomorphy proposals of Mester (1994), Burzio (1996), Kager (1996), and Tranel (1998), the Strong Lexicon Optimization of Sanders (2006) and the Selective Lexicon Optimization of van Oostendorp (2014). Yip (1996) notes four different types of lexical economy: economy of individual lexical entries (as in *SPEC (Prince and Smolensky 1993)); economy of phoneme inventory (involving judgements across the entire lexicon, through an as-yet unclear mechanism); economy of phonotactic combinations (which is traditionally known as Morpheme Structure Constraints); and economy of paradigms (the assumption that same underlying form is used for the whole paradigm even if it is not most economical for a single given form, such as in Tesar and Smolensky (1998).)

While lexical economy in its various forms does permit inputs that are not necessarily most faithful to the output, abstract underlying representations are broadly avoided in constraint based approaches, and in extreme cases some argue that there is no need for any underlying representations in Optimality Theory at all

(Hammond 1995; Russell 1995; Golston 1995; Yip 1995). This is not compatible with the abstractness of the catalectic mora proposed in this thesis. Similar concerns as to how constraint based approaches can account for segmentally non-overt material are raised in Hall (2007). For other criticisms of constraint-based approaches see Idsardi (1998) for the challenges of handling abstractness and opacity in OT, Nevins and Vaux (2003) for the issue of absolute ungrammaticality, and Paster (2013) for a discussion of the duplication problem.

5.4 Directions for Future Research

Nonetheless, this research has uncovered several directions for future research.

In terms of experimental psycholinguistic research, the catalexis analysis can be enriched by production timing experiments to explore whether final CVXC syllables are planned as one unit or two. The analysis here would predict that they are indeed planned as two units regardless of morphological structure, and this should be reflected by increased time needed for on-line planning.

In terms of historical research, while new patterns have been identified in this research including the relationship between syllable structure and stress, why and how these are linked is unclear. Exploring how varieties have changed over time in their stress and syllable structure, including whether change in one precipitates change in the other, would be an important direction of research, especially in light of the hitherto primarily segmental focus of Arabic historical linguistics. Furthermore, this thesis has uncovered further strong patterns of phonological behaviour. Whether these groupings are associated with other morphological and phonetic patterns may allow us to investigate the historical development of modern Arabic varieties from a new angle.

Future questions highlighted by this research need not be solely of interest to Arabists. The implications of this use for catalexis in accounting for the behaviour of extrasyllabic elements or superheavy syllables in other languages merits further exploration. Furthermore, why and how stress and extrametricality is linked to the tolerance of medial syllables suggests the need for further investigation of the relationship between stress and syllable structure. Finally, this thesis posits that some elements are clitics in some varieties and prefixes in others. Whether the direction of change is prefix to enclitic, as has occurred in German, or whether clitics have become prefixes, is a controversial area to explore.

Whilst the typology motivated here covers the nature of and interactions between syllable structure, foot structures and the prosodic word, Arabic cannot be unique in this variation. What this thesis has shown is that there are constraints on both the input and the output of the phonological grammar, and that the variation rests on the acceptability or violation of the constraints. This widespread synchronic variation in Arabic varieties reflects smaller diachronic changes that have had wide ranging impacts on the entire phonological system. Thus, the Arabic typology emphasizes two aspects: first, variation in phonology is not just on the surface, and second, abstract quantitative analyses including catalexis are fundamental to our understanding of how prosodic structure can survive across a variety of dialects, possibly across centuries.

Appendices

A
Data Appendix

Contents

This appendix contains the root and citation form for all the data used in this thesis. Note that where the form uses a biliteral root, I have noted this as C_1-C_2 rather than C_1-C_2-C_2.

A.1 Moroccan Casablanca

Table A.1: Data from Moroccan Casablanca

Root	Lexeme	Translation	Word-Type
b-n-y	bany	he built	verb
b-w-s	bas	he kissed	verb
f-x-r	faxər	coal	noun
d-a	dːa	he took	verb
d-y-r	daːr	he did	verb
dˤ-r-b	dˤarəb	hit	noun
dˤ-r-b	dˤərb	hitting	noun
ʤ-r-ʤ-r	ʤərʤər	he trailed	verb
g-l-b	gləb	he reversed sth.	verb
k-l	kal	he ate	verb
k-r-k-b	kərkəb	he rolled	verb
k-t-b	ktəb	he wrote	verb
k-t-f	ktəf	shoulder	noun
l-m-n	limuna	orange	noun
l-w-y	lawiy	twisting	adjective
l-ʕ-b	lʕəb	he played	verb
m-l-m-l	məlməl	he shook	verb
m-r-m-d	mərməd	he trailed sb. (in dust)	verb
n-m-l	nməl	ant	noun
n-z-r	minzara	sharpener	verb
q-f-z	qfəz	cage	noun
q-l-b	mqulb	deceiver	participle
q-t-l	qtəl	he killed	verb
r-b-l	rubl	he disturbed	verb
s-l-m	məslim	Muslim	noun
s-m-a	sma	sky	noun
s-m-ħ	sməħ	he forgave	verb
s-n-s-l	sənsla	zip	noun
s-r-q	srəq	he stole	verb
s-ˤ-f-r	sˤfər	yellow	adjective
ʃ-m-ʕ	ʃmiʕa	smile	noun
t-n-b-r	tənbər	stamp	noun
tˤ-r-q	mtˤərqa	hammer	noun
w-l	walu	nothing	noun
z-r	ʔaməzːru	Amezrou (place)	noun
z-r	məzːriwi	someone from Amezrou	noun
ʒ-r-ʒ-r	ʒərʒər	he trailed in dust	verb

A.2 Algerian

Table A.2: Data from Algerian

Root	Lexeme	Translation	Word-Type
b-n-t	bənt	girl	noun
b-ʕ-a	baʕa	he sent	verb
d-y	dːa	he took	verb
d-f-n	ndfan	he was buried	verb
d-n	dːan	he called to prayer	verb
ɣ-l-b	nɣlab	he was defeated	verb
h-t-r-f	hatraf	he dreamt	verb
h-t-r-f	mhatraf	being dreamt	adjective
ħ-b-s	ħabs	jail	noun
ħ-k-m	ħkem	he caught	verb
ħ-m-r	ħmaːr	donkey	noun
k-d-b	kadab	he lied	verb
k-d-b	kadːab	he accused (s.o.) of lying	verb
k-l-b	kəlb	dog	noun
k-t-b	ktəb	he wrote	verb
m-a	mːa	mother	noun
m-l-k	malk	property	noun
n-ʤ-m	nʤuːm	stars	noun
q-n-q-r	qanqra	toad	noun
q-t-l	qatl	killing	noun
q-t-r	qutr	territory	noun
r-g-d	ragːad	he put to sleep	verb
r-g-sˤ	rgesˤ	he danced	verb
r-z-q	razq	possession	noun
s-b-g	saːbig	overtaking	adjective
s-b-g	sabag	he overtook	verb
s-b-g	sabːag	he made sb. overtake	verb
s-b-t	sabt	Saturday	noun
s-n-s-l	sansla	chain	noun
ʃ-d	ʃədː	he held	verb
ʃ-f	ʃafː	he inspired pity	verb
ʃ-w-f	ʃaːf	he saw	verb
t-b-ʕ	tabːaʕ	he followed	verb
w-q-t	waqt	time	noun

A.3 Tunisian

Table A.3: Data from Tunisian

Root	Lexeme	Translation	Word-Type
b-l-d	bulada:	tiresome people	noun
b-r-d-g-n	birdga:n	oranges	noun
b-w-s	ba:s	he kissed	verb
d-a	d:a	he took	verb
d-h-n	dihn	paint	noun
ʤ-y-b	ʤa:b	he brought	verb
f-d-y	iftda	he was revenged	verb
f-r-q	far:aq	he distributed	verb
ħ-b-s	ħabs	jail	noun
k-f-r	kifir	he blasphemed	verb
ʔ-k-l	kal	he ate	verb
k-l-m	kal:am	he talked to	verb
k-n-z	kinz	treasure	noun
k-r-m-b	krumb	cabbage	noun
k-t-b	ktib	he wrote	verb
k-t-b	maktba	library	noun
l-b-s	lab:is	he clothed	verb
l-m-n	limuna	lemon	noun
l-ʕ-b	laʕ:ab	he made play	verb
m-a	m:a	mother	noun
m-q-t	maqt	rage	noun
m-r-a	mra	wife	noun
m-r	mr:a	itch	noun
n-w-r	nu:r	light	noun
q-l	qul:a	water jug	noun
q-r-y	qra:	he read	verb
q-w-m	qa:m	he rose	verb
r-q-b	raqba	neck	noun
r-q-sˤ	ra:qsˤa	dancer	noun
r-w-ħ	ra:ħ	he went	verb
r-w-ħ	ru:ħ	soul	noun
r-y-ħ	ri:ħ	wind	noun
s-k-m	sukma	cold	noun
s-l-s	salsla	chain	noun
s-m-a	sma	sky	noun
s-m-r	misma:r	nail	noun
s-r	sri:r	bed	noun
ʃ-r-k	ʃarika	company	noun
ʃ-r-y	ʃra:	bought	berb
tˤ-ħ-n	tˤaħu:na	mill	noun

continued ...

Root	Lexeme	Translation	Word-Type
tˁ-y-r	tˁaːr	he flew	verb
w-r-q	wraqː	paper	noun
x-b-y	mxubːi	hidden	adjective
x-d-m	xdiːma	maid	noun
x-w-l	xaːl	uncle	noun
ʒ-y-b	ʒab	he brought	verb
ʔ-q-l	ʔaqalː	less	elative adjective
ʕ-m-d	muʕtamadː	head of administrative district	noun

A.4 Libyan Tripoli

Table A.4: Data from Libyan Tripoli

Root	Lexeme	Translation	Word-Type
b-r-ʧ-q-l	bərʧqal	oranges	noun
b-w-s	bas	he kissed	verb
b-y-ʕ	baʕ	he sent	verb
d-n	dːən	he anointed	verb
d-r-k-b	dərkəb	he hit the drum	verb
dˁ-r-b	dˁrəb	he hit	verb
f-m-m	nəfəmː	he was understood	verb
f-r-ʃ-k	fərʃək	he relaxed	verb
f-tˁ	fːətˁ	towels	noun
g-d-b	gdəb	he lied	verb
ħ-m	stħəm	he bathed	verb
ħw-sˁ	ħusˁa	house	noun
k-l-b	kəlb	dog	noun
k-r-s	kərsi	chair	noun
k-s-r	ʧkəsːər	he was broken	verb
k-ʧ-b	kʧəb	he wrote	verb
l-z-r	luzir	minister	noun
m	əmː	mother	noun
m-l-y	mːali	parents	noun
m-r	stamarː	he continued	verb
m-r-d	mrəd	he crawled	verb
m-ʃ-y	mʃa / əmʃa	he went	verb
n-s-y	nisa	he forgot	verb
q-dˁ-y	qadˁi	judge	noun
q-l-q	qəlːəq	he bored	verb
q-t-l	qtəl	he killed	verb

continued ...

Root	Lexeme	Translation	Word-Type
q-y-l	qal	he said	verb
s-k-r	msək:rin	closed	adjective
s-m-n	sman	he became fat	verb
s-n-s-l	sənsla	necklace	noun
s-n-y-t	sanyətʃ	field	noun
s-t-n	stən	he waited	verb
ʃ-d	ʃəd:	he seized	verb
tˤ-b-ħ	ˈtˤəb:əħ	he called	verb
tˤ-w-sˤ	tˤasˤa	cup	noun
w-q-f	wqəf	he stopped	verb
x-b-z	xəbz	bread	noun
x-l-q	xləq	he created	verb
x-r-f	xər:əf	he told	verb
z-w-r	zar	he visited	verb
ʒ-y-b	ʒab	he brought	verb
ʕ-l-m	ʕəl:əm	he taught	verb

A.5 Cairene

Table A.5: Data from Cairene

Root	Lexeme	Translation	Word-Type
b-n-t	bint	girl	noun
b-n-y	bana	he built	verb
b-w-b	ba:b	door	noun
b-y-t	be:t	house	noun
dˤ-r-b	dˤarab	he hit	verb
f-h-m	fihim	he understood	verb
g-y-b	ga:b	he brought	verb
ħ-b	ħab:	he loved	verb
j-w-m	jo:m	day	noun
k-l	kul:	all	determiner
k-l-m	kal:am	he spoke	verb
k-n-k	kanaka	coffee pot	noun
k-t-b	katab	he wrote	verb
k-t-b	maktab	office	noun
k-t-b	maktaba	library	noun
k-t-b	kita:b	book	noun
k-w-n	ka:n	he was	verb
m-l	ʔamal:	most boring	elative adjective

continued ...

