جامعة محمد الخامس بالرباط
كلية الآداب والعلوم الإنسانية
Mohammed V University in Rabat
Faculty of Letters & Human Sciences

Center for Doctoral Studies
"The Human & Space in the Mediterranean World"

Doctoral Program
"Language Culture and Society"

Research Group / Laboratory
"Culture, Language, Education, Society and Development / Migration and Culture"

PATTERNS OF GEMINATION IN MOROCCAN ARABIC: ISSUES IN PHONOLOGY AND MORPHOLOGY

-A Constraint-based Analysis-

A thesis submitted in fulfillment of the requirements for the degree of 'Doctorat'

Prepared by
Ayoub Noamane

Jointly supervised by

Dr. Karim Bensoukas & Dr. Joe Pater
Mohammed V University, Rabat, Morocco | University of Massachusetts, Amherst, U.S.A.

November 2018

THESIS DEFENSE JURY

PATTERNS OF GEMINATION IN MOROCCAN ARABIC
ISSUES IN PHONOLOGY AND MORPHOLOGY

By

Ayoub Noamane

Approved on 10 November, 2018 by

Karim Bensoukas
Mohammed V University, Rabat
Supervisor

Joe Pater
University of Massachusetts Amherst, USA
Supervisor

Hassan Souali
Mohammed V University, Rabat
Chair

Aziz Boudlal
Hassan II University, Casablanca
Member

Abdellatif Al Ghadi
Mohammed V University, Rabat
Member

Mohamed Marouane
Chouaib Doukkali University, El Jadida
Member

Ikbal Zeddari
Mohammed V University, Rabat
Member

ABSTRACT

PATTERNS OF GEMINATION IN MOROCCAN ARABIC: ISSUES IN PHONOLOGY AND MORPHOLOGY

AYOUB NOAMANE

Geminate consonants in Moroccan Arabic (MA) are omnipresent. Besides their phonemic and distinctive nature, they also arise through phonological and morphological processes. In this dissertation, we examine in detail the phonological and morphological patterning of geminates in MA, using the constraint-based framework of Optimality Theory (Prince & Smolensky 1993/2004; McCarthy & Prince,1993a, 1993b, 1995). Along this line of investigation, the issues of geminate distribution, geminate behavior, geminate representation, word minimality, syllable structure, schwa epenthesis and morphological lengthening are scrutinized. This dissertation aspires to be the most comprehensive linguistic study of geminate patterns in MA. In this capacity, our work is meant to be an in-depth exploration of a significant, yet surprisingly understudied, part of the sound system of MA.

In the phonology part, this dissertation explores the special behavior of geminate consonants vis-á-vis short consonants and consonant clusters. By way of comparison, it is shown that geminates exhibit properties that are reminiscent of both unit structures and cluster structures. In order to characterize the patterning of geminates in MA, it is suggested that geminates should be depicted as two root nodes which are underlyingly associated with a mora. Evidence for the moraicity of geminates is illustrated by data pertaining to word minimality, syllabification and word formation. Furthermore, the syllabification of geminates in their different positions is investigated. It is argued that hetrosyllabicity characterizes geminates in all positions and follows from their moraic structure instead of their bi-positional structure.

In the morphology part, the focus is placed on the effect of morphology on geminate integrity and the role of consonant lengthening in word formation. As far as geminate integrity is concerned, the idea entertained is that geminate integrity in MA is systematically subject to breaking by means of morphologically motivated vowels. On the contrary, phonological vowels (i.e. schwa epenthesis) seem to fail in splitting geminates. Furthermore, it is shown that gemination can be employed by morphology to mark the derivation of various forms, namely morphological causatives, agent nouns and instrument nouns. In analyzing this case of morphological gemination, a moraic affix is posited, whose landing position is defined by constraints on syllabic well-formedness.

In addition to investigating the phonological and morphological patterning of geminates in MA, this work also addresses the morphological status of the root as a base of derivation in the morphology of MA. In particular, a chapter is devoted to motivating a root-based approach to morphology based on theoretical evidence from the formation of morphological causatives and the comparative form.

ملخص

مظاهر التضعيف في العربية المغربية
إشكالات صواتية وصرفية

من إعداد
أيوب نعمان

تسلط هذه الأطروحة الضوء على مظاهر التّضعيف وأشكاله في العربية المغربية في إطار نظرية المفاضلة. وتبيّن هذه الدراسة مدى تميّز النظام الصوتي للعربية المغربية بإنتاجية التّضعيف، بحيث يمكن أن يضعّف أي صامت بغض النظر عن طبيعته الصوتية أو عن موقعه داخل الكلمة. وتناقش الدراسة محفّزات التّضعيف والتي تنقسم إلى محفزات معجمية وصواتية وصرفية. وتهدف هذه الأطروحة بالأساس إلى تقديم تحليل صواتي وصرفي يتم الكشف من خلاله عن القيود والتفاعلات التي تتحكم في ظهور التّضعيف في العربية المغربية وكذا توزيعه وتقطيعه.

من الناحية الصّواتية، تتطرق هذه الأطروحة إلى السّلوك الصّواتي للصّوامت المضعّفة مقارنة بسلوك الصّوامت الغير مضعّفة والصّوامت المتتالية، بحيث توضح أن سلوك الصّوامت المضعّفة يمكنها من الجمع بين صفات هذين الأخيرين. كما تهدف الدراسة إلى اقتراح تمثيل صِواتي يجمع بين مستوى تركيبي ومستوى فوق تركيبي، وذلك لحل المشاكل التي تواجهها التمثيلات الأخرى بخصوص السلوك المزدوج للصوامت المضعفة وعلاقتها بالبنية الفوق تركيبية. إضافة إلى ذلك، تتناول الدراسة إشكالية التقطيع الفوق تركيبي للصوامت المضعفة وخاصة الهامشية منها.

أما من الناحية الصرفية، تهدف هذه الدراسة إلى استجلاء طبيعة البناء الصرفي والصوتي لأفعال السبب وأسماء الأفعال وأسماء الآلة المشتقة في العربية المغربية. وتتميز هذه الكلمات بالتضعيف الصرفي لحروفها الوسطى. وبالرغم من بساطة بنائها الصرفي، فإنه من غير السهل تحديد طبيعة اللاصق المسؤول عن اشتقاق هذه الكلمات وكذلك الأسباب وراء تضعيف حروفها الوسطى بالتحديد. انطلاقا من هذا الإشكال، تسعى هذه الأطروحة إلى تحقيق ثلاث أهداف رئيسية. أولا، تحديد طبيعة الاشتقاق في هذه الكلمات. ثانيا، تحديد الدوافع وراء التضعيف الاشتقاقي للحروف الوسطى تحديدا. ثالثا، استنتاج القيود المسؤولة عن هذا النوع من الاشتقاق وكذا سلمها التراتبي.

كما تتناول الأطروحة طبيعة الجذور في العربية المغربية والدور الذي تلعبه في الاشتقاق. وبهذا الخصوص، تبرز الدراسة أن الجذر يلعب دورا محوريا في اشتقاق العديد من الكلمات، بحيث يسمح بتفادي الكثير من مشاكل التطابق التي قد يطرحها اشتقاق كلمة من كلمة أخرى.

ACKNOWLEDGMENTS

There are many people who deserve my utmost gratitude for having contributed to the successful completion of this research project.

Above all else, I am sincerely and deeply grateful to my supervisor, Dr. Karim Bensoukas, whose commitment to the adequate fulfillment of this work has gone beyond the call of duty. It was thanks to his unwavering support and highly professional academic guidance that I could endure all the hardships that come with academic research, in general, and writing a doctoral dissertation, in specific. Particularly, Dr. Bensoukas should be credited for having so generously discussed with me every single detail and issue in this dissertation, and also for having read and commented on earlier versions of this work.

I would also like to acknowledge the role that Dr. Bensoukas has played in my academic career: first, as my phonetics and phonology teacher in my undergraduate studies, then, as my morphology teacher for three semesters during my master's studies, and as the supervisor of my master's thesis prior to my doctoral studies. This is to say that it was Dr. Bensoukas who introduced me to the fields of phonology and morphology as well as the framework of Optimality Theory, which is used here as my framework of analysis. What I have learnt from Dr. Bensoukas during all these years has eventually qualified me to join the University of Massachusetts at Amherst to work with some of the finest linguists in the world.