Root	Lexeme	Translation	Word-Type
m-sˤ-r	masˤr	Egypt	noun
r-m-y	rama	he threw	verb
ʃ-k-l	muʃkila	problem	noun
t-ħ-t	taħt	under	preposition
z-k-r	zaːkir	he studied	verb
ʔ-b	ʔabː	father	noun
ʔ-m-l	ʔamal	hope	noun
ʔ-x	ʔaxː	brother	noun
ʔ-w-l	ʔaːl	he said	verb
ʕ-j-n	ʕajːaːn	sick	adjectives

A.6 Lebanese

Table A.6: Data from Lebanese

Root	Lexeme	Translation	Word-Type
d-ʕ-w	daʕa	he summoned	verb
d-r-b	darb	road	noun
dˤ-r-b	dˤarab	he hit	verb
ħ-f-r	ħafar	he dug	verb
ɣ-l	mistɣilː	taking advantage of	adjective
ħ-b	ħabː	he loved	verb
ħ-m-l	ħamal	he neglected	verb
k-t-b	kitab	he wrote	verb
k-t-b	maktuːb	letter	noun
k-t-b	maktabe	library	noun
k-t-b	kaːtib	writer	noun
n-d-r	nidi̲r	vow	noun
n-m-r	nimr	tiger	noun
n-s-l	nasi̲l	progeny	noun
n-s-r	nisi̲r	eagle	noun
n-z-l	nizil	he descended	verb
n-z-l	nazːal	he brought down	verb
n-z-l	naːzal	he taught	verb
r-s-l	mraːsil	correspondent	noun
s-k-r	sakar	he intoxicated	verb
s-k-r	sakːar	he closed	verb
s-m-ʕ	simʕa	reputation	noun
s-t	sitː	grandmother	noun
ʃ-t	ʃatːat	he spread	verb

continued ...

Root	Lexeme	Translation	Word-Type
tˤ-w-l	tˤaːwle	table	noun
x-w-f	xawf	fear	noun
ʔ-s-m	ʔisi̲m	name	noun
ʕ-d-d	mistʕidː	ready	adjective
ʕ-m-l	ʕaːmle	working class	noun
ʕ-m-l	maʕaːmli	employee	noun
ʕ-m-l	ʕimil	he made	noun

A.7 Palestinian

Table A.7: Data from Palestinian

Root	Lexeme	Translation	Word-Type
b-n-k	bank	bank	noun
b-l-tˤ	baːltˤo	coat	noun
b-r-d-ʔ-n	burdʔaːn	orange	noun
d-k-n	duk:aːn	shops	noun
d-r-s	darːas	he taught	verb
d-r-s	dars/dari̲s	lesson	noun
d-r-s	madrasa	school	noun
ʤ-h-l	ʤahil	ignorance	noun
ʤ-n-b	ʤanb	side	noun
ʤ-y-b	ʤaːb	he brought	verb
f-h-m	fihim	heunderstood	verb
f-k	mfakː	screwdriver	noun
f-r-d	fard/fari̲d	pistol	noun
f-x-r	miftxir	proud	adjective
k-l-b	kalb/kali̲b	dog	noun
k-r-m	kari̲m	orchard	noun
k-r-s	kursi	chair	noun
k-t-b	katab	he wrote	verb
k-t-b	ktaːb	book	noun
k-z-b	kizi̲b	lying	noun
n-ʤ-ħ	naːʤiħ	successful	adjective
r-b	rabː	god	noun
r-s-l	raːsal	he corresponded	verb
s-l-m	stalam	he received	verb
s-m-n	sami̲n	shortening	noun
s-t	sitː	grandmother	noun
s-ʔ-l	saʔal	he asked	verb

continued . . .

Root	Lexeme	Translation	Word-Type
ʃ-b-k	ʃubːaːk	window	noun
ʃ-w-f	ʃaːf	he saw	verb
t-m-r	tami̠r	dates	noun
tˤ-l-b	tˤalab	he ordered	verb
ʔ-b-l	ʔabil	he accepted	verb
ʔ-b-l	ʔabi̠l	before	preposition
ʔ-b-n	ʔibi̠n	son	noun
ʔ-m	ʔimː	he removed	verb
ʔ-m	ʔimː	mother	noun
ʔ-x-t	ʔuxt	sister	noun
ʕ-b-d	ʕabi̠d	slave	noun
ʕ-l-m	mʕalːim	teacher	noun
ʕ-l-m	ʕalːam	he taught	verb
ʕ-r-b	ʕarab	Arabs	noun

A.8 Wadi Ramː

Table A.8: Data from Wadi Ramː

Root	Lexeme	Translation	Word-Type
d-h-n	dihin	fat	noun
d-r-s	draːs	harvest	noun
ʤ-y-b	ʤaːb	he brought	verb
ʤ-b-l	ʤibal	mountain	noun
ʤ-b-n	ʤubu̠n	coward	noun
g-s-m	gaːsam	he shared	verb
g-tˤ-m	gitˤam	he broke	verb
ħ-l-b	ħ-l-b	he milked	verb
ħ-r-m	ħurma	wife	noun
k-t-b	kitab	he wrote	verb
l-f-y	lifa	he came	verb
m-r-y	mraːy	mirror	noun
m-ɣ-r	maɣar	he mixed	verb
m-s-k	masak	he caught	verb
n-k-s	nikas	he returned	verb
n-w-l	tanaːwal	he cooperated to catch	verb
r-f-g	raːfag	he accompanied	verb
n-ʕ-j	naʕaja	ewe	noun
r-y	mraːy	mirror	noun
r-y-ʕ	miryaʕ	ram leading the herd	noun

continued ...

Root	Lexeme	Translation	Word-Type
s-b-g	sibag	he foreran	verb
s-k-n	sakan	he lived	verb
sˤ-x-l	sˤaxal	kid	noun
ʃ-d	aʃtadː	he warmed	verb
ʃ-ʕ-r	ʃaʕa̲r	hair	noun
tˤ-l-b	tˤalab	he asked	verb
w-s-m	wasi̲m	brand	noun
ʔ-b-u	ʔabu	father	noun
ʕ-d	astaʕadː	he was ready	verb
ʕ-g-l	ʕagal	he tied	verb
ʕ-l-m	mʕalːim	teacher	noun
ʕ-l-m	mitʕalːim	educated	adjective
ʕ-r-f	ʔ̲i̲mʕaruːf/mʕaruːf	known	adjective
ʕ-ʃ-ʃ	ʔ̲i̲mʕaʃːaːh/mʕaʃːaːh	pastures	noun
ʕ-w-d	ʕaːd	he repeated	verb

A.9 Iraqi

Table A.9: Data from Palestinian

Root	Lexeme	Translation	Word-Type
b-d-l	badːal	he changed	verb
b-n-ʤ	banʤ/bani̲ʤ	anesthetic	noun
b-ħ-r	baħa̲r	sea	noun
b-y-t	beːt	house	noun
d-b-r	dabːar	he arranged	verb
d-r-s	dars	lesson	noun
d-z	dazː	he sent	verb
ʤ-y-b	ʤaːb	he brought	verb
ʤ-m-l	ʔiʤmaːl	camel	noun
f-k-f-k	mfakfuk	having taken apart	adjective
g-b-l	gbaːl	facing	preposition
g-l-b	galu̲b	heart	noun
ˈħ-l	maħalː	place	noun
k-b-r	kbiːr	big	adjective
k-t-b	kitab	he wrote	verb
k-t-b	kitaːb	book	noun
k-t-b	kaːtib	writer	noun
k-w-n	kawn	universe	noun
m-r-y	mara	woman	noun

continued . . .

Root	Lexeme	Translation	Word-Type
n-s-y	nisa	he forgot	verb
q-r-y	qira	he read	verb
q-y-l	mustaqil	independent	adjective
q-w-s	qaws	curve	noun
r-k-b	rtikab	he committed	verb
r-ʔ-s	raːs	head	noun
s-h-l	sahi̲l	easy	adjective
s-l-k	silk/sili̲k	wire	noun
s-m	samːa	he named	verb
ʃ-f-y	mustaʃfa	hospital	noun
w-l-y	wlaːya	city	noun
ʔ-j-y	ʔija	he came	verb
ʔ-r-d	arad	he wanted	verb

A.10 Muscat

Table A.10: Data from Muscat

Root	Lexeme	Translation	Word-Type
b-ɣ-y	bɣa	he wanted	verb
b-n-t	bint	daughter	noun
b-y-t	beːt	house	noun
b-k-y	bakaː	he cried	verb
d-l	dalː	he guided	verb
d-ʃ-d-ʃ	diʃdaʃah	dishdasha	noun
d-x-l	ndixal	he entered	verb
f-k-r	ftakar	he thought	verb
g-l-s	galas	he sat	verb
g-w-h-r	goːhra	jewel	noun
k-b	nkabː	he poured	verb
k-l	kal	he ate	verb
k-l-m	tkalːam	he talked	verb
k-n-s-l	kansal	he cancelled	verb
k-n-z	kanz	pressed dates	noun
k-s-f	kasaf	he hemmed	verb
k-t-b	katab	he wrote	verb
l-b-n	lubaːn	frankincese	verb
l-b-q	labːaq	he turned on	verb
m-r	stamarː	he continued	verb
n-d-y	naːdaː	he called	verb

continued . . .

Root	Lexeme	Translation	Word-Type
q-b-l	qabiːlah	tribe	noun
q-h-w	tqahwaː	he had coffee	verb
q-n-d-l	qandiːla	lantern	noun
r-b-ʕ	mraːbiʕ	accompanying	adjective
ʃ-b	muʃaːbː	handfan	noun
ʃ-l	ʃalː	he carried	verb
t-r-g-m	targam	he translated	verb
t-r-s	ntiras	he was filled	verb
t-sˤ-l	tːasˤal	he got in touch	verb
w-d-y	waːdi	valley	noun
w-r-d	ward	flowers	noun
x-b-z	xubz	bread	noun
x-y-r	xtaːr	he chose	verb
ʕ-ʃ-y	ʕaʃa	supper	noun
ʕ-l-m	tʕalːam	he learnt	verb
ʕ-m-l	istaʕmal	he used	verb
ʕ-ʃ-b	ʕaʃub	grass	noun

A.11 Qatari

Table A.11: Data from Qatari

Root	Lexeme	Translation	Word-Type
b-ʕ-d	baʕd	and also	conjunction
d-f-y	difay	he got warm	verb
d-h-n	dihi̠n	fat/grease	noun
d-w-r	dawːar	he looked for	verb
dˤ-r-b	dˤarab	he hit	verb
f-r-g	farg	difference	noun
f-h-m	fihim	he understood	verb
g-l-m	galam	pen	noun
g-m-ʃ	gmaːʃ	pearls	noun
h-d-h-d	hudhud	kingfisher	noun
ħ-d-g	hidaħ	he fished	verb
ħ-m-r	ħamiːr	donkey	noun
ħ-z-m	ħzaːm	belt	noun
k-b-r	kibra	noun	size
k-l	kilː	all	determiner
k-l-m	takalːam	he talked	verb
k-t	katː	he spilt	verb

continued ...

Root	Lexeme	Translation	Word-Type
k-t-b	ktaːb	book	noun
l-h-m	ilhaːm	he accused	verb
m-θ-l	maθalan	for example	noun
m-r-r	amarː	bitter	adjective
n-ʃ-d	niʃad	he asked	verb
q-n-d-l	qandiːla	lantern	noun
r-d	radː	he returned (sth.)	verb
r-m-l	rami̱l	sand	noun
s-b-t	sabt	Saturday	noun
s-g-f	sagf	ceiling	noun
s-ħ-b	sħib	he pulled	verb
s-m-r	samːar	he hammered	verb
ʃ-h-r	maʃhuːr	famous	adjective
ʃ-r-y	ʃara	he bought	verb
r-y-ʃ	riːʃ	feathers	noun
t-r-s	tiras	he filled up	verb
w-sˤ-l	wisˤal	he arrived	verb
w-z-n	wazi̱n	weight	noun
x-n-f-s	xunfus	beetle	noun
z-tˤ	zatˤ:	he devoured	verb
z-x	zaxː	he touched	verb
ʕ-r-f	ʕiraf	he knew	verb
ʕ-z-m	ʕizam	he invited	verb

A.12 Rufaidah

Table A.12: Data from Rufaidah

Root	Lexeme	Translation	Word-Type
b-n	ʔibn	son	noun
b-n-t	bint	daughter	noun
b-y-t	be:t	house	noun
ð-h-n	ðahn	he woke up	verb
ð-r-b	ðarab	he hit	verb
g-w-l	ga:l	he told	verb
g-b-l	gaːbal	he met	verb
ħ-r-m	ʔaħtaram	he respected	verb
k-l	kala	he ate	verb
k-l-m	kalːam	he spoke	verb
k-l-m	tkalːam	he spoke	verb

continued . . .

Root	Lexeme	Translation	Word-Type
k-s-r	kasːar	he smashed	verb
k-t-b	katab	he wrote	verb
l-b-s	labs	he wore	verb
l-ħ-g	liħig	he caught up	verb
m-d	madː	he spread	verb
m-r	ʔastamarː	he continued	verb
m-ʃ-y	maʃa	he went	verb
n-s-y	nasi	he forgot	verb
r-b-tˤ	rabtˤ	he tied	verb
r-k-b	rakb	he rode	verb
r-m-y	rama	he threw	verb
s-k-b	sakb	he poured	verb
s-l-m	salːam	he handed	verb
s-l-m	ʔastalam	he received	verb
s-m-ʕ	samaʕ	he heard	verb
s-y-r	saːr	he went	verb
ʃ-r-b	ʃarb	he drank	verb
ʃ-r-y	ʔiʃir	he bought	verb
ʃ-w-f	ʃaːf	he saw	verb
ʃ-y-l	ʃaːl	he carried	verb
ʃ-y-r	ʃaːr	he bought	verb
x-ð-y	xaða	he took	verb
x-sˤ-m	mxaːsˤamah	quarrel	noun
x-sˤ-m	txaːsˤam	he had a quarrel	verb
z-r-g	zarrag	he became blue	verb
z-ʕ-l	zaʕl	he became upset	verb

A.13 Rwaili

Table A.13: Data from Rwaili

Root	Lexeme	Translation	Word-Type
b-d-l	badːal	he changed	verb
b-g-r	bgirih	cow	noun
ð-b-ħ	ðbaħ	he killed	verb
ð-r-b	ðarab	he hit	verb
g-b-l	gaːbil	he met	verb
g-r-b	garːab	he drew near	verb
g-y-l	gal	he said	verb
ħ-z-m	ħizim	tie	noun

continued ...

Root	Lexeme	Translation	Word-Type
k-sˤ-m	kasˤːam	he broke	verb
k-t-b	kitab	he wrote	verb
m-d	madː	he spread	verb
m-ʃ-y	miʃa	he went	verb
m-r	astamarː	he continued	verb
n-s-y	nisi	he forgot	verb
a-r-y	ara	he saw	verb
s-f-r	saːfar	he travelled	verb
s-m-ʕ	simiʕ	he heard	verb
ʃ-r-b	ʃirb	he drank	verb
ʃ-r-y	aʃtar	he bought	verb
ʃ-w-f	ʃaf	he saw	verb
ʃ-y-l	ʃal	he carried	verb
w-z-n	wazan	he weighed	verb
x-l	xala	desert	noun
z-m	zamː	he carried	verb
ʕ-m-l	istaʕmil	he used	verb

A.14 Haːyil

Table A.14: Data from Haːyil

Root	Lexeme	Translation	Word-Type
b-d-l	badːal	he changed	verb
b-r	bar	open desert	noun
ð-b-ħ	ðbaħ	he killed	verb
f-sˤ-l	fasˤal	he set apart	verb
ɣ-s-l	ɣasal	he washed	verb
g-b-l	gaːbil	he met	verb
g-r-b	garːab	he drew near	verb
ɣ-n-m	ɣnima	sheep	noun
ɟ-l-s	ɟalas	he sat down	verb
ɟ-m-l	ɟimal	camel	noun
k-l	kala	he ate	verb
k-s-r	kasːar	he smashed	verb
k-t-b	kitab	he wrote	verb
l-g-y	liga	he found	verb
m-d	madː	he spread	verb
m-r	ʔastamarː	he continued	verb
m-ʃ-y	miʃa	he went	verb

continued ...

Root	Lexeme	Translation	Word-Type
n-s-y	nisi	he forgot	verb
s-f-r	saːfir	he travelled	verb
s-m-ʕ	simiʕ	he heard	verb
ʃ-r-b	ʃarab	he drank	verb
ʃ-w-f	ʃaf	he saw	verb
tˤ-l-b	tˤalab	he demanded	verb
ʕ-m-l	staʕmal	he used	verb
w-z-n	wazan	he weighed	verb

A.15 Makːan

Table A.15: Data from Makːan

Root	Lexeme	Translation	Word-Type
b-g-r	bagara	cow	noun
b-w-b	baːb	door	noun
d-r-s	mudarːis	teacher	noun
g-l-a	gala	he fried	verb
g-s-m	magsuːm	destined	adjective
g-t-l	gatal	he killed	verv
k-l-b	kalb	dog	noun
k-l-m	kalːam	he spoke to	verb
k-s-r	kasar	he broke	noun
k-t-b	katab	he wrote	verb
k-t-b	kitaːb	book	noun
k-t-b	maktuːb	written	adjective
m-d	madd	he extended	verb
m-r-y	maːra	he bet/challenged	verb
n-d-y	naːdi	club	noun
r-m-y	rama	he threw	verb
r-w-ħ	raːħ	he went	verb
s-f-r	saːfar	he travelled	verb
sˤ-ħ-b	sˤaːħib	friend	noun
ʃ-k-l	muʃkila	problem	noun
ʃ-r-b	ʃirib	he drank	verb
ʃ-y-l	ʃaːl	he carried	verb
w-d-y	waːdi	valley	noun
ʕ-l-m	tʕalːam	he learnt	verb

A.16 Sˤanʕaːniː

Table A.16: Data from Sˤanʕaːniː

Root	Lexeme	Translation	Word-Type
b-n-t	bint	girl	noun
b-s-tˤ	basaːtˤiːn	gardens	noun
b-y-n	bayn	among	preposition
b-y-ʕ	baːʕ	sold	verb
d-m	dam	blood	noun
d-r-s	madrasih	school	noun
d-r-s	mudarːis	teacher	noun
d-r-y	diri	he knew	verb
d-w-r	dawːir	he looked for	verb
dˤ-r-b	dˤarab	he hit	verb
f-h-n	iftahan	he rested	verb
g-m-b-r	gambar	he sat	verb
g-w-l	gaːl	he said	verb
h-r-b	hirib	he fled	verb
ħ-b	ħabː	he loved	verb
ħ-tˤ-b	ħ tˤab	wood	noun
ħ-ɟ	ħaɟː	pilgrim	noun
j-d	jad	hand	noun
ɟ-l-s	ilis	he sat	verb
ɟ-m-ʕ	ɟamːaʕ	he collected	verb
ɟ-r-r	ɟarː	he took	verb
k-s-b	iktasab	he earned	verb
k-t-b	kitaːb	book	noun
k-t-b	maktuːb	letter	noun
k-t-b	maktaba	library	noun
k-w-n	kaːn	he was	verb
l-b-s	labːas	he dressed	verb
l-f-l-f	laflaf	he collected	verb
r-g-b	ragaba	neck	noun
r-g-ʕ	ragːaʕ	he sew	verb
r-m-y	ramaː	he threw	verb
r-x-sˤ	ʔarxasˤ	cheaper	elative adjective
r-w-ħ	raːħ	he went	verb
s-m	ʔism	name	noun
x-ʃ-b	xʃab	wood	noun
z-w-r	zaːr	he visited	verb
ʔ-b	ʔab	father	noun

Reading is not optional ...

— Walter Dean Myers

References

Abboud-Haggar, Soha (2015). "Syllabic Structure and Stress in Cairene Arabic: A Didactic Approach". In: *Miscelánea de Estudios Árabes y Hebraicos. Sección Árabe-Islam* 64, pp. 19–30.

Abdel-Massih, Ernest, Zaki Abdel-Malik, and El-Said Badawi (1979). *A Comprehensive Study of Egyptian Arabic, Volume 3 (A Preliminary Edition): A Reference Grammar of Egyptian Arabic.* Ann Arbor, MI: Center for Near Eastern and North African Studies, University of Michigan.

Abdelmassih, Ernest (1973). *An Introduction to Moroccan Arabic.* Ann Arbor, MI: The University of Michigan.

Abdo, Daud (1969). *On Stress and Arabic Phonology, A Generative Approach.* Beirut: Khayats.

Abdul-Karim, Kamal Wadih (1980). "Aspects of the Phonology of Lebanese Arabic". PhD thesis. Illinois.

Abu-Haidar, Farida (1991). *Christian Arabic of Baghdad.* Wiesbaden: Harrassowitz.

Abu-Mansour, Mahasen Hasan (1987). "A Nonlinear Analysis of Arabic Syllabic Phonology, with Special Reference to Makkan". en. Ph.D. University of Florida.

— (1992). "Closed-Syllable Shortening and Morphological Levels". In: *Perspectives on Arabic Linguistics IV.* Ed. by Ellen Broselow, Mushira Eid, and John J. McCarthy. Amsterdam & Philadelphia: John Benjamins, pp. 47–75.

Abu-Salim, Issam M (1987). "Vowel Harmony in Palestinian Arabic: A Metrical Perspective". In: *Journal of Linguistics* 23.1, pp. 1–24.

— (1982a). "A Reanalysis of Some Aspects of Arabic Phonology: A Metrical Approach". PhD thesis. University of Illinois.

— (1982b). "Syllable Structure and Syllabification in Palestinian Arabic". In: *Studies in the Linguistic Sciences* 12, pp. 1–28.

Abul-Fadl, Fahmi (1961). *Volkstümliche Texte in Arabischen Bauerndialekten Der Ägyptischen Provinz Sarqiyya.* Münster: Westfälische Wilhelms-Universität.

Abumdas, Abdul Hamid Ali (1985). "Libyan Arabic Phonology". en. PhD thesis. University of Michigan.

Aguadé, Jordi (2006). "Writing Dialect in Morocco". In: *Estudios de dialectologia norteafricana y andalusi* 10, pp. 253–274.

— (2018). "The Maghrebi Dialects of Arabic". In: *Arabic Historical Dialectology.* Ed. by Clive Holes. Oxford: Oxford University Press, pp. 29–63.

Ahmed, Amer (2015). "On Agreement Affixes, Incorporated Pronouns, and Clitics in Standard Arabic". en. In: *SKY Journal of Linguistics* 28, pp. 67–102.

Al Mashaqba, Bassil Mohammad (2015). "The Phonology and Morphology of Wadi Ramm Arabic". DPhil. University of Salford.

Al Sharkawi, Muhammad (2017). *History and Development of the Arabic Language.* London & New York: Routledge.

Al Solami, Majed Abdullah (2020). “Vowel Elision, Epenthesis and Metrical Systems in Bedouin Arabic Dialects”. Ph.D. Toronto, Canada: University of Toronto.

Alber, Birgit, Geert Booij, and J. van de Weijer (1997). “Quantity Sensitivity as the Result of Constraint Interaction”. In: *Phonology in Progress-Progress in Phonology, HIL Phonology Papers III.* The Hague: Holland Academic Graphics.

Albirini, Abdulkafi and Peter Powesland (Nov. 2011). “The Sociolinguistic Functions of Codeswitching between Standard Arabic and Dialectal Arabic”. In: *Language in Society* 40, pp. 537–562.

Alhammad, Reham (2018). “The Role of the Syllable Contact Law-Semisyllable (SCL-SEMI) in the Coda Clusters of Najdi Arabic and Other Languages”. English. Ph.D. United States – Wisconsin: The University of Wisconsin - Milwaukee.

Ali, Marwa (2013). “The Variations of Stem-Final Vowel in Cairene Egyptian Colloquial Arabic: Pre-Suffix Vowel Lengthening Process”. Unpublished Ms. Ain Shams University.

Allen, William Sidney (1973). *Accent and Rhythm.* Vol. 12. Cambridge: Cambridge University Press.

Anderson, Stephen R. (2005). *Aspects of the Theory of Clitics.* Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

— (2011). “Clitics”. In: *The Blackwell Companion to Phonology.* Ed. by Marc van Oostendorp et al. Chichester: Wiley-Blackwell.

Angoujard, Jean-Pierre (1981). “Contribution à l’analyse Prosodique (Parlers de Tunis, Du Caire et de Damas)”. In: *Analyses, Théorie* 1, pp. 66–121.

— (1990). *Metrical Structure of Arabic.* Dordrecht: Foris.

Aoun, Youssef (1979). “Is the Syllable or the Supersyllable a Constituent?” In: *MIT Working Papers in Linguistics* 1, pp. 140–148.

Aoun, Youssef, Elabbas Benmamoun, and Lina Choveiri (2010). *Arabic Syntax.* Cambridge: Cambridge University Press.

Appleyard, David (2003). “Berber Overview [and] Semitic Overview”. In: *Selected Comparative-Historical Afrasian Linguistic Studies.* Ed. by M Bender, David Appleyard, and G Takacs. Germany: Lincom, pp. 23–26.