I wish also to express my heartfelt gratitude to my co-supervisor, Dr. Joe Pater, for accepting to co-supervise this dissertation as well as for strongly supporting my Fulbright application. It is thanks to his support that I could pursue parts of my doctoral studies in the US under his supervision. I am also gratefully indebted to Dr. Pater for his valuable contributions in helping improve many of the OT analyses presented herein and solve many of the problems that got me stuck on many occasions. Besides, the meetings that I had with Dr. Pater have opened my eyes to many of the issues that I have ignored before and allowed me to further expand the scope of my research.

I am also thankful to Dr. John McCarthy, who made a great impact on every aspect of this thesis thanks to his illuminating comments, probing questions, stimulating discussions and tremendous knowledge of linguistics in general and Arabic data in particular. Therefore, I am very grateful to him for having accepted to meet with me on a regular basis despite his busy schedule as Dean of the Graduate School at the time of my Fulbright visit. I should say that Dr. John McCarthy's modesty and generosity have taught me that being a successful researcher takes more than just an exceptional intellect.

I could not have imagined having better advisors and mentors for my doctoral studies than these very distinguished and renowned linguists. However, as the disclaimer goes, any errors of fact or analysis that may have remained in the final version of this thesis are my entire responsibility.

While at UMass, I have benefitted substantially from personal and academic interactions with a number of very brilliant phonology students, including Ivy Hauser, Coral Rose, Aleksei Nazarov, Leland Kusmer and Brandon Loy Prickett. I would like to thank them all for the insightful phonology discussions we had during our Phonology Reading Group meetings as well as for making me feel like one of their own. There are other people from the linguistics community at UMass who deserve my special thanks for their hospitability, namely: John Kingston, Kyle Johnson, Magda Oiry, Gaja Jarosz, Deniz Özyildiz, Alex Göbel, David Erschler, Sakshi Bhatia and Tracy Conner.

It has been a great privilege for me to audit some linguistics courses with a number of very distinguished professors at UMass. Particularly, I wish to thank Gaja Jarosz, Alice Harris and Joe Pater for allowing me to audit their Generative Phonology course, Morphology course and Phonology Seminar, respectively. I should acknowledge that these courses have had a direct role in improving the quality of this dissertation since they enabled me to deepen and expand my knowledge of Optimality Theory and the fields of phonology and morphology.

On a more personal level, I wish to thank Vincent Homer, Rajesh Bhatt and Ilaria Frana for making my time at UMass and in Northampton very enjoyable and unforgettable through their amazing friendship and companionship. More specifically, I wish to thank them for the many trips, parties, musical shows and dinners they have invited me to around Northampton, Amherst and Boston.

Let me also take this opportunity to thank the Moroccan American Commission for Educational and Cultural Exchange, the Fulbright program and AMIDEAST for working together to award me the Fulbright scholarship, which led to the funding of my stay in the US.

For commenting on and reacting to earlier presentations of some parts of this work, I would like to thank the audiences of the 3rd International Doctoral Seminar on Comparative Linguistics, the 1st Annual Doctoral Symposium on Issues in Language, Culture and Translation Studies, the 6th Annual Doctoral Symposium on Arabic Linguistics, held at Ibn Tofail University, Chouaib Doukkali University, Hassan II University, respectively.

Last but not least, I would like to thank my friends, namely Ayoub Loutfi, Mohamed Cherif, Abdessamad Ait Dada and Hamza Bouanani, for their constant support and encouragement. Thank you all for standing by my side through thick and thin.

PHONETIC TRANSCRIPTION

Throughout this dissertation, transcription symbols have the conventional International Phonetic Alphabet (IPA) values, with the exception of emphatic consonants which are indicated here with a dot underneath the consonant in question, for example:

Emphatics	Non-emphatics
ṛ	r
ṣ	s
ṭ	t
ḍ	d

Geminates in IPA are transcribed as a sequence of two identical consonants, indicating their increased length vis-á-vis singletons:

Geminates	Singletons
rr	r
ss	s
tt	t
dd	d

However, note that the IPA value for geminates is purely conventional and has no theoretical implications whatsoever. It is a common misconception to view geminates as clusters solely based on the way they are transcribed.

LIST OF ABBREVIATIONS

AIC	Adjacency Identity Constraint
AN	Agent Noun
C	Consonant
CL	Compensatory Lengthening
CM	Composite Model
Con	Constraint set
CT	Correspondence Theory
EA	Egyptian Arabic
Eval	Evaluator (component of Optimality Theory)
FP	Fixed Prosody
G	Any geminate
GA	Generalized Alignment
GCat	Grammatical Category
Gen	Generator (component of Optimality Theory)
GL	Geminate Law
GTT	Generalized Template Theory
IH	Integrity Hypothesis
IN	Instrument Noun
IPA	International Phonetic Alphabet
MA	Moroccan Arabic
MCat	Morphological Category
MCs	Morphological Causatives
MG	Morphological Gemination
MT	Moraic Theory
OCP	Obligatory Contour Principle
OT	Optimality Theory
PA	Palestinian Arabic
PC	Prosodic Circumscription
PCat	Phonological Category
PP	Past Participle
RM	Realize Morpheme
RMT	Realizational Morphology Theory
RN	Root Node
Rt	Root
RTC	Rule Type Condition
RP	Root-prosody
Pwd	Prosodic Word
SA	Standard Arabic
SPE	The Sound Pattern of English
V	Vowel

LIST OF SYMBOLS

σ	Syllable
μ	Mora
√	Root
☞	Optimal candidate
☜	Multiple optimal forms
☹	Wrong output form
*	One violation
*!	Fatal violation
>>	Domination

TABLE OF CONTENTS

	Page
ABSTRACT IN ENGLISH	**iii**
ABSTRACT IN ARABIC	**iv**
ACKNOWLEDGEMENTS	**v**
PHONETIC TRANSCRIPTION	**vii**
LIST OF ABBREVIATIONS	**viii**
LIST OF SYMBOLS	**ix**
TABLE OF CONTENTS	**x**

CHAPTER ONE
OVERVIEW AND THEORETICAL PRELIMINARIES

1. Purpose of the study	1
2. Scope of the study	4
3. Optimality Theory	10
3.1. Principles of OT	11
3.2. The architecture of OT grammar	12
3.3. Correspondence theory	15
4. Organization of the study	19

CHAPTER TWO
A TAXONOMY OF GEMINATES IN MOROCCAN ARABIC

1. Introduction	24
2. What is a geminate?	25
3. The timing of geminates	26
4. Geminates in MA	28
4.1. The featural composition of geminates	28
4.2. The position of geminates	30
4.3. Lexical geminates	33
4.4. Phonological geminates	39
4.4.1. Gemination through assimilation	39
4.4.2. Gemination as prosodic structure improvement	50
4.5. Morphological geminates	51
5. Fake vs. true geminates	52
6. The lexical distribution of geminates	57
7. Conclusion	59

CHAPTER THREE
GEMINATES IN MOROCCAN ARABIC: BEHAVIOR AND REPRESENTATION

1. Introduction 60
2. The dual behavior of geminates in MA 61
   2.1. Geminates as sequence structures 61
   2.2. Geminates as unit structures 72
3. The phonological representation of geminates in MA 78
   3.1. The linear approach 78
   3.2. The non-linear approach 80
      3.2.1. The autosegmental approach 80
      3.2.2. The moraic approach 92
      3.2.3. The two-root node theory of length 104
      3.2.4. The proposed model: a moraic two-root node model 106
4. Conclusion 110

CHAPTER FOUR
EVIDENCE FROM MOROCCAN ARABIC FOR THE MORAIC NATURE OF GEMINATES

1. Introduction 112
1. Evidence from word-minimality 115
3. Evidence from syllable structure and moraic structure 130
   3.1. Gemination by syllabification analysis 130
   3.2. Compensatory lengthening analysis 133
3. Evidence from word-formation 141
4. Defusing the nonmoraic view of geminates 144
5. Conclusion 146

CHAPTER FIVE
THE SYLLABIFICATION OF GEMINATES IN MOROCCAN ARABIC

1. Introduction 147
2. Geminates as non-complex margins in MA 149
3. The syllabification of geminates 155
   3.1. Word-internal geminates 155
   3.2. Word-initial geminates 160
   3.3. Word-final geminates 171
4. The empty-headed syllable hypothesis 177
   4.1. [i]-epenthesis 177
   4.2. Degemination 184
5. Conclusion 189