Aquil, Rajaa (2011). “The Role of the Syllable in the Segmentation of Cairene Spoken Arabic”. In: *Journal of Psycholinguistic Research* 41, pp. 141–158.

Árnason, Kristján (1980). *Quantity in Historical Linguistics.* Cambridge: Cambridge University Press.

Auger, Julie and Richard D. Janda (1994). “Clitics as Affixes and the Process Morphology of Romance Verbs”. In: *Linguistic Society of America Annual Meeting.* Boston.

Azzabi, Ahad M. EL. (2001). “Syllabic Structure of Arabic”. In: *Aspects of Arabic Phonology.* Ed. by A.R. Fatihi. Delhi: Kalinga Publications, pp. 111–115.

Azzajaj (1979). *Maa’niul Qur’an.* Cairo: General Association of Princely Publication.

Bachra, Bernard N. (2001). *The Phonological Structure of the Verbal Roots in Arabic and Hebrew.* Leiden: Brill.

Badawi, El-Said (1973). *Levels of Modern Arabic in Egypt.* Cairo: Dar al-Maarif.

Badawi, El-Said and Kurt R. Jankowsky (1985). “Educated Spoken Arabic: A Problem in Teaching Arabic as a Foreign Language”. In: *Scientific and Humanistic Dimensions of Language: Festschrift for Robert Lado.* Amsterdam: John Benjamins, pp. 15–22.

Bafile, Laura (1994). “La Riassegnazione Postlessicale Dell’accento Nel Napoletano”. In: *Quaderni del dipartimento di linguistica – Università di Firenze* 5, pp. 1–23.

Bagemihi, Bruce (1991). "Syllable Structure in Bella Coola". In: *Proceedings of the New England Linguistics Society* 21, pp. 16–30.

Bamakhramah, Majdi A (2010). "Syllable Structure in Arabic Varieties with a Focus on Superheavy Syllables". DPhil Thesis. Indiana University.

Bassiouney, Reem (2006). *Functions of Code Switching in Egypt.* Leiden & Boston: Brill.

— (2009). *Arabic Sociolinguistics.* Edinburgh: Edinburgh University Press.

Behnstedt, Peter (1987). *Die Dialekte Der Gegend von a'dah (Nord-Jemen).* Wiesbaden: Harrassowitz Verlag.

— (1988). "Die Agyptisch-Arabischen Dialekte, Part 3, Texte II. Niltaldialekte. III. Oasendialekte". In: *Tubinger Atlas des Vorderen Orients, Beihefte, Reihe B: Geistewissenschaften* 50.3.

— (1992). "Die Nordjemenitischen Dialekte. Part 2, Glossar: Alif-Dal". In: *Jemen-Studien* 3, pp. 43–96.

— (1997). *Sprachatlas von Syrien.* Wiesbaden: Harrassowitz Verlag.

Behnstedt, Peter and Manfred Woidich (1985). *Die Ägyptisch-Arabischen Dialekte.* Wiesbaden: Ludwig Riechert.

— (2005). *Arabische Dialektgeographie.* Leiden: Brill.

— (2013). "Arabic Dialectology". In: *The Oxford Handbook of Arabic Linguistics.* Ed. by Jonathan Owens. Oxford: Oxford University Press.

Bell, Alan and Joan B. Hooper (1978). "Issues and Evidence in Syllabic Phonology." In: *Syllable and Segments.* Ed. by Alan Bell and Joan B. Hooper. Amsterdam: North-Holland, pp. 3–22.

Bell, Alan and Mohamad Saka (1983). "Reversed Sonority in Pashto Initial Clusters". In: *Journal of Phonetics* 11, pp. 259–275.

Benkato, Adam (2014). "The Arabic Dialect of Benghazi, Libya: Historical and Comparative Notes". In: *Zeitschrift für Arabische Linguistik* 59, pp. 57–102.

Benmamoun, Elabbas (2011). "Agreement and Cliticization in Arabic Varieties from Diachronic and Synchronic Perspectives". In: *al-'Arabiyya* 44/45, pp. 137–150.

Bennett, Patrick R. (1998). *Comparative Semitic Linguistics: A Manual.* Winona Lake, Indiana: Eisenbrauns.

Bensoukas, Karim (2012). "The Prosody of Moroccan Amazigh and Moroccan Arabic: Similarities in the Phonology of Schwa (2012)". en. In: *Prosody Matters: Essays in Honor of Lisa Selkirk.* Ed. by Toni Borowsky et al. London: Equinox, pp. 3–42.

Bensoukas, Karim and Abdelaziz Boudlal (2012). "An Amazigh Substratum In Moroccan Arabic: The Prosody Of Schwa". en. In: *Langues et Litteratures* 22, pp. 179–221.

Benyagoub, Lahcene (Apr. 2017). "A Moraically-Based Model of the Syllable in Bechar Arabic". en. In: *Arabic Language, Literature & Culture* 2.2, p. 40.

Berendsen, Egon (1986). *The Phonology of Cliticization.* Dordrecht: Foris.

Berent, Iris and Joseph Shimron (2003). "Co-Occurrence Restrictions on Identical Consonants in the Hebrew Lexicon: Are They Due to Similarity?" In: *Journal of Linguistics* 39, pp. 31–55.

Bergsträßer, Gotthelf (1915). "Sprachatlas von Syrien Und Palästina". In: *Zeitschrift des Deutschen Palästina-Vereins* XXXVIII, pp. 169–222.

Birkeland, Harris (1952). *Growth and Structure of the Egyptian Arabic Dialect.* Oslo: I kommisjon Hos Jacob Dybwad.

Bithell, Jethro (1952). *German Pronunciation and Phonology.* London: Methuen & Co.

Black, Alexis (2011). "Ostension and Definiteness in the Kwak'wala Noun Phrase". In: *University of British Columbia Working Papers in Linguistics* 34, pp. 34–52.

Blanc, H. (1960). “Style Variations in Spoken Arabic: A Sample of Inter-Dialectal Conversation”. In: *Contributions to Arabic Linguistics.* Ed. by Charles A Ferguson. Cambridge, MA: Harvard University Press, pp. 81–158.

Blanc, Haim (1964). *Communal Dialects in Baghdad.* Harvard Middle Eastern Monograph Series. Cambridge, MA: Harvard University Press.

Blau, Joshua (1988). *Studies in Middle Arabic and Its Judaeo-Arabic Variety.* Jerusalem: Magnes.

Blevins, Juliet (1995). “The Syllable in Phonological Theory”. In: *Handbook of Phonological Theory.* Ed. by J Goldsmith. London: Basil Blackwell, pp. 206–244.

Bohas, Georges (1978). “Quelques Processus Phonologiques Dans l’Arabe de Damas”. In: *Analyses, Théorie* 1, pp. 87–152.

Bohas, Georges, Jean-Patrick Guillaume, and Djamel Eddine Kouloughli (1990). *The Arabic Linguistic Tradition.* Routledge Library Editions: Language and Literature of the Middle East. Oxford: Routledge.

Booij, Geert (1995). *The Phonology of Dutch.* Oxford: Clarendon Press.

— (1996). “Cliticization as Prosodic Integration: The Case of Dutch”. In: *The Linguistic Review* 13, pp. 219–42.

Börjars, Kersti (1994). “Swedish Double Determination in a European Typological Perspective”. In: *Nordic Journal of Linguistics* 17, pp. 219–252.

— (1998). “Clitics, Affixes and Parallel Correspondence”. In: *Proceedings of the LFG ’98 Conference.* Ed. by Miriam Butt and Tracy Holloway King.

Bosisio, Nicole (2003). “Vocoids and Their Prosodic Districtuion, with Special Reference to Italian and Arabic”. en. Ph. D. Thesis. Durham University.

Boudlal, Abdelaziz (2001). “Constraint Interaction in the Phonology and Morphology of Casablanca Moroccan Arabic”. eng. Ph.D. Thesis. Rabat: Université Mohammed V.

Bouhadiba, Farouk A.N. (1988). “Aspects of Algerian Arabic Verb Phonology and Morphology”. PhD thesis. University of Reading.

Boussofara-Omar, Naima (2003). “Revisiting Arabic Diglossic Switching in Light of the Matrix Language Frame Model and Its Sub-Models: The 4-M and the Abstract Level Models”. In: *Bilingualism: Language and Cognition* 6.1, pp. 33–46.

Brame, Michael K (1971). “Stress in Arabic and Generative Phonology [Review Article]”. In: *Foundations of Language* 7, pp. 556–591.

— (1973). “On Stress in Two Arabic Dialects”. In: *A Festschrift for Morris Halle.* Ed. by Stephen R. Anderson and Paul Kiparsky. New York: Holt, Rinehart & Winston., pp. 14–25.

— (1974). “The Cycle in Phonology: Stress in Palestinian, Maltese, and Spanish”. In: *Linguistic Inquiry* 5.1, pp. 39–60.

Brinkmann, Hennig (1966). *Studien Zur Geschichte Der Deutschen Sprache Und Literature I.* Dusseldorf: Padagogischer Verlag Schwann.

Brockelmann, Carl (1908). *Grundriß Der Vergleichenden Grammatik Der Semitischen Sprachen.* Berlin: Reuther and Reichrad.

Broselow, Ellen (1976). “The Phonology of Egyptian Arabic”. Ph.D. Thesis. University of Massachusetts, Amherst.

— (1992). “Parametric Variation in Arabic Dialect Phonology”. In: *Perspectives on Arabic Linguistics IV.* Ed. by Mushira Eid and John McCarthy. Amsterdam & Philadelphia: John Benjamins Publishing Company, pp. 7–47.

— (June 2001). “Stress-Epenthesis Interactions”. English. In.

Broselow, Ellen, Su-I Chen, and Marie Huffman (1997). “Syllable Weight: Convergence of Phonology and Phonetics”. In: *Phonology* 14.1, pp. 47–82.

Broselow, Ellen, Su-I Chen, Marie Huffman, and Ruohmei Hsieh (1995). “The Timing Structure of CVVC Sylalbles”. In: *Perspectives on Arabic Linguistics VIII*. Ed. by Mushira Eid. Amsterdam & Philadelphia: John Benjamins, pp. 119–138.

Buckley, Eugene (1997). “Tigrinya Root Morphology and the OCP”. In: *University of Pennsylvania Working Papers in Linguistics* 4.3, pp. 19–51.

— (1998). “Alignment in Manam Stress”. In: *Linguistic Inquiry* 29, pp. 476–496.

Burzio, Luigi (1987). “English Stress”. In: *Certamen Phonologicum*. Ed. by P.M. Bertinetto and Michele Loporcaro. Torino: Rosenberg and Sellner.

— (1996). “Surface Constraints versus Underlying Representations”. In: *Current Trends in Phonology: Models and Methods*. Ed. by Jacques Durand and Bernard Laks. Salford, Manchester: European Studies Research Institute, University of Salford, pp. 92–122.

Bye, Patrik (1997). “A Generative Perspective on ’Overlength’ in Estonian and Sammi”. In: *Estonian Prosody: Papers from a Symposium*. Tallinn, estonia, pp. 36–70.

Cadora, F. J. (1992). *Bedouin, Village and Urban Arabic: An Ecolinguistic Study*. Leiden: Brill.

Calabrese, Andrea (2005). *Markedness and Economy in a Derivational Model of Phonology*. Berlin: Mouton de Gruyter.

Cantineau, Jean (1940). *Les Parlers Arabes Du (orân. Atlas*. Paris: Klincksieck.

— (1946). “Esquisse d’une Phonologie de l’arabe Classique”. In: *Bulletin de la Societe de linguistique de Paris* 43, pp. 93–140.

— (1960). *Études de Linguistique Arabe*. Paris: Klincksieck.

Charette, Monik (1985). “The Appendix in Parametric Phonology”. In: *Studies in Affircan Linguistics* 9.Suppl. Pp. 49–53.

Cho, Young-mee Yu and Tracy Holloway King (2003). “Semisyllables and Universal Syllabification”. In: *The Syllable in Optimality Theory*. Ed. by Caroline Fery and Ruben van de Vijver. Cambridge: Cambridge University Press.

Chomsky, Noam and Morris Halle (1968). *The Sound Pattern of English*. New York & London: Harper & Row.

Chomsky, Noam, Morris Halle, and Fred Lukoff (1956). “On Accent and Juncture in English”. In: *For Roman Jakobson: Essays on the Occasion of His Sixtieth Birthday, 11 October, 1956*. Ed. by Morris Halle, Horace Lunt, and H. Maclean. The Hague: Mouton, pp. 65–80.

Chtatou, Mohamed (2009). “The Influence of the Berber Language on Moroccan Arabic”. In: *International Journal of the Sociology of Language* 123.1, pp. 101–118.

Clements, George Nickerson (1987). “Phonological Feature Representation and the Description of Intrusive Stops.” In: *Papers from the Annual Regional Meeting, Chicago Linguistic Society* 23.2, pp. 29–50.

— (1990). “The Role of the Sonority Cycle in Core Syllabification”. In: *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Ed. by John Kingston and Mary Beckman. Cambridge: Cambridge University Press, pp. 283–333.

Clements, George Nickerson and Samuel Jay Keyser (1983). *CV Phonology*. Cambridge, MA: MIT Press.

Cohen, David (1970). “Koine, Langues Communes et Dialectes Arabes”. In: *Etudes de Linguistique Semitiques et Arabe*. Ed. by David Cohen. The Hague and Paris: Mouton, pp. 105–125.

Corriente, Federico (2008). *Romania Arábica: Tres Cuestiones Básicas: Arabismos, "Mozárabe" y "Jarchas"*. Madrid: Editorial Trotta.
Crowhurst, Megan and Mark Hewitt (1994). "Directional Footing, Degeneracy, and Alignment". In: *ROA-65*.
Crowhurst, Megan and Lev Michael (2005). "Iterative Footing and Prominence-Driven Stress in Nanti (Kampa)". In: *Language* 81.1, pp. 47–95.
Cutler, Anne et al. (1986). "The Syllable's Different Role in Segmentation of French and English". In: *Journal of Memory and Language* 25, pp. 385–400.
Danecki, Janusz (2018). "Idġām". In: *Encyclopedia Arabic Language and Linguistics*.
Davis, Stuart (1999). "On the Representation of Initial Geminates". In: *Phonology* 16, pp. 93–104.
— (2011). "Geminates". In: *The Blackwell Companion to Phonology*. Ed. by Marc van Oostendorp et al. Oxford: Wiley Blackwell, pp. 873–897.
Davis, Stuart and Marwa Ragheb (2014). "Geminate Representation in Arabic". In: *Perspectives on Arabic Linguistics XXIV-XXV*. Amsterdam & Philadelphia: John Benjamins, pp. 3–19.
Davis, Stuart and Seung-Hoon Shin (1999). "The Syllable Contact Constraint in Korean: An Optimality-theoretic Analysis." In: *Journal of East Asian Linguistics* 8, pp. 285–312.
Dell, Francois (1995). "Consonant Clusters and Phonological Syllables in French". In: *Lingua* 95, pp. 5–26.
Diem, Werner (1979). "Studien Zur Frage Des Substrats Im Arabischen". In: *Der Islam* LVI.1, pp. 12–80.
Dixon, R. M. W. (2002). "The Eclectic Morphology of Jarawara, and the Status of Word". In: *Word: A Cross-Linguistic Typology*. Ed. by R. M. W. Dixon and Alexandra Y. Aikhenvald. Cambridge: Cambridge University Press, pp. 125–152.
Donegan, Patricia J. and David Stampe (1983). "Rhythm and the Holistic Organisation of Language Structure". In: *Papers from the Parasession on the Interplay of Phonology, Morphology and Syntax*. Ed. by John F. Richardson, Mitchell Marks, and Amy Chukerman. Chicago: Chicago Linguistics Society, pp. 337–353.
Eberhard, David M., Gary F. Simons, and Charles D. Fennig, eds. (2020). *Ethnologue: Languages of the World*. en. 23rd. Dallas, Texas: SIL International.
Edzard, Lutz (1998). *Polygenesis, Convergence and Entropy: An Alternative Model of Linguistic Evolution Applied to Semitic Linguistics*. Wiesbaden: Harrassowitz.
Eid, Mushira (1982). "The Non-Randomness of Diglossic Variation in Arabic". In: *Glossa* 16.1, pp. 54–84.
— (1988). "Principles for Code-Switching between Standard and Egyptian Arabic". In: *Al-'Arabiyya* 21.1-2, pp. 51–79.
Eisner, Jason (1997). "Efficient Generation in Primitive Optimality Theory". In: *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*. East Stroudsburg, PA: Association for Computational Linguistics, pp. 313–320.
El Aissati, Abderrahman (2006). "Berber Loanwords". In: *Encyclopedia of Arabic Language and Linguistics* I, pp. 293–9.
El Zarka, Dina (2018). "Reduplication". In: *Encyclopedia Arabic Language and Linguistics*. Ed. by Edward Lutz and Rudolf de Jong.
Elenbaas, Nine (1999). *A Unified Account of Binary and Ternary Stress*. Utrecht: Netherlands Graduate School of Linguistics.

Embick, David and Robert Noyer (2001). "Movement Operations after Syntax". In: *Linguistic Inquiry* 32, pp. 555–595.
Erwin, Wallace. M. (1963). *A Short Reference Grammar of Iraqi Arabic*. Washington, DC: Georgetown University Press.
Faber, Alice (1997). "Genetic Subgrouping of the Semitic Languages". In: *The Semitic Languages*. Ed. by Robert Hetzron. London: Routledge, pp. 3–15.
Farrag, Nesrine Saad, Lawrence J Cheskin, and Mohamed Farag (2017). "A Systematic Review of Childhood Obesity in the Middle East and North Africa (MENA) Region: Health Impact and Management". In: *Advances in Pediatric Research*.
Farwaneh, Samira (1995). "Directionality Effects in Arabic Dialect Syllable Structure". PhD thesis. University of Utah.
— (1997). "Augmentation and Reduction Domains in Arabic: The Role of Identity". en. In: *North East Linguistics Society* 27, p. 11.
— (2009). "Towards a Typology of Arabic Dialects: The Role of Final Consonantality". In: *Journal of Arabic and Islamic Studies* 9, pp. 82–109.
Fassi Fehri, Abdelkader (1990). *al-Binaau l-Muwaazi: Naariytun Fii Binaai l-Kalimati Wa l-Jumlati [The Parallel Structure: A Theory of Word and Sentence Structures]*. Casablanca: Tuubqaal Publishing House.
— (1993). *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.
Faust, Noam and Shanti Ulfsbjorninn (2018). "Arabic Stress in Strict CV, with No Moras, No Syllables, No Feet and No Extrametricality". en. In: *The Linguistic Review* 35.4, pp. 561–600.
Ferguson, Charles A. (1959a). "Diglossia". In: *Word* 15, pp. 325–340.
— (1959b). "The Arabic Koine". In: *Language* 35.4, pp. 616–630.
Féry, Caroline and Ruben van de Vijver (2003). *The Syllable in Optimality Theory*. New York, UNITED STATES: Cambridge University Press.
Fischer, Wolfdietrich and Otto Jastrow, eds. (1980). *Handbuch Der Arabischen Dialekte*. Wiesbaden: Harrassowitz.
Fischer-Jorgensen, Eli (1952). "On the Definition of Phoneme Categories on a Distributional Basis". In: *Acta Linguistica*, pp. 8–39.
Frisch, Stefan, Michael Broe, and Janet Breckenridge Pierrehumbert (1997). "Similarity and Phonotactics in Arabic". University of Michigan, Ann Arbor.
Frisch, Stefan, Janet Breckenridge Pierrehumbert, and Michael Broe (2004). "Similarity Avoidance and the OCP". In: *Natural Language & Linguistic Theory* 22.1, pp. 179–228.
Fück, Johann (1950). *Arabiya. Untersuchungen Zur Arabischen Sprach- Und Stilgeschichte*. Berlin: Akademie-Verlag.
Gairdner, William Henry Temple (1925). *The Phonetics of Arabic*. Oxford: Oxford University Press.
Garbell, Irene (1958). "Remarks on the Historical Phonology of an Eastern Mediterranean Dialect". In: *Word* 14, pp. 303–337.
Garbini, Giovanni (1984). *Le Lingue Semitiche: Studi Di Storia Linguistica*. 2nd editio. Naples: Istituto Universitario Orientale.
Giegerich, Heinz J. (1985). *Metrical Phonology and Phonological Structure: German and English*. Cambridge: Cambridge University Press.
Glover, Bonnie Carol (1988). "The Morphophonology of Muscat Arabic". English. Ph.D. United States – California: University of California, Los Angeles.
Golston, Chris (1995). "Direct Optimality Theory: Representation as Constraint Violation". Heinrich-Heine Universitat Dusseldorf.

Golston, Chris and Tomas Riad (1997). “The Phonology of Classical Arabic Meter”. In: *Linguistics* 35, pp. 111–132.
Gordon, Matthew K. (2002). “A Factorial Typology of Quantity-Insensitive Stress”. In: *Natural Language & Linguistic Theory* 20, pp. 491–552.
— (2006). *Syllable Weight: Phonetics, Phonology and Typology.* Routledge.
— (2014). “Disentangling Stress and Pitch-Accent: A Typology of Prominence at Different Prosodic Levels”. In: *Word Stress: Theoretical and Typological Issues.* Ed. by Harry van der Hulst. Cambridge: Cambridge University Press.
Green, Antony D. (2005). “Word, Foot and Syllable Structure in Burmese Syntax”. In: *Studies in Burmese Linguistics* 570. Ed. by J. Watkins, pp. 1–26.
Green, Thomas (1995). “The Stress Window in Pirahã: A Reanalysis of Rhythm in Optimality Theory”. en. Unpublished Ms. MIT.
Greenberg, Joseph Harold (1950). “The Patterning of Root Morphemes in Semitic”. In: *Word* 6, pp. 162–181.
Gussenhoven, Carlos (1986a). “English Plosive Allophones and Ambisyllabicity”. In: *Grammars* 10.2.
— (1986b). “Over de Fonologie van Nederlandse Clitica”. In: *Spektator* 15, pp. 180–200.
Haddad, Ghasan (1984). “Problems and Issues in the Phonology of Lebanese Arabic”. English. Ph.D. United States – Illinois: University of Illinois at Urbana-Champaign.
Haelewyck, Jean-Claude (2007). *Grammaire Comparée Des Langues Sémitiques.* Brussels: Safran.
Hall, Daniel Currie (2007). “The Role and Representation of Contrast in Phonological Theory (2007)”. en. In: *Toronto Working Papers in Linguistics.*
Hall, Tracy Alan (1999). “Phonotactics and the Prosodic Structure in German Function Words”. In: *Studies on the Phonological Word.* Ed. by Tracy Alan Hall and Ursula Kleinhenz. Amsterdam & Philadelphia: John Benjamins, pp. 99–131.
Halle, Morris and Jean-Roger Vergnaud (1979). *Metrical Phonology.* Tech. rep. MIT.
— (1987). *An Essay on Stress.* Cambridge, MA: MIT Press.
Hamid, Abdel Halim (1984). “A Descriptive Analysis of Sudanese Colloquial Arabic Phonology”. Ph.D. Dissertation. University of Illinois.
Hammond, Michael (1995). “There Is No Lexicon!” en. University of Arizona, Tucson.
Hannahs, Stephen J. (1995). “The Phonological Word in French”. In: *Linguistics* 33, pp. 1125–1144.
Harrell, Richard S. (1962). *A Short Reference Grammar of Moroccan Arabic.* Washington, DC: Georgetown University Press.
Hayes, Bruce (1981). “A Metrical Theory of Stress Rules”. PhD Thesis. MIT.
— (1982). “Extrametricality and English Stress”. In: *Linguistic Inquiry* 13, pp. 227–276.
— (1984). “The Phonology of Rhythm in English”. In: *Linguistic Inquiry* 15, pp. 33–74.
— (1986). “Inalterability in CV Phonology”. In: *Language* 62.2, pp. 321–351.
— (1987). “A Revised Parametric Metrical Theory”. In: *North East Linguistics Society* 17.1, pp. 274–289.
— (1989). “Compensatory Lengthening in Moraic Phonology”. In: *Linguistic Inquiry* 20, pp. 253–306.
— (1995). *Metrical Stress Theory: Principles and Case Studies.* London: The University of Chicago Press, Ltd.
Herr, Kristen Elizabeth (2011). “The Phonological Interpretation of Minor Syllables, Applied to Lemi Chin”. en. MA thesis. Chaing Mai, Thailand: Payap University.