CHAPTER SIX
ON THE INTEGRITY OF GEMINATES IN MOROCCAN ARABIC

1. Introduction 191
2. Schwa epenthesis in MA 192
3. Geminate integrity: A review 188
    3.1. Benhallam (1980): Rule types and geminate integrity 202
    3.2. Benhallam (1991): Limitations of the GL 209
4. An OT account of geminate integrity 211
    4.1. Schwa epenthesis in the context of geminate integrity 211
    4.2. Broken plurals 214
    4.3. The past participle 219
5. Some persisting issues 226
6. The fate of split geminates 234
7. Extending the past participle analysis 235
8. Conclusion 245

CHAPTER SEVEN
MOTIVATING THE ROOT IN MOROCCAN ARABIC: DERIVING THE CAUSATIVE AND THE COMPARATIVE

1. Introduction 246
2. Setting the Scene 247
3. The case against the root in Semitic 250
4. Root-based vs. word-based approaches to deriving MCs in MA 256
    4.1. Demotivating a word-based approach 257
        4.1.1. Base inconsistency 259
        4.1.2. Base to output nonconformity 262
        4.1.3. Base unavailability 265
    4.2. Motivating a root-based approach 266
        4.2.1. Underspecified roots 269
        4.2.2. Minimal roots 271
        4.2.3. Abstract roots 273
5. The status of geminates in the root 275
6. The status of vowels in the root 277
6. The comparative: further evidence for the morphological status of the root 280
7. Conclusion 284

CHAPTER EIGHT
A MORAIC APPROACH TO MORPHOLOGICAL GEMINATION IN MOROCCAN ARABIC

1. Introduction 285
2. Morphological gemination in MA: Data description 287
    2.1. Morphological causatives 287
    2.2. Agent nouns 289
    2.3. Instrument nouns 291
    2.4. Basic assumptions 292
        2.4.1. Extending the root-based approach 292
        2.4.2. Templatic morphology 295
        2.4.3. The moraic two-root node representation 298
3. Analysis 301
    3.1. The unified theory of morphological gemination 301
    3.2. Generalized template theory: A root and prosody approach 306
    3.3. Identifying our constraints and their interactions 308
        3.3.1. Affixation module 308
        3.3.2. Selection module 312
        3.3.3. Constraint interaction 313
        3.3.4. Output selection: the causative 318
        3.3.5. Output selection: the agent noun 319
        3.3.6. Output selection: the instrument noun 321
4. Earlier accounts of morphological gemination in MA 323
    4.1. The prosodic circumscription analysis 323
    4.2. The skeletal analysis 325
    4.3. The reduplication analysis 325
5. Conclusion 327

GENERAL CONCLUSION 329
REFERENCES 338

CHAPTER ONE

OVERVIEW AND THEORETICAL PRELIMINARIES

1. Purpose of the study

The motivation behind writing this dissertation is many-fold. Most importantly, this work was spurred by our keen interest in the phonology of geminate consonants. Geminates have been the subject of study and discussion for phonologists since the advent of generative phonology (Chomsky and Halle, 1968). This is mainly because such segments are shrouded in controversy and characterized by a great deal of variability and inconsistency with regards to their phonological behavior and distribution. What also makes studying geminates appealing for phonologists is their ability to affect every aspect of the phonological system of languages that include them (e.g. stress, word minimality, syllable structure and so on). Up to the present, studying geminates remains an intriguing area of research. Phonologists are still trying to figure out a universal underlying structure of geminates that can stand firm in the face of the notable cross-linguistic variations (for recent works on the issue of gemination see Kubozono (2017) and Zimmermann (2017)).

In relation the latter point, two main views have been competing for the status of a universal representation of geminates. Under one view, geminates are conceived of as a pair of skeletal positions sharing a single melodic element. Under the other view, geminates are believed to be monolithic segments associated with an underlying mora. In this context, this dissertation aims to join and contribute to the on-going research about the nature of geminate consonants by highlighting new insights about the various issues associated with one view or the other. The ultimate goal is to help develop a flexible universal model of representing geminates that can comply with the observed individual crosslinguistic variations.

What also makes studying geminates interesting to us is the fact that Moroccan Arabic belongs to a category of languages that make a distinction between short and long consonants (e.g. Russian (Dmitrieva, 2017), Tashlhit (Bensoukas, 2001), Japanese (Kubozono et al., 2008), Trukese (Hart, 1991), Somali (Blevins, 2008)). Geminate patterns in MA permeate the language system. Interestingly, not only do geminates function as distinctive segments for lexical contrast in MA, but they also arise from morphological as well as phonological activity. Therefore, the issue of gemination in MA stands out as an intriguing area of research that needs to be explored for a better understanding of the language's morphological and phonological systems. However, very few linguists have seriously concerned themselves with the study of geminate patterns in MA. Therefore, such a work is especially needed in order to fill up a gap in the body of research on MA phonology concerning the study of geminates. While the issue of geminates has been addressed in a few academic articles and dissertations (Benhallam, 1980, 1991; Boudlal, 2001), especially in the context of schwa epenthesis and syllable structure, there are, to our knowledge, no works that are fully devoted to the phonological patterning of geminate consonants, much less their morphological patterning. Thus, this work is set to be the first and most comprehensive thesis fully devoted to the study of geminate patterns of MA. Also, the current work differs from its predecessors in terms of the range of the issues it raises and the contributions it provides.

Besides providing a different perspective on a number of previously studied issues such as geminate integrity, the formation of morphological causatives, the formation of the past participle form as well as the morphological status of the root, we also provide some first-time analyses for a number of phonological and morphological patterns in MA. These include the syllabification of edge geminates within moraic theory, [i]-epenthesis after final-geminated verbs, the degemination of the passive affix /tt-/, gemination by syllabification in the context of the feminine suffix /-at/ and the formation of the agent form and the instrument form. Having said that, the novelty of this work stems from dealing with issues that are at the forefront of phonological research and that have far-reaching implications on central discussions about the phonological representation of consonantal length, the interplay between morphology and phonology, the nature and role of the root in Semitic as well as the morphological legitimacy of templates in nonconcatenative morphology.

While the central aim of this dissertation is to uncover the constraint interactions responsible for the creation, distribution and parsing of geminate consonants, the work also promises other equally important contributions. More specifically, we also seek to put the data of MA geminates in its cross-linguistic context by comparing it to data of geminates from other languages. Besides, we intend to offer a sound critical review of the issues and questions that have kept phonologists so occupied about geminates. Another major goal for us is to make MA data theoretically and empirically useful for supporting or refuting previous accounts and analyses. We can also add that this dissertation represents a continuation of and an update on previous OT work on MA (e.g. Hammari (2000); Boudlal (2001); Rguibi (2001); Imouzaz (2002); Laaboudi (2003)).

2. Scope of the Study

The scope of this study spans a wide range of issues in the phonology and morphology of MA, but with a specific focus on the issue of gemination. In particular, we will seek to define the interacting constraints and the crucial rankings that govern the occurrence, distribution, behavior, representation and syllabification of geminates as well as their role in word formation.

As previously mentioned, the sound pattern of Moroccan Arabic is characterized by the notable presence of long consonants, technically referred to as geminates. These can be lexical, phonologically coerced or morphologically derived. The lexicality of geminates is determined by their distinctive role in establishing semantic contrast between different words. This is put into evidence by the following minimal pairs:

(1) Lexical geminates

ħluf	‘oath’	ħə**ll**uf	‘boar’
bna	‘he built’	bə**nn**a	‘flavor’
ħaʒa	‘thing’	ħa**ʒʒ**a	‘a female pilgrim’
ʃalal	‘paralysis’	ʃə**ll**al	‘waterfall’
ʒafaf	‘drought’	ʒə**ff**af	‘a floor mop’

In light of minimal pairs such these, we intend to determine the constraints responsible for the phonemic status of geminates in MA. Besides, we will also look into those geminate structures which are phonologically created through processes of total assimilation and prosodic lengthening. Such structures are exemplified below:

(2) Phonologically-derived geminates

a.	/l-ʃita/	**ʃʃ**ita	‘the broom’
	/l-ṣif/	**ṣṣ**if	‘the summer’
	/l-ʒaṛ/	**ʒʒ**aṛ	‘the neighbor’
	/l-taman/	**tt**aman	‘the price’
	/l-dib/	**dd**ib	‘the wolf’
b.	/ʒbəd-t/	ʒbə**tt**	‘I pulled out’
	/ṣrəṭ-t/	ṣrə**tt**	‘I swallowed’
	/ɦbəṭ-t/	ɦbə**tt**	‘I came down’
	/ħfəḍ-t/	ħfə**tt**	‘I memorized’
	/mrəḍ-t/	mrə**tt**	‘I got sick’
c.	ḍəṛbatu	ḍəṛbə**tt**u	‘she hit him’
	ʃəṛbatu	ʃəṛbə**tt**u	‘she drank it
	qətlatu	qətlə**tt**u	‘she killed him’
	nəʃṛatu	nəʃṛə**tt**u	‘she hung it’
	ʕəṛfatu	ʕəṛfə**tt**u	‘she knew him’

In this concern, we will be interested in investigating how sequences of non-identical consonants can be transformed into geminate structures.