Hetzron, Robert (1974). “La Division Des Langues Sémitiques”. In: *Actes Du Premier Congrès International de Linguistique Sémitique et Chamito-Sémitque, Paris 16 19 Juillet, 1969*. Ed. by A Caquot and David Cohen. The Hague, Paris: Mouton, pp. 181–194.
— (1975). “Genetic Classification and Ethiopian Semitic”. In: *Hamito-Semitica*. Ed. by J Bynon and T Bynon. The Hague: Mouton, pp. 103–127.
— (1976). “Two Principles of Genetic Reconstruction”. In: *Lingua* 38, pp. 89–108.
Hockett, Charles (1955). *Manual of Phonology*. Baltimore: Waverly Press.
Holes, Clive (2004). *Modern Arabic: Structures, Functions and Varieties*. Washington, DC: University of Georgetown Press.
— ed. (2018). *Arabic Historical Dialectology: Linguistic and Sociolinguistic Approaches*. Oxford: Oxford University Press.
Holt, D. Eric (1997). “The Role of the Listener in the Historical Phonology of Spanish and Portuguese: An Optimality-theoretic Account”. Ph.D. Dissertation. Georgetown University.
Hooper, Joan B. (1976). *An Introduction to Natural Generative Phonology*. New York: Academic Press.
Hwaidi, Tamader Hassan (2016). “Syllable Structure and Syllabification in Al’ain Libyan Arabic”. D.Phil. Newcastle University.
Hyman, Larry M. (1985). *A Theory of Phonological Weight*. Dordrecht: Foris.
Ibn Jinnii, Abul-Fath (1970). *A. Al-Khasā’is*. Ed. by Mohammad ’Ali L-Najjaar. Beirut: Dar al-Huda.
Idsardi, William J. (1998). “Tiberian Hebrew Spirantization and Phonological Derivations”. In: *Linguistic Inquiry* 29.1, pp. 37–73.
Inkelas, Sharon (1989). “Prosodic Constituency in the Lexicon”. Ph.D. Stanford University.
Itô, Junko (1986). “Syllable Theory in Prosodic Phonology”. Ph.D. University of Massachusetts, Amherst.
— (1989). “A Prosodic Theory of Epenthesis”. In: *Natural language and linguistic theory* 7.2, pp. 217–260.
Jakobson, Roman (1957). “Mufaxxama: The ‘emphatic’ Phonemes in Arabic”. In: *Selected Writings 1: Phonological Studies*. The Hague: Mouton, pp. 510–522.
— (1971). *Selected Writings. Vol. 1: Phonological Studies*. The Hague: Mouton.
Jamary, Abderrahim (1993). “Clitic Phenomena in Arabic”. Ph.D. University of London.
Jarrar, Mustafa et al. (2017). “Curras: An Annotated Corpus for the Palestinian Arabic Dialect”. In: *Language Resources and Evaluation* 51.3, pp. 745–775.
Jespersen, Otto (1904). *Lehrbuch Der Phonetik*. Leipzig and Berlin: B.G. Teubner.
Kabrah, Rawiah S. (2004). “Opacity and Transparency in the Phonology of Makkan Arabic: A Stratal Optimality -Theoretic Approach”. English. Ph.D. United States – Massachusetts: Boston University.
Kager, René (1989). *A Metrical Theory of Stress and Destressing in English and Dutch*. Dordrecht: Foris.
— (1995a). “Consequences of Catalexis”. In: *Leiden in Last: HIL Phonology Papers I*. Ed. by Harry van der Hulst and J. van de Weijer. The Hague: Holland Academic Graphics, pp. 269–298.
— (1995b). “The Metrical Theory of Word Stress”. In: *The Handbook of Phonological Theory*. Ed. by John Goldsmith. First. Oxford: Blackwell, pp. 205–244.
— (1996). “On Affix Allomorphy and Syllable Counting”. In: *Interfaces in Phonology*. Ed. by Ursula Kleinhenz. Studia Grammatica 41. Berlin: Akademie-Verlag.

Kager, René (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
— (2007). “Feet and Metrical Stress”. In: *Cambridge Handbook of Phonology*. Ed. by Paul De Lacy. Cambridge: Cambridge University Press, pp. 195–228.
Kanerva, Jonni (1987). “Morphological Integrity and Syntax: The Evidence from Finnish Possessive Suffixes.” In: *Language* 63.3, pp. 498–521.
Kang, Ongmi (1992). “Korean Prosodic Phonology”. PhD Thesis. University of Washington.
Kaye, Alan S (1991). “The Hamzat Al-Wal in Contemporary Modern Standard Arabic”. In: *Journal of the American Oriental Society* 111.3, pp. 572–574.
Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud (1990). “Constituent Structure and Government in Phonology”. In: *Phonology* 7, pp. 193–232.
Kelkar, Ashok Ramchandra (1968). *Studies in Hindi-Urdu*. Poona: Deccan College.
Kenstowicz, Michael (1970). “On the Rotation of Vowel Length in Lithuanian”. In: *Paper in Linguistics* 3.1, pp. 73–113.
— (1986). “Notes on Syllable Structure in Three Arabic Dialects”. In: *Revue quebecoise de linguistique* 16.1, pp. 101–127.
— (1991). “On Metrical Constituents: Unbalanced Trochees and Degenerate Feet”. Unpublished Ms. MIT.
— (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.
— (1996). “Base Identity and Uniform Exponence: Alternatives to Cyclicity”. In: *Current Trends in Phonology: Models and Methods*. Ed. by Jacques Durand and Bernard Laks. Salford: European Studies Research Institute, University of Salford, pp. 363–394.
Kenstowicz, Michael and Charles Kisseberth (1979). *Generative Phonology: Description and Theory*. New York & London: Academic Press.
Kenstowicz, Michael and Charles Pyle (1973). “On the Phonological Integrity of Geminate Clusters”. In: *Issues in Phonological Theory*, pp. 27–43.
Keydana, Götz (2012). “Evidence for Non-Linear Phonological Structure in Indo-European: The Case of Fricative Clusters”. In: *The Sound of Indo-European, Phonetics, Phonemics and Morphophemics*. Ed. by Benedicte Nielsen Whitehead et al. Vol. 4. Copenhagen Studies in Indo-European. Copenhagen.
Khattab, Ghada and Jalal Al-Tamimi (2014). “Geminate Timing in Lebanese Arabic: The Relationship between Phonetic Timing and Phonological Structure”. In: *Laboratory Phonology* 5.2, pp. 231–269.
Kiparsky, Paul (1977). “The Rhythmic Structure of English Verse”. In: *Linguistic Inquiry* 8, pp. 189–247.
— (1982). “From Cyclic Phonology to Lexical Phonology”. In: *The Structure of Phonological Representations. Vol.1*. Ed. by Harry van der Hulst and Norval Smith. Dordrecht: Foris, pp. 131–175.
— (1991). “Catalexis”. Unpublished Ms. Stanford University and Wissenschaftskolleg zu Berlin.
— (2000). “Opacity and Cyclicity”. In: *The Linguistic Review* 17, pp. 351–367.
— (2003). “Syllables and Moras in Arabic”. In: *The Syllable in Optimality Theory*. Ed. by Caroline Fery and Ruben van de Vijver. Cambridge: Cambridge University Press, pp. 147–182.
Klavans, Judith (1985). “The Independence of Syntax and Phonology in Cliticization”. In: *Language* 61, pp. 95–120.
Koshal, Sanyukta (1979). *Ladakhi Grammar*. Delhi: Motilal Banarsidass.

Kraehenmann, Astrid (2003). *Quantity and Prosodic Asymmetries in Alemannic: Synchronic and Diachronic Perspectives.* Berlin/Boston, Germany: De Gruyter, Inc.

Kubozono, Harao (1989). "The Mora and Syllable Structure in Japanese: Evidence from Speech Errors". In: *Language and Speech* 32, pp. 249–278.

— (1999). "Mora and Syllable". In: *The Handbook of Japanese Linguistics.* Ed. by Natsuko Tsujimura. Oxford: Wiley-Blackwell.

Labov, William (2007). "Transmission and Diffusion". In: *Language* 83, pp. 344–387.

Lahiri, Aditi (2001). "Metrical Patterns". In: *Language Typology and Language Universals.* Ed. by E König and Martin Haspelmath. Berlin: Mouton, pp. 1347–1367.

Lahiri, Aditi and Jennifer Fitzpatrick-Cole (1999). "Emphatic Clitics in Bengali". In: *Phrasal Phonology.* Ed. by René Kager and Wim Zonneveld. Dordrecht: Foris.

Lahiri, Aditi and Jacques Koreman (1988). "Syllable Quantity and Stress in Dutch". In: *Proceedings of the West Coast Conference on Formal Linguistics* VII, pp. 217–228.

Lahiri, Aditi and William Marslen-Wilson (1992). "Lexical Processing and Phonological Representation". In: *Gesture, Segment, Prosody.* Ed. by D. Robert Ladd and Gerard J. Docherty. Papers in Laboratory Phonology. Cambridge: Cambridge University Press, pp. 229–260.

Lahiri, Aditi and Frans Plank (2010). "Phonological Phrasing in Germanic: The Judgement of History, Confirmed through Experiment". In: *Transactions of the Philological Society* 108.3, pp. 370–398.

— (to appear). "Phonological Phrasing". In: *The Handbook of the History of Phonology.* Ed. by B. Elan Dresher and Harry van der Hulst.

Lahiri, Aditi, Allison Wetterlin, and Elisabet Jönsson-Steiner (2005). "Sounds Definite-Ly Clitic: Evidence from Scandinavian Tone". In: *Lingue e Linguaggio* IV, pp. 243–262.

Lahrouchi, Mohamed (2018). "Syllable Structure and Vowel/Zero Alternations in Moroccan Arabic and Berber". In: *The Routledge Handbook of African Linguistics.* Ed. by Augustine Agwele and Adams Bodomo. 1st. London: Routledge-Taylor & Francis Group, pp. 168–180.

Lahrouchi, Mohamed and Rachid Ridouane (2016). "On Diminutives and Plurals in Moroccan Arabic". In: *Morphology* 26.3-4, pp. 453–475.

Laks, Lior (2011). "Morpho-Phonological and Morpho-Thematic Relations in Hebrew and Arabic Verb Formation". PhD thesis. Tel Aviv University.

Larcher, Pierre (2001). "Le Parler Des Arabes de Cyrenaique vu Par Un Voyageur Marocain Du XIIIe Siecle". In: *Arabica* 48, pp. 368–82.

Leben, William (1980). "A Metrical Analysis of Length". In: *Linguistic Inquiry* 11.3, pp. 479–509.

Lecerf, Jean (1932). "Litterature Dialectale et Renaissance Arabe Moderne". In: *Bulletin d'etudes orientales de l'Institut Francais de Damas.*

Lehiste, Ilse (1966). *Consonant Quality and Phonological Units in Estonian.* The Hague: Mouton.

Levin, Juliette (1985). "A Metrical Theory of Syllabicity". Ph.D. Cambridge, MA: Massachusetts Institute of Technology.

Liberman, Mark (1975). "The Intonational System of English". PhD Thesis. MIT.

Liberman, Mark and Alan Prince (1977). "On Stress and Linguistic Rhythm". In: *Linguistic Inquiry* 8.2, pp. 249–336.

Lipiński, Edward (2001). *Semitic Languages: Outline of a Comparative Grammar.* Second. Leuven: Peeters.

Loporcaro, Michele (2000). “Stress Stability under Cliticization and the Prosodic Status of Romance Clitics”. In: *Phonological Theory and the Dialects of Italy.* Ed. by Lori Repetti. Philadelphia & Amsterdam: John Benjamins, pp. 137–168.

Lowe, John J. (2016). “English Possessive ’s: Clitic and Affix”. en. In: *Natural Language & Linguistic Theory* 34.1, pp. 157–195.

Luis, Ana R. (2004). “Clitics as Morphology”. PhD thesis. University of Essex.

Maamouri, Mohamed (1967). “The Phonology of Tunisian Arabic”. English. Ph.D. United States – New York: Cornell University.

MacDonald, Lorna (1990). *A Grammar of Tauya.* Berlin: Mouton de Gruyter.

Majdi, Basim Badi (1988). “Iraqi Arabic Morphophonemics”. English. Ph.D. United States – Connecticut: University of Connecticut.

Mansour, Jacob (1991). *The Jewish Baghdadi Dialect: Studies and Texts in the Judaeo-Arabic Dialect of Baghdad.* Or-Yehuda: The Babylonian Jewry Heritage Centre.

Marçais, William (1930). “La Diglossie Arabe”. In: *L’Enseignement Public - Revue Pedagogique* 104.12, pp. 401–409.

— (1938). “Comment l’Afrique Du Nord a Été Arabisée”. In: *Annales de l’Institut d’Etudes Orientales d’Alger* 4, pp. 1–22.

Martinet, A (1955). *Ėconomie Des Changements Phonétiques.* Francke: Berne.

Mashaqba, Bassil and Anas Huneety (2018). “Emergence of Iambs in Eastern Arabic: Metrical Iambicity Dominating Optimal Nonfinality”. en. In: *SKASE Journal of Theoretical Linguistics* 15.3, p. 22.

McCarthy, John J (2007). “What Is Optimality Theory?” en. In: *Language and Linguistics Compass* 93, p. 29.