On a different note, we will devote some space to investigating the distribution and parsing of geminate consonants in MA. What is interesting about MA is the fact that geminates may appear medially as well as at the edges. This is exemplified by the following items:

(3) The distribution and syllabification of geminates in MA

Word-initially		*Word-medially*		*Word-finally*	
<u>dd</u>i	'to take'	sə<u>nn</u>a	'tooth'	sə<u>dd</u>	'to close'
<u>ss</u>uq	'the market'	gə<u>ff</u>a	'bag'	ʒə<u>rr</u>	'to pull
<u>ll</u>imun	'the orange'	sa<u>ʕʕ</u>a	'clock'	ʕə<u>ḏḏ</u>	'to bite'

In dealing with this issue, we are looking towards developing an account of geminate syllabification that is compatible with the latest development in phonological representation, namely moraic theory.

In terms of their phonological behavior, we will show that geminates in MA demonstrate an inconsistent behavior in relation to the process of schwa epenthesis. We ask if geminates do get split up in MA, and when and how does that happen? Note that, in the data below, the epenthetic position of the schwa with regards to the geminates in question is characterized by variability:

(4) Schwa epenthesis and geminate integrity

(i)

ɦṛəb	'escape'	sədd	'close'
tləf	'disappear'	ħəll	'open'
ʃṛəb	'drink'	ʃədd	'catch'
kbər	'grow'	dəqq	'knock'

(ii)

a.

	Sg.	*Pl.*	
	səkka	sk**ə**k	‘rail’
	dərra	dr**ə**r	‘veil’
	dəffa	df**ə**f	‘door’
	ʕəʃʃ	ʕʃ**ə**ʃ	‘nest’

b.

	Sg.	*Pl.*	
	bəqqa	b**ə**qq	‘bug’
	ħəbba	ħ**ə**bb	‘grain’
	bəẓẓa	b**ə**ẓẓ	‘brat’
	qəʃʃa	q**ə**ʃʃ	‘junk’

Therefore, we will try to determine why geminates and schwa epenthesis behave the way they do in relation to each other.

The relevance of geminates in MA extends to word formation and derivation. Derived forms like causative verbs, agent nouns, and instrument nouns all involve a medial geminate that is not part of the base form and that functions as a morphological marker.

(5) Morphological gemination

(i) The causative verb

xsər	‘lose’	xə**ss**ər	‘to make lose’
zləq	‘slip’	zə**ll**əq	‘to make slip’
ʃṛəb	‘drink’	ʃə**ṛṛ**əb	‘to make drink’

(ii) The agent noun

gzər	‘to butcher’	gə**zz**ar	‘butcher’
bni	‘to build’	bə**nn**aj	‘mason’
ʃuf	‘to see’	ʃə**ww**af	‘fortuneteller’

(iii) The instrument noun

skət	‘to be quiet’	sə**kk**ata	‘pacifier’
kwi	‘to weld’	kə**ww**aja	‘a welding machine’
sqi	‘to water’	sə**qq**aja	‘fountain’

Nonetheless, it remains unclear what morphological process in particular motivates this type of geminates and what determines their position in the derived forms. In this regard, we will show that gemination can take on the role of a morphological marker in the derivation of the word forms cited above.

In addition, one of the central ideas entertained in this dissertation is that word formation in MA is basically root-based, unless otherwise proven. In this regard, we will be presenting data from the formation of causative verbs and comparative forms which supports a root-based approach to derivation. A sample of the data to be discussed is given below:

(6)

Root	*Noun*	*Causative*	
√ʕdb	ʕadab	ʕəddəb	‘torture’
√ɣwt	ɣawt	ɣəwwət	‘screaming’
√ħli	ħəlwa	ħəlli	‘candy’
√jtm	jtim	jəttəm	‘orphan’

√ṣbn	ṣabun	ṣəbbən	‘soap’
√fʃʃ	fʃuʃ	fəʃʃəʃ	‘pampering’

Root	*Adj.*	*Comp.*	
√tql	tqil	tqəl	‘heavy’
√ɣlḍ	ɣliḍ	ɣləḍ	‘fat’
√ṭwl	ṭwil	ṭwəl	‘tall’
√ʕṛḍ	ʕṛiḍ	ʕṛəḍ	‘wide’
√qsṛ	qsiṛ	qsəṛ	‘short’
√ṛxs	ṛxis	ṛxəs	‘cheap’

With reference to this data, we will point out to the irregularities posed by a word-based approach to derivation. Then, we will demonstrate how a theory of roots can eschew those irregularities.

Our account for the various issues pertaining to gemination in MA will be couched within the constraint-based framework of Optimality Theory (OT) (Prince & Smolensky, 1993/2004; McCarthy & Prince, 1993a, 1993b, 1995, 1999). The OT framework allows for typologically generalizable analyses by appealing to universal and cross-linguistically common constraints. In the section that follows, a brief review of Optimality Theory, in its classical parallel version, will be provided.

3. Optimality theory

OT (Prince & Smolensky, 1993/2004; McCarthy & Prince, 1993a, 1993b, 1995, 1999 and others; see also Kager, 1999; McCarthy 2002, 2008 for good introductions) is a development of generative grammar that embodies an alternative theoretical framework to modeling and representing human linguistic knowledge. OT takes as a starting point the Chomskyan belief that languages are fundamentally similar as a result of a genetic predisposition in humans to acquire LANAGUAGE (Chomsky (1957) and others).

Central to OT is the assumption that surface language structures are the outcome of interaction between competing universal, yet violable, constraints. The universality of constraints offers a way of characterizing the cross-linguistic similarities that correspond to universal grammar (UG). Violability, on the other hand, makes it possible to account for cross-language variations. This being the case, OT dispenses with the assumption of parametric setting of universal principles. Specific to OT is the assumption that constraints are ranked differently from one language to another with priority being given to some constraints over the others. A surface form is qualified as being optimal (i.e. the most harmonic), as compared to other candidates, if it incurs the least serious violations of the relevant set of constraints, with respect to their hierarchical ranking. The wide acclaim of OT emanates from the fact that it offers in-depth insights into language typology and language universals.

3.1. Principles of OT

At the heart of OT are the following assumptions (McCarthy and Prince, 1993a):

(a) Universality

(b) Violability

(c) Ranking

(d) Inclusiveness

(e) Parallelism

The first principle suggests that constraints, provided by Universal Grammar, are made universally accessible to, and are present in, all grammars. That is, grammars of individual languages have recourse to the same basic and limited set of constraints, reflecting the general structure of 'LANGUAGE'. Violability means that constraints are violable, yet violation should be kept minimal, taking into consideration the ranking of constraints. Ranking indicates that constraints are ranked on a language-particular basis; a grammar is a ranking of the constraint set. In this view, each language thus reflects a specific way, out of many, of arranging and ranking the constraint set. The fourth principle of OT, inclusiveness, proposes that the generation of candidate analyses is governed by very general considerations of structural well-formedness. This means that "there are no specific rules or repair strategies with specific structural descriptions or structural changes or with connections to specific constraints" (McCarthy and Prince, 1993a: 6). Parallelism postulates that satisfaction or violation of the constraint hierarchy is computed over the whole hierarchy and the whole candidate set, with no intermediate representations between the input and the output. In this approach, OT abandons the idea of serial derivation according to which a surface form is derived by applying a series of rewrite-rules to a given underlying form.

3.2. The architecture of OT grammar

The OT grammar comprises four main components namely the lexicon, the generator, the constraint set and the evaluator. First, the lexicon is defined as a general repertoire or inventory containing the underlying forms of morphemes. This component is responsible for supplying the grammar with the essential lexical representations which are then handed to the generator to function as input forms. Crucial to understanding the lexicon, under the auspices of OT, is the notion of 'Richness of the Base'. In this model, no prohibitions, restrictions or linguistic demands can hold at the level of the underlying representations contained in the lexicon. To put it differently, all aspects of structural well-formedness are determined by the constraint hierarchy, regardless of the forms posited by the lexicon (Kager, 1999). Kager (1999) further elucidates that OT's Richness of the Base is a way out of the early generative grammar duplication problem where 'Morpheme Structure Constraints' were argued to reproduce information predicted by rewrite rules.

Second, the Generator (Gen for short) is a function that generates for a given input an unlimited number of logically possible output candidates $\{Cand_1, Cand_2, Cand_n\}$ from which the optimal one is singled out (see Prince & Smolensky, 1993/2004; McCarthy & Prince, 1993a:21, 1993b:9).

The function of Gen can be best simplified by the following diagram:

(7) Gen (Input) $\rightarrow \{Cand_1, Cand_{2,...}Cand_n\}$

Third, the constraint set (hereafter Con) is claimed to be part of Universal Grammar (UG for short) and hence universally shared by the grammars of all individual languages. A constraint is a structural demand that may be either satisfied or violated by an output form. Two central types of constraint families are recognized in OT, namely markedness and faithfulness[1]. On the one hand, markedness constraints dictate some criteria of structural well-formedness that output structures are required to obey. These constraints may either prohibit the occurrence of certain structures (e.g. vowels must not be nasal) or prefer given forms over others (e.g. sonorants must be voiced). Generally speaking, markedness constraints militate against marked structures and favor their unmarked counterparts.

On the other hand, faithfulness constraints demand that output forms preserve the featural, segmental and prosodic identity of their basic forms, thus enforcing the maintenance of similarity between the output and its input. In other words, this set of constraints militates against the eroding effects of markedness constraints, therefore protecting lexical contrasts from being neutralized. As a result, it appears that these two constraint families demonstrate contrasting demands of structural well-formedness, in the sense that the satisfaction of markedness implies the violation of faithfulness and vice versa.

In theory, the total absence of faithfulness constraints would lead to the neutralization of all lexical contrasts via the acute minimization of markedness. Contrariwise, the absolute lack of markedness constraints would result in grammars with a maximum degree of markedness as well as a complete matching between outputs and their inputs. Accordingly, there must be a way of compromise, reconciliation and balance between the conflicting demands of preserving contrast

[1] The contrast between what is marked and what is unmarked is deeply rooted in the articulatory and perceptual systems of humans. Generally, unmarked structures are basic and easy to articulate and perceive (e.g. non-round front vowels), while marked ones are avoided and used mainly to maintain contrast (e.g. round front vowels) (see Kager, 1999:2).

and minimizing markedness. Here comes the role of the fourth component of OT grammar, the evaluator (henceforth Eval).

Eval is a function which selects the optimal output from a number of candidate analyses via the assessment of these against a hierarchy of ranked constraints in a parallel and inclusive fashion. To this end, Eval applies a language-specific hierarchy of universal constraints to the set of candidates. Within this hierarchy, there is a transitivity of ranking which stipulates that: "if C_1>>C_2 and C_2>>C_3, then C_1>>C_3" (Kager, 1999:21) (where A>>B reads A dominates B). Each constraint in the hierarchy eliminates some candidate output that violates it, until only one candidate remains.

In this approach, Eval is responsible for assigning violation marks to those candidates which incur them. Yet, violation of lower-ranked constraints must always be kept to a minimum as far as it is enough to obviate the violation of some higher-ranked constraint. In a related vein, no candidate is expected to satisfy all the constraints of the grammar since these have inherent conflicting demands on structural well-formedness. This principle is referred to as the 'Fallacy of Perfection', according to which no perfect output is expected to occur. In compliance with this view, "no candidate is good or bad in itself; it's only good or bad in relation to other candidates from the same input" (McCarthy, 2008:21); thus, the notion of ungrammaticality is relative in nature.

Additionally, it is worth noting that the dominance relation between constraints is strict, hence the notion of strict domination. This means that the satisfaction of lower-ranked constraints cannot compensate for the violation of any higher-ranked constraint, nor can the violation of multiple lower-ranked constraints be evaluated as more serious than the violation of a higher-ranked constraint. The role of Eval can be best exemplified in the following chart:

(8) EVAL {$Cand_1$, $Cand_2$, $Cand_n$} → [Output]

To sum up this section, an OT grammar can be schematically represented as follows:

(9)

/Input/ ➔ $\boxed{\text{GEN}}$ ➔ {$Cand_1$, $Cand_{2,...}$$Cand_n$} ➔ $\boxed{\text{EVAL}}$ ➔ [Output]

3.3. Correspondence theory

Correspondence theory (hereafter CT) was proposed by McCarthy and Prince (1995, 1999) expanding on Prince and Smolensky (1993) and McCarthy and Prince (1993a). In this theory, two elements S_1 and S_2 stand in a correspondence relation defined as in (12). Constraints in OT are of two major types: markedness constraints and faithfulness constraints. The former penalize phonological marked structures, whereas the latter requires identity between corresponding elements. Thus, CT concerns faithfulness constraints as it governs identity relations between various corresponding forms, namely between input-output, base-reduplicant and output-output. CT claims that faithfulness and identity requirements holding between the diverse aforementioned representations stem from the same exact set of constraints listed in (13).

(10) Correspondence:

Given two strings S_1 and S_2, correspondence is a relation R from the elements of S_1 to those of S_2. Elements $\alpha \in S_1$ and $\beta \in S_2$ are referred to as correspondents of one another when α R β.

In OT earliest work, the faithfulness constraints PARSE and FILL (see Prince and Smolensky, 1993; McCarthy and Prince, 1993a) militate against any divergence between input and output. This is abandoned in favor of the formalism of correspondence, which is originally developed for the reduplicant-base relation and aims at providing a better account for those identity-preserving interactions termed underapplication and overapplication. Consequently, CT eliminates the need for a separate theory for each of these relations and as a result unifies the formal treatment of corresponding relations (McCarthy and Prince, 1999).

Constraints on correspondence relations refer to two corresponding structures (S_1, S_2) and a relation ($\Re$) between them. Each constraint is a general family of domain-specific constraints for input-output, base-reduplicant or output-output correspondence.

(11) Constraints on correspondence relation: (McCarthy and Prince, 1995, 1999)

a. Maximality (MAX): Every element of S_1 has a correspondent in S_2 (No deletion).

b. Dependence (DEP): Every element of S_2 has a correspondent in S_1 (No epenthesis).

c. Identity-[F] (IDENT-F): Correspondent segments in $S_1$and S_2 have identical values for feature [F] (No change).

d. Contiguity (CONTIG):

i. I-CONTIG (No Skipping): The portion of S_1 standing in correspondence forms a contiguous string.

ii. O-CONTIG (No Intrusion): The portion of S_2 standing in correspondence forms a contiguous string.

e. {Right, Left}-Anchor (S_1, S_2): Any element at the designated periphery of S_1 has a correspondent at the designated periphery of S_2.

f. Linearity (No Metathesis): S_1 is consistent with the precedence structure of S_2, and vice versa.

g. Uniformity (No Coalescence): No element of S_2 has multiple correspondents in S_1.

h. Integrity (No Breaking): No element of S_1 has multiple correspondents in S_2.

MAX and DEP have opposite requirements. First, the MAX constraint family prohibits and militates against phonological deletion of S_1 material. It simply requires any element present in S_1 to have a matching element in S_2. For instance, the output /CV/ is said to violate MAX assuming that the input is /CVC/. Second, the DEP constraint family, on the other hand, is violated in case S_2 includes an element which has no correspondent in S_1. In other words, DEP bans and militates against any type of phonological epenthesis, demanding every element in S_2 to be present in S_1. For example, the output /CVCC/ of the input /CVC/ is ruled out by DEP.

The IDENT constraint family demands that elements of S_1 and those of S_2 be identical feature-wise. It militates against feature alternation between correspondent segments. Hence, any candidate where feature alternation occurs is punished by IDENT; for example, [kid] violates IDENT supposing that /kit/ is the input and that [d] and /t/ are correspondent elements.

CONTIG and LIN deal with linearity in different ways. On the one hand, the Contiguity constraint family comprises two categories of Contiguity constraints: I-Contiguity and O-Contiguity. Generally, it demands the preservation of the linear order of elements. More specifically, I-Contiguity bans deletion of elements internal to the input string (e.g. xyz → xz) (i.e. no skipping). O-Contiguity, on the other hand, prohibits insertion of elements internal to the input string (e.g. xz → xyz) (i.e. no intrusion). Linearity, on the other hand, is a constraint family which excludes metathesis as it requires the maintenance of the linear order of corresponding structures. Hence, forms like [zyx] and [yxz], for the input /xyz/, violate this constraint because the sequential order of the three elements is not preserved.