— (1979). “On Stress and Syllabification”. In: *Linguistic Inquiry* 10.3, pp. 443–465.

— (1981). “A Prosodic Theory of Nonconcatenative Morphology”. In: *Linguistic inquiry* 12.January, pp. 373–418.

— (1986). “OCP Effects: Gemination and Antigemination”. In: *Linguistic Inquiry* 17, pp. 207–263.

— (1993). “Template Form in Prosodic Morphology”. In: *FLSM III: Papers from the Third Annual Meeting of the Formal Linguistics Society of Midamerica.* Ed. by Laurel Smith Stvan. Evanston, Illinois: Bloomington, Ind.

— (1994). “The Phonetics and Phonology of Semitic Pharyngeals”. In: *Papers in Laboratory Phonology III.* Ed. by Patricia A. Keating. Cambridge: Cambridge University Press.

— (2003). “Metrical Phonology”. In: *International Encyclopedia of Linguistics.* Ed. by William J. Frawley. Second. Oxford: Oxford University Press, pp. 54–57.

— (2005). “The Length of Stem-Final Vowels in Colloquial Arabic”. In: *Perspectives on Arabic Linguistics XVII-XVIII.* Ed. by Mohammad T Alhawary and Elabbas Benmamoun. Amsterdam: John Benjamins Publishing Company.

McCarthy, John J. and Alan Prince (1990a). “Prosodic Morphology and Templatic Morphology”. In: *Perspectives on Arabic Linguistics II.* Ed. by Mushira Eid and John J. McCarthy. Amsterdam & Philadelphia: John Benjamins Publishing Company, pp. 1–54.

— (1990b). “Prosodic Morphology and Templatic Morphology”. en. In: *Current Issues in Linguistic Theory.* Ed. by Mushira Eid and John J. McCarthy. Vol. 72. Amsterdam: John Benjamins Publishing Company, p. 1.

— (1998). “Prosodic Morphology”. In: *The Handbook of Morphology.* Ed. by Andrew Spencer and Arnold Zwicky. Oxford & Malden, MA: Blackwell, pp. 283–305.

Mehler, J. et al. (1981). "The Syllable's Role in Speech Segmentation". In: *Journal of Verbal Learning and Verbal Behavior* 20, pp. 298–305.

Meiseles, Gustav (1980). "Educated Spoken Arabic and the Arabic Language Continuum". In: *Archivum Linguisticum* 11, pp. 118–143.

Mejdell, Gunvor (2018). "Diglossia". In: *The Routledge Handbook of Arabic Linguistics*. Ed. by Reem Bassiouney and Elabbas Benmamoun. London: Routledge, pp. 332–344.

Mester, Armin (1994). "The Quantitative Trochee in Latin". In: *Natural Language and Linguistic Theory* 12, pp. 1–61.

Mester, Armin and Jaye Padgett (1994). "Directional Syllabification in Generalised Alignment". In: *Phonology at Santa Cruz 3*. Ed. by Jason Merchant, Jaye Padgett, and Rachel Walker. Vol. 3, pp. 79–85.

Miller, Catherine (2007). "Arabic Urban Vernaculars: Development and Change". In: *Arabic in the City*. London: Routledge-Taylor, pp. 1–30.

Miller, Wick R. (1965). *Acoma Grammar and Texts*. Berkely & Los Angeles: University of California Press.

Mitchell, Terence Frederick (1960). "Prominence and Syllabification in Arabic". In: *Bulletin of the School of Oriental and African Studies* 23, pp. 369–389.

Mohammad, Mohammad (1990). "The Problem of Subject-Verb Agreement in Arabic: Towards a Solution". In: *Perspectives on Arabic Linguistics*. Ed. by Mushira Eid. I. Amsterdam: John Benjamins, pp. 95–125.

Mohanan, K.P. and Tara Mohanan (1984). "Lexical Phonology of the Consonant System in Malayalam". In: *Linguistic Inquiry* 15.4, pp. 575–602.

Al-Mohanna, Faisal (2004). *Paradoxical Non-Finality: Stress Assignment in Three Arabic Dialects*. Tech. rep. King Saud University.

Morelli, Frida (2003). "The Relative Harmony of /S+Stop/ Onsets: Obstruent Clusters and the Sonority Sequencing Principle". In: *The Syllable in Optimality Theory*. Ed. by C. Fery Caroline and Ruben van de Vijver. New York: Cambridge University Press, pp. 356–371.

Moren, Bruce (2000). "The Puzzle of Kashmiri Stress". In: *Phonology* 17, pp. 365–396.

Murray, Robert W. and Theo Vennemann (1983). "Sound Change and Syllable Structure in Germanic Phonology." In: *Language* 59, pp. 514–528.

Mustawafi, Eiman (2018). "Arabic Phonology". In: *The Routledge Handbook of Arabic Linguistics*. Ed. by Reem Bassiouney and Elabbas Benmamoun. London: Routledge, pp. 11–31.

Nebes, Norbert (1994). "Zur Form Der Imperfektbasis Des Unvermehrten Grundstammes Im Altsüdarabischen". In: *Festschrift Ewald Wagner Zum 65. Geburtstag, Vol. 1: Semitische Studien Unter Besonderer Berücksichtigung Der Südsemitistik*. Ed. by W Heinrichs and G Schoeller. Beirut, Stuttgart: F. Steiner, pp. 59–81.

Nespor, Marina and Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.

Nevins, Andrew and Bert Vaux (2003). "Metalinguistic Shmetalinguistic: The Phonology of Shm-Reduplication". en. In: *Proceedings from the Annual Meeting of the Chicago Linguistic Society*. Vol. 39. Chicago Linguistic Society, pp. 702–721.

Nevis, Joel (1986). "Finnish Particle Clitics and General Clitic Theory". PhD thesis. The Ohio State University.

Newman, Stanley (1947). "Bella Coola I: Phonology". In: *International Journal of American Linguistics* 13, pp. 129–134.

Ngonyani, Deo (2006). “Resumptive Pronominal Clitics in Bantu Languages”. In: *Selected Proceedings of the 36th Annual Conference on African Linguistics.* Ed. by Olaoba F. Arasanyin and Michael F. Pemberton. Somerville, MA: Cascadilla Proceedings Project.
Nöldeke, Theodor (1899). *Die Semitischen Sprachen: Eine Skizze.* Leipzig: Tauchnitz.
— (1911). “Semitic Languages”. In: *The Encyclopaedia Britannica, Vol. 24.* London: Encyclopaedia Britannica, pp. 617–630.
Nouveau, Dominique (Jan. 1994). “Catalexis and Extrametricality in Dutch Word Stress”. en. In: *The Linguistic Review* 11.3-4, pp. 315–328.
Obrecht, Dean (1965). “Three Experiments in the Perception of Geminate Consonants in Arabic”. In: *Language and Speech* 8.1, pp. 31–41.
Odden, David (1978). “Aspects of Iraqi Verbal Phonology”. In: *Studies in the Linguistic Sciences* 8.1, pp. 137–152.
Owens, Jonathan (2003). “Arabic Dialect History and Historical Linguistic Mythology”. In: *Journal of the American Oriental Society* 123.4, pp. 715–740.
— (2006). *A Linguistic History of Arabic.* Oxford: Oxford University Press.
Palva, H (2006). “Dialects: Classification”. In: *Encyclopedia of Arabic Language and Linguistics Vol.1.* Ed. by K. Versteegh. Vol. 1. Leiden: Brill, pp. 604–613.
Paradis, Carole (1988). “On Constraints and Repair Strategies”. In: *The Linguistic Review* 6, pp. 71–96.
Paster, Mary (2013). “Rethinking the ‘Duplication Problem’”. en. In: *Lingua* 126, pp. 78–91.
Peperkamp, Sharon Andrea (1997). “Prosodic Words”. PhD Thesis. University of Amsterdam.
— (2004). “Lexical Exceptions in Stress Systems: Arguments from Early Language Acquisition and Adult Speech Perception”. en. In: *Language* 80.1, pp. 98–126.
Pierrehumbert, Janet Breckenridge (1980). “The Finnish Possessive Suffixes”. In: *Language* 56, pp. 603–21.
— (1993). “Dissimilarity in Arabic Verbal Roots”. In: *Proceedings of NELS 23.* Amherst, MA: Department of Linguistics, University of Massachusetts, pp. 367–381.
Piggott, Glyne (1999). “At the Right Edge of Words”. In: *The Linguistic Review* 16, pp. 143–185.
Piggott, Glyne and Rajendra Singh (1985). “The Phonology of Epenthetic Segments”. In: *Canadian Journal of Linguistics* 30, pp. 415–451.
Porkhomovsky, Victor (1997). “Modern South Arabian Languages from a Semitic and Hamito-Semitic Perspective”. In: *Proceedings of the Seminar for Arabian Studies* 27, pp. 219–223.
Prince, Alan (1980). “A Metrical Theory of Estonian Quantity”. In: *Linguistic Inquiry* 11, pp. 511–562.
— (1983). “Relating to the Grid”. In: *Linguistic Inquiry* 14, pp. 19–100.
Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar.* Tech. rep. Rutgers University and University of Colorado, Boulder.
Prochazka Jr., Theodor (1988). *Saudi Arabian Dialects.* Oxford: Routledge.
Pyle, Charles (1971). “West Greenlandic Eskimo and the Representation of Vowel Length”. In: *Papers in Linguistics* 3, pp. 115–146.
Qafisheh, Hamdi A. (1977). *A Short Reference Grammar of Gulf Arabic.* Tucson: University of Arizona Press.
Rabin, Chaim (1951). *Ancient West Arabian.* London: Taylor’s Foreign Press.

Raffelsiefen, Renate (1999). “Diagnostics for Prosodic Words Revisited”. In: *Studies on the Phonological Word*. Ed. by Tracy Alan Hall and Ursula Kleinhenz. Amsterdam & Philadelphia: John Benjamins, pp. 133–201.

Rakhieh, Belal (2009). “The Phonology of Ma’ani Arabic: Stratal or Parallel OT”. Ph.D. Thesis. University of Essex.

Retsö, Jan (2000). “Kaškaša, t-Passives and the Ancient Dialects in Arabia”. In: *Studi Di Dialettologia Araba*. Ed. by Lidia Bettini. Roma: Istituo per l’Oriente, C.A. Nallino, pp. 111–118.

Rialland, Annie (1994). “The Phonology and Phonetics of Extrasyllabicity in French”. In: *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*. Ed. by Patricia A. Keating. Cambridge: Cambridge University Press, pp. 136–159.

Rischel, Jorgen (1972). “Compound Stress in Danish without a Cycle”. In: *Annual Report of the Institute of Phonetics: University of Copenhagen (ARIPUC)* 6, pp. 211–228.

Rosenbaum, Gabriel (2004). “Egyptian Arabic as a Written Language”. In: *Jerusalem Studies in Arabic and Islam* 29, pp. 281–340.

Rosenhouse, Judith (May 2011). “Bedouin Arabic”. en. In: *Encyclopedia of Arabic Language and Linguistics*.

Rosenthall, Sam and Harry van der Hulst (1999). “Weight-by-Position by Position. Natural Language and Linguistic Theory”. In: *Natural language and linguistic theory* 17, pp. 499–540.

Rosenthall, Samuel (2018). “Gemination”. In: *Encyclopedia Arabic Language and Linguistics*.

Rubach, Jerzy (1993). “Skeletal versus Moraic Representation in Slovak”. In: *Natural Language & Linguistic Theory* 11.4, pp. 625–653.

Rubin, Aaron (2011). “Phyla and Waves: Models of Classification of the Semitic Languages”. In: *The Semitic Languages: An International Handbook*. Ed. by Weninger Stefan et al. Berlin, Germany: De GruyterMouton.

Rubin, Aaron D. (Jan. 2008). “The Subgrouping of the Semitic Languages”. In: *Language and Linguistics Compass* 2.1, pp. 79–102.

Russell, Kevin (1995). “Morphemes and Candidates in Optimality Theory”. University of Manitoba.

Sanders, Nathan (Nov. 2006). *Strong Lexicon Optimization*. en.

Saporta, Sol and Heles Contreras (1962). *A Phonological Grammar of Spanish*. Seattle: University of Washington Press.

Al-Saqqaf, Abdullah Hassan Shaikh (1999). “A Descriptive Linguistic Study of the Spoken Arabic of Wadi Hadramawt, Yemen”. PhD thesis. University of Exeter.

Schaade, A (1911). *Sibawayhis Lautlehre*. Leiden: Brill.

Selkirk, Elisabeth O (1972). “The Phrase Phonology of English and French”. Ph.D. MIT.

— (1978). “On Prosodic Structure and Its Relation to Syntactic Structure”. In: *Nordic Prosody*. Ed. by T. Fretheim. Trondheim: TAPIR, pp. 111–140.