The Anchoring constraints (also called Alignment constraints, see Prince & McCarthy, 1993b) are a single family of well-formedness constraints. They regulate the coincidence between the diverse prosodic categories (PCat) and grammatical ones (GCat). In this model, alignment constraints specify the designated edges of the relevant constituents by referring either to the right edge (i.e. ALIGN-R) or the left edge (i.e. ALIGN-L).

Uniformity and Integrity are two general constraint families against multiple correspondences of elements. In particular, uniformity bans elements of S_2 from being multiply represented in S_1 (i.e. no coalescence). That is to say, it militates against fusing and merging the elements of S_1 as they emerge in S_2. Integrity, on the other hand, prohibits the splitting of S_1 elements in S_2 (i.e. no breaking). It rules out any candidate where two elements correspond to one element in the input.

Specific examples of constraints and their interaction will be given in due time as we begin analyzing our data. In the subsequent section, the structure of this dissertation will be mapped out and the content of each chapter will be overviewed.

4. Organization of the study

In addition to this introductory chapter, the main body of this dissertation consists of seven other chapters, each of which contains some amount of analysis and problem solving. Even though all the chapters overlap with each other to a great extent, this thesis can be roughly divided into two parts. The first part can be said to be devoted mainly to phonological issues concerning geminate patterns, whereas the second part can be described as being concerned with the morphological implications of gemination and issues of word derivation. The first part subsumes chapters 2 through 5, and the second part incorporates chapters 7 and 8. Chapter 6, however, lies right at the interface between morphology and phonology as it deals with the impact of phonological and morphological processes on geminate integrity in MA.

Chapter two, entitled ‘A taxonomy of geminates in Moroccan Arabic’, lays out and describes the data to be the treated in the course of this study. It categorizes the geminate patterns attested in MA based on their origin, featural make-up and position. In this chapter, it is shown that geminates in MA may descend from every major class feature and occur in all word positions. In addition, we highlight the basic constraint interactions that make underlying geminates possible in MA. We also examine the ways geminates can be derived phonologically. In particular, we make the claim that geminates can arise through the processes of total assimilation and prosodic lengthening. Moreover, we present data demonstrating that geminates in MA can also be morphologically derived. Also, in this chapter, we introduce some of the criteria that can be used to differentiate between fake geminates and true geminates. In this regard, we argue that derived geminates can also be described as ‘true’ as long as they exhibit the same properties of underlying geminates, namely in reaction to schwa epenthesis. Finally, in this chapter, we briefly describe the extent to which the major lexical categories may allow geminate consonants.

Chapter three, entitled ‘Geminates in Moroccan Arabic: behavior and representation’, examines the phonological behavior and representation of geminates in MA vis-á-vis consonant clusters and singletons. On this matter, we show that the behavior of geminates is characterized by inconsistency. In particular, it is shown that geminates tend to pattern with both consonant clusters and singletons on different occasions. On this basis, we proceed to evaluate the possible ways geminates can be represented in phonological theory in light of the observed behavior. In this context, the skeletal model and the moraic model of geminate representation are contrasted, and the empirical implications that follow from each are highlighted. This has led us to the conclusion that each model is inadequate to account for the range of challenges posed by the MA data about geminates, in specific, and prosody, in general. As an alternative, a composite model, where geminates are represented as bi-positional and moraic, is proposed to explain the dual behavior of geminates and their impact on the prosodic system of the language.

Chapter four, entitled ‘Evidence from Moroccan Arabic for the moraic nature of geminates’, offers firsthand pieces of evidence that can support the moraic conception of geminates in MA. The evidence presented comes from word minimality, compensatory lengthening and word formation. First, we argue that that the existence of GV words (e.g. ddi ‘to take’), to the exclusion of CV words (e.g. *di), in MA proves that a geminate is underlyingly moraic, in that a monovocalic word with an initial geminate is qualified as bimoraic. We also claim that the fact that MA involves a case of compensatory lengthening that leads to the creation of geminate consonants represents additional evidence in support of the moraic representation of geminates. This is because that compensatory lengthening has been shown to be mainly a moraic preserving process across many languages. The third type of evidence we introduce consists in demonstrating that a moraic approach to consonantal length proves to be the best option to account for morphological gemination in MA and elsewhere.

In chapter five, ‘The syllabification of geminates in Moroccan Arabic’, we first address the various issues posed by our proposed model for representing geminates in the context of syllable structure. In this regard, it is argued that geminates in MA are parsed under moraic analysis rather than segmental analysis. With this in mind, we proceed to uncover the constraint interactions responsible for the syllabification of medial geminates and peripheral ones. Our analysis characterizes the syllabification of geminates by means of the same constraints regardless of their position in the word. Also, this chapter zooms in on the issue of final geminate syllabification. We demonstrate that, just like initial and medial geminates in MA, final geminates are also heterosyllabic, creating a structure of empty-headed syllables to prevent the geminate from being contained under a single syllabic node. Proving this hypothesis are two phonological processes: [i]-epenthesis after final-geminated verbs and degemination of the reflexive prefix [tt-].

Chapter six, entitled ‘On the integrity of geminates in Moroccan Arabic’, discusses the issue of geminate integrity in MA. It concerns itself with the observation that the integrity of geminates in MA is preserved on some occasions and compromised on others. To explain this behavior, Benhallam (1980) argued that the integrity of geminates in MA is subject to the nature of the rule targeting the geminate, such that geminates can be split by morphological rules but not by phonological ones (especially schwa epenthesis). Building on this, the purpose of this chapter is two-fold. First, we intend to characterize Benhallam's condition in terms of constraint interaction. Second, we aim to account for the data that challenges the condition under the same terms. In particular, we claim that when the integrity of geminates is compromised by means of schwa epenthesis, that might be because of morphological pressure as well.

Chapter seven, entitled 'Motivating the root in Moroccan Arabic: deriving the causative and the comparative' deals with the morphological status of the root in word formation in MA. This chapter paves the way to the next one by arguing that the derivation of morphological causatives, and by extension agent nouns and instrument nouns, is better accounted for if a root-based approach is adopted. We argue that a root-based approach explains why some of the unpredictable variations occur between some causatives and their corresponding output forms. On the contrary, we demonstrate that a word-based approach fails to account for many of the variations between some causatives and their output base forms. To further substantiate the root-based approach, this chapter also deals with the derivation of comparative forms in MA, which seems to be better explained if roots are taken as base forms.

Chapter eight, entitled 'A moraic approach to morphological gemination in Moroccan Arabic', is concerned with the issue of morphological gemination. The focus is on determining the nature of the affix that leads to consonant lengthening for the sake of deriving forms such as causative verbs, agent nouns and instrument nouns. Also, the chapter tries to account for the conditions that determine the position of the designated affixes in the derivation of causatives, agent nouns and instrument nouns. To this end, it will be shown that morphological gemination occurs as the result of affixing a moraic morpheme to a root. As for the linearization of this moraic affix, it will be argued that it is decided by constraints of prosodic markedness. The latter will also be used to explain the invariant templates of the grammatical forms under study in this chapter.

CHAPTER TWO

A TAXONOMY OF GEMINATE CONSONANTS IN MOROCCAN ARABIC

1. Introduction

Geminate consonants appear in many languages around the globe. Previous studies have shown that these languages tend to vary in terms of their geminate inventories. Differences may concern the ways through which geminates can surface in the phonological systems of these languages or the restrictions placed on geminates' featural composition and lexical position (Kawahara, 2007; Pajak, 2009). On the extreme sides of this cross-linguistic variation of geminate inventories, there exist languages that allow geminates from all major class features, and in all positions (e.g. Moroccan Amazigh; see Saib (1973) for data from Tamazight and Bensoukas (2001) for data from Tashlhit), and languages that permit only the least marked geminate structures, namely intervocalic voiceless obstruent geminates (e.g. Japanese (Kawahara, 2005, 2015; Kawahara and Melanie, 2017; Kubozono et al., 2008)).