— (1981). “Epenthesis and Degenerate Syllables in Cairene Arabic”. In: *Theoretical Issues in the Grammar of Semitic Languages. MIT Working Papers in Linguistics 3*. Ed. by H. Borer and J. Aoun. Cambridge, MA: Department of Linguistics and Philosophy, MIT, pp. 209–232.

— (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

— (1986). “On Derived Domains in Sentence Phonology”. In: *Phonology* 3, pp. 371–405.

Selkirk, Elisabeth O (1995). “The Prosodic Structure of Function Words”. In: *University of Massachusetts Occasional Papers* 18. Ed. by J. Beckman, L. Dickey, and S. Urbanczyk, pp. 439–470.

Seo, Misun (Apr. 2011). “Syllable Contact”. In: *The Blackwell Companion to Phonology.* Major Reference Works, pp. 1–18.

Shaaban, Kassim Ali (1977). “The Phonology of Omani Arabic”. Ph.D. University of Texas, Austin.

Shattuck-Hufnagel, Stefanie and Alice Turk (1996). “A Prosody Tutorial for Investigators of Auditory Sentence Processing”. In: *Journal of Psycholinguistic Research* 25.2.

Shaw, Jason, Adamantios I. Gafos, et al. (2009). “Syllabification in Moroccan Arabic: Evidence from Patterns of Temporal Stability in Articulation*”. en. In: *Phonology* 26.1, pp. 187–215.

Shaw, Jason A., Adamantios I. Gafos, et al. (2011). “Dynamic Invariance in the Phonetic Expression of Syllable Structure: A Case Study of Moroccan Arabic Consonant Clusters”. en. In: *Phonology* 28.03, pp. 455–490.

Shaw, Patricia (1992). “The Prosodic Constituency of Minor Syllables”. In: *Proceedings of the 12th West Coast Conference on Formal Linguistics.* Ed. by Erin Duncan, Donka Farkas, and Philip Spaelti. University of California, Santa Cruz.

— (1993). “The Prosodic Constituency of Minor Syllables”. In: *Proceedings of the 12th West Coast Conference on Formal Linguistics* 12, pp. 117–132.

Sherer, Tim (1994). “Prosodic Phonotactics”. Dissertation. Amherst, MA: University of Massachusetts, Amherst.

Sibawayhi, Abū Bišr Amr b. Umān b. Qanbar and ’A. S. M. Hārūn (1982/AH 1402). *Kitāb Sībawayhi.* Vol. 4. Cairo: Dār al-Ǧīl.

Sievers, Eduard (1881). *Grundzuge Der Phonetik.* Leipzig: Breitkopf and Hartel.

Singh, Rajendra (1987). “Well-Formedness Conditions and Phonological Theory”. In: *Phonologica 84.* Ed. by Wolfgang Dressler. Cambridge: Cambridge University Press.

Smolensky, Paul (1993). *Harmony, Markedness and Phonological Activity.* Tech. rep. Presented at the Rutgers Optimality Workshop 1.

— (1995). “On the Structure of the Constraint Component Con of UG”. ROA-86.

— (1996). *The Initial State and "Richness of the Base" in Optimality Theory.* en. Technical JHU-CogSci-96-4. Baltimore: John Hopkins Univeristy, p. 31.

Spencer, Andrew (1996). *Phonology.* Oxford: Blackwell.

Spencer, Andrew and Ana R. Luis (2012). *Clitics: An Introduction.* en. Cambridge University Press.

Spitta, Wilhelm (1880). *Grammatik Des Vulgär-Arabischen Dialekts von Ägypten.* Leipzig: Hinrichs.

Stempel, R (1999). *Abriß Einer Historischen Grammatik Der Semitischen Sprachen.* Frankfurt am Main: Peter Lang.

Steriade, Donca (1982). “Greek Prosodies and the Nature of Syllabification”. Doctoral. MIT.

— (1991). “Solution to Palestinian Arabic Stress”. Class Handout. University of California at Los Angeles.

Al-Sulaiti, Latifa Mubarak (1993). “Some Aspects of Qatari Arabic Phonology and Morphology”. PhD thesis. University of Lancaster.

Swadesh, Morris (1937). “The Phonemic Interpretation of Long Consonants”. In: *Language* 13, pp. 1–10.

Taine-Cheikh, Catherine (2007). "Hassaniyya Arabic". In: *Encyclopedia of Arabic Language and Linguistics* II, pp. 240–5.

Tesar, Bruce and Paul Smolensky (1998). "Learnability in Optimality Theory". In: *Linguistic Inquiry* 29, pp. 229–268.

Torres-Tamarit, Francesc and Clàudia Pons-Moll (2019). "Enclitic-Induced Stress Shift in Catalan". en. In: *Journal of Linguistics* 55.2, pp. 407–444.

Tranel, Bernard (1998). "On the Issue of the Syntax/Phonology Interaction". In: *Proceedings of the West Coast Conference on Formal Linguistics 16*. Ed. by E. Curtis, J. Lyle, and G. Webster. Stanford, California: CSLI, pp. 415–429.

Trommelen, Mieke (1983). *The Syllable in Dutch, with Special Reference to Diminutive Formation*. Dordrecht: Foris Publications.

Trubetzkoy, Nikolaj (1938). "Die Phonologischen Grundlagen Der Sogennaten 'Quantitat' in Den Verschiedenen Sprachen". In: *Scritti in Onore Di Alfredo Trombetti*. Ed. by Alfredo Trombetti. Milan: Hoepli, pp. 155–74.

— (1939). *Grundzuge Der Phonologie*. Prague: Travaux du cercle linguistique de Prague.

Tyler, Stephen A. (1969). *Koya: An Outline Grammar: Gommu Dialect*. Berkely & Los Angeles: University of California Press.

Ullendorff, Edward (1970). "Comparative Semitics". In: *Current Trends in Linguistics, Vol. VI: Linguistics in South West Asia and North Africa*. Ed. by Thomas Sebeok. The Hague: Mouton, pp. 261–73.

van der Hulst, Harry (1984). *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.

van der Hulst, Harry and Sam Hellmuth (2010). "Word Accent Systems in the Middle East". In: *A Survey of Word Accentual Patterns in the Languages of theWorld*. Ed. by Harry van der Hulst, R. Goedemans, and E. van Zanten. Berlin; New York: Mouton de Gruyter.

van Oostendorp, Marc (2014). "Selective Lexicon Optimization". en. In: *Lingua*. SI: Usage-Based and Rule-Based Approaches to Phonological Variation 142, pp. 76–84.

Vaux, Bert and Andrew Wolfe (2009). "The Appendix". In: *Contemporary Views on Architecture and Representations in Phonology*. Ed. by Eric Raimy and Charles E. Cairns. MIT Press.

Vennemann, Theo (1988). *Preference Laws for Syllable Structure and the Explanation of Sound Change: With Special Reference to German, Germanic, Italian, and Latin*. Berlin: Mouton de Gruyter.

Versteegh, Kees (1984). "Arab Grammatical Studies before Sībawayhi". In: *Matériaux Pour Une Histoire Des Théories Linguistiques*. Ed. by Sylvain Auroux. Lille: Lille: Université de Lille III Diffusion P. U. L.

— (2014). *The Arabic Language*. Second. Edinburgh: Edinburgh University Press.

Vollers, K (1906). *Volkssprache Und Schriftsprache Im Alten Arabien*. Strassburg: Trübner.

Wackernagel, Jacob (1892). "Über Ein Gesetz Der Indogermanischen Wortstellung". In: *Indogermanische Forschungen* 1, pp. 335–436.

Walters, Keith (1996a). "Diglossia, Linguistic Variation, and Language Change". In: *Perspectives on Arabic Linguistics VIII*. Ed. by Mushira Eid. Amsterdam & Philadelphia: John Benjamins, pp. 157–197.

— (1996b). "Intrasentential Codeswitching in Diglossic Settings and Implications for Linguistic Variation and Change". In: *Sociolinguistic Variation: Data, Theory, and Analysis: Selected Papers from NWAV23 at Stanford*. Ed. by Jennifer Arnold. Stanford, California: CSLI, pp. 401–416.

Watson, Janet C E (2007). “Syllabification Patterns in Arabic Dialects: Long Segments and Mora Sharing”. In: *Phonology* 24.2, pp. 335–356.

— (2011a). “Word Stress in Arabic”. In: *The Blackwell Companion to Phonology*. Ed. by MV Oostendorp et al. Oxford: Wiley-Blackwell.

— (Dec. 2011b). “Arabic Dialects”. en. In: *The Semitic Languages: An International Handbook*. Ed. by Stefan Weninger et al. Berlin: De Gruyter Mouton. Chap. The Semitic Languages, pp. 851–896.

Watson, Janet C. E. and Abdullah Musallam Al-Mahri (2018). “A Stratal OT Account of Word Stress in Teh Mehri of Bit Thuwar”. In: *Re-Engaging Comparative Semitic and Arabic Studies*. Ed. by D Birnstiel and N Pat-El. Gottingen, Germany: Harrassowitz Verlag, pp.151–174.

Watson, Janet C.E. (1992). “Kashkasha with Reference to Modern Yemeni Dialects”. In: *Zeitschrift für Arabische Linguistik* 24, pp. 60–81.

— (2002). *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press. arXiv: `1011.1669v3`.

Welden, Ann. (1980). “Stress in Cairo Arabic”. In: *Studies in the Linguistic Sciences* 10.2, pp. 99–120.

Weninger, Stefan et al., eds. (2011). *The Semitic Languages: An International Handbook*. Berlin: De Gruyter Mouton.

Al-Wer, Enam and Rudolf de Jong (2018). “Dialects of Arabic”. In: *The Handbook of Dialectology*. Ed. by Charles Boberg, John Nerbonne, and Dominic Watt. Hoboken, New Jersey: Wiley Blackwell.

Wheeldon, Linda and Aditi Lahiri (1997). “Prosodic Units in Speech Production”. In: *Journal of Memory and Language* 37, pp. 356–381.

Woidich, Manfred (1980). “Illi Als Konjunktion Im Kairenischen”. In: *Studien Aus Arabistik Und Semitistik. Anton Spitaler Zum Siebzigsten Geburtstag von Seinen Schülernüberreicht*. Ed. by Werner Diem and Stefan Wild. Wiesbaden: Harrassowitz, pp. 224–238.

— (1993). “Die Dialekte Der Agyptischen Oasen: Wetliches Oder Ostliches Arabisch?” In: *Zeitschrift für Arabische Linguistik* 25, pp. 340–359.

— (2010). “Von Der Wörtlichen Rede Zur Sachprosa: Zur Entwicklung Der Ägyptisch-Arabischen Dialektliteratu”. In: *Dialektliteratur Heute*. Ed. by Horst Haider Munske. Online.

Yip, Moira (1995). “Identity Avoidance in Phonology and Morphology”. In: *Proceedings of the Conference on Morphology and Its Relation to Syntax and Phonology*. UC Davis.

— (1996). “Lexicon Optimization in Languages without Alternations”. en. In: *Current Trends in Phonology: Models and Methods*. Ed. by Jacques Durand and Bernard Laks. Salford, Manchester: European Studies Research Institute, University of Salford, p. 32.

Yoda, Sumikazu (2005). *The Arabic Dialect of the Jews in Tripoli (Libya): Grammar, Text and Glossary*. en. Wiesbaden: Otto Harrassowitz Verlag.

Younes, Munther (1995). “On Vowel Shortening in Palestinian Arabic”. In: *Perspectives on Arabic Linguistics VII*. Ed. by Mushira Eid. Amsterdam & Philadelphia: John Benjamins Publishing Company, pp. 157–172.

Zaborski, Andrzej (2018). “Article, Definite”. In: *Encyclopedia Arabic Language and Linguistics*.

Zec, Draga (1995). “Sonority Constraints on Syllable Structure”. In: *Phonology* 12, pp. 85–129.

— (2007). “The Syllable”. en. In: *The Cambridge Handbook of Phonology.* Ed. by Paul De Lacy. Cambridge: Cambridge University Press, p. 709.

Zec, Draga and Sharon Andrea Inkelas (1991). “The Place of Clitics in the Prosodic Hierarchy.” In: *Proceedings of the 10th West Coast Conference on Formal Linguistics.* Stanford, California: CSLI Publications, pp. 505–19.

Zwicky, Arnold (1977). “On Clitics”. In: *3rd International Phonologie-Tagung.* University of Vienna: IULC.

— (1985). “Clitics and Particles”. In: *Language* 61.2, pp. 283–305.

Zwicky, Arnold and Geoffrey Pullum (1983). “Cliticization vs. Inflection: English n’t”. In: *Language* 59, pp. 502–13.