In this cross-linguistic context, the purpose of this chapter is to provide a bird's eye view of the issue of gemination in MA through establishing a taxonomy of the geminate structures that arise in the language. To this end, we will show that geminates in MA are fundamentally lexical, but can also be phonologically and morphologically derived. We will also show that MA patterns with languages that put no restrictions on geminates, and hence allows all kinds of geminate structures both feature-wise and position-wise.

The remainder of this chapter is structured as follows. Sections 2 and 3 offer brief phonological and phonetic definitions of geminate consonants. Section 4 introduces the different sound classes that geminates can belong to and the word positions they can occur in. Section 4 also describes the various ways through which geminates can arise in MA. It specifically shows that the patterns of geminates attested in MA can be lexical, phonological as well as morphological. Section 5 draws a distinction between fake geminates and true geminates. Section 6 describes the lexical distribution of geminates by showing what major lexical categories are marked by gemination and to what extent. Section 7 sums up the content of this chapter.

2. What is a geminate?

A geminate is generally defined as a "long or doubled consonant that contrasts phonemically with its shorter or singleton counterpart" (Davis, 2011). The terms 'geminate', 'doubled consonant', 'long consonant' and 'strong consonant' are often all used interchangeably in the literature to refer to the same kind of segments. Sometimes, using one term over the other can be theoretically motivated. For instance, the term 'doubled consonant' could be used to imply that a geminate is a sequence of two identical consonants. The terms 'long consonant' and 'strong consonant' clearly refer to geminates as single consonants. The term 'geminate' is the most theoretically neutral one, although sometimes the phrase 'geminate cluster' is used to clarify the author's assumptions about the nature of the geminate structure.

Giving a formal phonological definition of geminates requires a theoretical framework of analysis. In linear phonology, a geminate is regarded either as a single segment specified for the feature [+long] (Chomsky and Halle, 1968) or as a sequence of two identical consonants (Kenstowicz and Pyle, 1973). In non-linear phonology, a geminate is defined in terms of prosodic association to two skeletal positions (McCarthy, 1979, 1981) or a mora (Hyman, 1985; McCarthy and Prince, 1986; Hayes, 1989). In principle, a geminate is not always a unit of phonemic contrast as it may have other roles defined by the various ways geminates can be created. For example, in many languages, geminates function as morphological markers, derived to serve different morphological roles.

3. The phonetic timing of geminates

Phonetically speaking, the difference between a geminate and a singleton is described in terms of tenseness, with the former being [+tense]. This means that geminates are articulated with longer constriction in the vocal tract compared to 'normal' consonants. Durations may vary depending on the quality of the geminate (e.g. stop or fricative geminate) (Cohn et al., 1999), the position of the geminate (i.e. medial or peripheral) or the elements surrounding the geminate (e.g. vowels or consonants) (Pajak, 2009a, b). It has been shown that on average the ratio between geminates and singletons ranges between 1:1.5 to 1:3 (see Hankamer and Lahiri, 1988; Ladefoged and Maddieson, 1996; Thurgood, 2001, 1993; Ridoune, 2010; Khattab and Al-Tamimi, 2014 for more on the phonetic properties of geminates).[2] For concreteness, the

[2] This means that a geminate can be 1.5 to 3 times longer than its corresponding singleton.

following tableau from Cohn et al., (1999) shows the mean closure durations for word-medial geminates and singleton consonants in Buginese, an Indonesian language:

(1) Buginese mean closure durations in V_V

	Singletons	*Geminates*	*Ratio*
vl. stops	102ms	169ms	1:1.7
vd. stops	68ms	116ms	1:1.7
vl. frics.	138ms	192ms	1:1.4
nasals	88ms	144ms	1:1.6
Lateral liquid	84ms	151ms	1:1.8
rhotic	19ms	73ms	1:3.8
glide	105ms	151ms	1:1.8

Later in this chapter, we will be demonstrating that geminates are omnipresent in the phonological system of MA, appearing in different positions and covering a broad spectrum of featural qualities. Therefore, studying the phonetic properties of geminates in MA is a very attractive research project. Particularly, it would be quite interesting to measure and compare the closure durations of geminates based on their quality (i.e. obstruent vs. sonorant geminates), position (i.e. peripheral vs. internal geminates) and origin (lexical vs. derived), and then see if these factors have any influence on the perception and acquisition of geminates. To the extent of our knowledge, no such measurements have been attempted before by any study. Unfortunately, this work will not address the phonetic measuring of the closure durations of geminates in MA. Thus, for now, we leave this matter for future research.

4. Geminates in Moroccan Arabic

4.1. The featural composition of geminates

According to Perlmutter (1995), languages where phonological contrasts occur between single and geminate consonants or short and long vowels, are said to have contrasts in phonological quantity, in addition to the more common contrasts in phonological quality:

(2) Two basic types of phonological contrast:

Quality contrast	Quantity contrast
/b/ vs. /p/	/b/ vs. /bb/
/m/ vs. /n/	/m/ vs. /mm/
/i/ vs. /e/	/i/ vs. /i:/
/u/ vs. /o/	/u/ vs. /u:/

Drawing on data that we have gathered, it has been observed that MA does not impose any restrictions on the featural composition of long consonants, allowing all kinds of geminates. The result is that every singleton in the segmental inventory of MA has a geminate counterpart, which can be either lexical or derived. To be more specific, the geminate inventory of MA contains geminate stops, geminate fricatives, geminate nasals, geminate liquids and geminate glides.[3] For illustrative purposes, examples for each class of sounds are provided in (3):

(3) The segmental content of geminates in MA

<u>Obstruents</u>

Stops:	*Voiceless*	sə**kk**a	'rail'
	Voiced	ɣə**dd**a	'tomorrow'
Fricatives:	*Voiceless*	gə**ff**a	'sack'
	Voiced	ħa**ʒʒ**a	'a female pilgrim'

[3] Geminate glides in MA are restrictedly derived through morphological gemination. For example, morphological causatives that are based on medial weak roots with high vowels appear with corresponding geminate glides.

Sonorants

Nasals:	bə**nn**a	'flavor'
	sə**mm**an	'quail'
Liquids:	fə**ll**us	'hatchling'
	mə**ṛṛ**a	'one time'
Glides:	sə**ww**əl	'to ask'
	xə**jj**əṭ	'to sew'

MA is one of the few languages that allow all types of geminates. Other languages that put no segmental restrictions on the occurrence of their geminates include, Buginese (Podesva, 2000), a Western Austronesian language spoken in South Sulawesi, Hungarian (Polgardi, 2005) and Tashlhit (Boukous, 1982)[4]. Different other languages vary in the extent to which they constrain the occurrence of geminates. In Trukese, all consonants have geminate counterparts with the exception of glides (Hart, 1991). In Japanese, voiced consonants, be they sonorant or obstruent, may not geminate (Kubozono et al. 2008). Selayarese, which is closely related to Buginese, does not allow voiced obstruents (Podesva, 2000). In Somali, the only type of geminates permitted are voiceless stops (Blevins, 2008). In fact, it was cross-linguistically observed that the number of geminates, in languages that allow them, is always less than or equal to the number of singletons (Blevins, 2005).

Based on these language-specific observations, typological studies of geminates have suggested a number of universal generalizations. Cross-linguistically, geminate obstruents were found to be more frequent than geminate sonorants. The presence of the latter in a language usually implies the existence of obstruent geminates as well (Podesva, 2002; Kawahara, 2007).

[4] In Tashlhit, geminate glides can be underlying. Bensoukas (2001) provides the following items: ajjis 'horse' and awwun 'stone'.

Within the obstruent category itself, voiceless stops were observed to be more common than voiced stops and fricatives (Podesva, 2002; Maddieson, 2008).

4.2. The position of geminates

It has also been noted that geminates in MA can arbitrarily occur in word-medial, word-final as well as word-initial positions. MA has only a handful of words with lexical initial geminates. Yet, peripheral (or edge) geminates (i.e. initial and final geminates) may arise elsewhere due a process of total assimilation between the definite article and the first radical element of some nouns or the perfective 1st person pronoun and the final radical element of some verbs, as in ***ssuq*** and *ʃə**tt*** in (4).

(4) The position of geminates in MA

a. Word-initially		*b. Word-medially*		*c. Word-finally*	
mm^{w} i	'my mother'	sə**nn**a	'tooth'	ʒə**ṛṛ**	'to pull'
bb^{w}a	'my father'	gə**ff**a	'sack'	mu**xx**	'brain'
dda	'took'	ħə**nn**a	'henna'	ṛə**ʃʃ**	'to spray'
ssuq	'the market'	fə**zz**əg	'to make wet'	ʃə**tt**	'I saw'

Peripheral geminates are believed to be cross-linguistically less common while medial, more specifically intervocalic, geminates are the most preferred (Davis and Tropintzi, 2017; Pajak, 2013; Dmitrieva, 2012). Like MA, Tashlhit has geminate consonants in all positions (Bensoukas, 2001). Hungarian allows only medial intervocalic geminates and final ones (Polgradi, 2005). In Italian and Japanese, geminates can only occur intervocalically (Curtis, 2003). This implies that a language with peripheral geminates is more likely to contain word-internal geminates whereas a language with intervocalic geminates does not necessarily tolerate edge geminates. (Thurgood, 1993).

Two main approaches have been suggested to explain the crosslinguistic variation of geminate inventories. In this concern, variations in the nature and position of geminates have been argued to be the result of constraints on phonetic markedness or otherwise the outcome of distinct diachronic origins. The phonetically-based approach attributes the variation to perceptual difficulties. For instance, Kawahara (2007) suggests that geminate sonorants are disfavored because they are less perceptible than geminate obstruents. Similarly, Pajak (2009a, b) argues that positional asymmetry of geminates is also believed to follow from perceptual restrictions, in that geminates adjacent to consonants are harder to perceive by listeners than geminates adjacent to vowels.

According to the historical approach, variation in geminate inventories is believed to be related to the various historical sources and the diachronic evolution of geminates in the languages they include them (Blevins, 2004a-c). For Blevins, it is true that geminate consonants are more marked than short ones, but no geminate type or position should be seen as less or more exceptional than others, for it is equally possible that any geminate type in any position can evolve if the appropriate conditions are available. Blevins (2004a) identifies seven diachronic pathways through which geminates can evolve into phonemic segments. These are listed below:

(1) assimilation in consonant clusters

(2) assimilation between consonants and adjacent vowels/glides

(3) vowel syncope

(4) lengthening under stress

(5) boundary lengthening

(6) reinterpretation of a voicing contrast

(7) reanalysis of identical C+C sequences

All of these pathways refer to phonological processes that affect a sequence of non-identical or identical segments. However, we suggest that geminates can evolve from morphological change as well. To substantiate our claim, we refer to a category of verbs in MA that come in the same morphological shape of causative verbs, without expressing the meaning of causativity. Instead, they surprisingly behave more like simple verbs. Put differently, these verbs comprise medial geminates that seem to be central to their basic lexical meaning rather than being morphologically expressive. Some of these verbs are provided in (5):

(5) Apparent causatives

ɦə**ww**əd	'to get down'
nə**qq**əz	'to jump'
sə**ll**əm	'to greet'
wə**ll**əf	'to get used to'
wə**rr**i	'to show'

Nevertheless, if considered from a diachronic perspective, these forms can be claimed to be originally causative verbs that have lost their meaning of causativity and turned into simple verbs over time. We can adduce support for this claim by comparing the relevant verbs with some of their semantically-related corresponding nouns, like the ones cited below for each verb in (5):

(6) Corresponding nouns

ɦə**w**da	'a slope'
tən**q**iza	'a jump'
təs**l**ima	'a handshake'
wə**l**f	'getting used to'
twə**r**ja	'showing'

These forms suggest that what actually appears to be lexical geminates in the verbs in (5) have singleton counterparts in some distant semantically-related words. This somehow hints at their derivable nature. If that is really the case, then this can be seen as a striking example of morphological and semantic discrepancy, whereby some morphological pattern has diverged from its original meaning to express a new one. In this case, verbs that were initially morphological causatives have lost some of their semantic meaning but kept their templatic shape intact. Along that process, the relevant geminates have become morphologically inactive. By the same reasoning, we predict that more causative verbs would undergo this semantic shift, and therefore more geminates would become lexical-like segments.

4.3. Lexical geminates

In order to emphasize the above-mentioned observations, we proceed to show that geminates in MA are basically underlyingly-motivated, in the sense that a singleton-geminate alternation can invoke a change in meaning. This fact would firmly establish the contrastive nature of geminates in MA as phonemes that should be equally placed into the phonemic inventory of MA. The basic consonant inventory of MA looks as follows:

(7) The consonant inventory of MA

	Labial	*Alveolar*	*Alveo-palatal*	*Velar*	*Uvular*	*Pharyngeal*	*Glottal*
Stops	b, (ḅ)	t, ṭ d, ḍ		k, g	q		(ʔ)
Fricatives	f	s, ṣ z, ẓ	ʃ, ʒ		x, ɣ	ħ, ʕ	ɦ
Nasals	m, (ṃ)	n					
Liquids		l, (ḷ) r, ṛ					
Glides	w		j				

In order to securely establish the phonemic nature of geminates in MA, the data in (8) lists a number of minimal pairs -or near minimal pairs- that demonstrate how substituting a singleton with a geminate can lead to a total alteration in lexical meaning.

(8) The distinctiveness of geminates in Moroccan Arabic:[5]

a. *Word-medial contrast*

ħna	'we'	ħənna	'henna'
flus	'money'	fəllus	'hatchling'
mṛa	'woman'	məṛṛa	'one time'
ɣla	'being expensive'	ɣəlla	'crop'
gfa	'back of neck'	gəffa	'bag'
ħluf	'oath'	ħəlluf	'boar'

[5] In IPA, a geminate is commonly transcribed as two adjacent identical consonants (i.e. CC, where C stands for any consonant) e.g. [tt], [ss], [ll]. This way of transcribing geminates is conventionally motivated, and hence does not make any theoretical claims about how geminates are phonologically represented. All it tries to convey is that geminates are somehow phonologically and phonetically different from their singletons. In other systems of transcription, a geminate can also be transcribed with the length diacritic (e.g. /t:/) or with capitalized symbols (e.g. /T/).

bna	‘he built’	bənna	‘flavor’
ħaʒa	‘thing’	ħaʒʒa	‘a female pilgrim’
ʃalal	‘paralysis’	ʃəllal	‘waterfall’
ħmam	‘dove’	ħəmmam	‘bath’
ʒafaf	‘drought’	ʒəffaf	‘a floor mop’
brika	‘lighter’	bərraka	‘corrugated-metal house’
skər	‘to get drunk’	sukkəṛ	‘sugar’
smən	‘salted butter’	səmman	‘quail’
mika	‘plastic’	mikki	‘a cartoon character’
ħda	‘next’	ħədda	‘body building’
ɣda	‘lunch’	ɣədda	‘tomorrow’
bṛa	‘letter’	bəṛṛa	‘outside’

b. *Word-final contrast*

fik	‘in you’	fəkk	‘separate’
ḍar	‘house’	ḍərr	‘hurt’
ʒaṛ	‘neighbor’	ʒəṛṛ	‘drag’
ħal	‘situation’	ħəll	‘open’
mal	‘he leaned’	məll	‘get bored’
duq	‘taste’	dəqq	‘knock’
xud	‘take’	xədd	‘cheek’
dub	‘melt’	dubb	‘bear’
kul	‘eat’	kull	‘all’
ʕum	‘swim’	ʕəmm	‘uncle’
dum	‘last’	dəmm	‘blood’
buq	‘loudspeaker’	bəqq	‘bedbugs’
ʒud	‘generosity’	ʒədd	‘grandfather’
ʕiʃ	‘live’	ʕəʃʃ	‘nest’

c. *Word-initial contrast*

mat	'he died'	mm^{w}ət	'the mother of'
bat	'spend the night'	bb^{w}at	'the father of'
ssaṭ	'the sound of slapping'	saṭ	'boy' (slang)
bbaq	'the sound of falling'	baqi	'not yet'

The data demonstrates that geminates in MA phonemically contrast with singletons in all positions: initially, medially and finally. Like MA, languages such as Tashlhit, Italian, Japanese and Finnish also have contrasts between geminates and singletons. Some Examples are shown below:

(9) Geminate contrast in other languages:

Tashlhit: i**m**i 'mouth' – i**mm**i 'mother'
Italian: fa**t**o 'fate' – fa**tt**o 'fact'
Japanese: sa**k**a 'hill' – sa**kk**a 'author'
Finnish: ta**k**a 'back' – ta**kk**a 'fireplace'

We proceed by showing how these segments are produced by an OT grammar in MA. In OT, the interaction between markedness constraints and faithfulness ones determines the inventory of structures that are permitted to surface in output forms. For every feature specified in the input, there should be a faithfulness constraint that requires it to stay and a markedness constraint that militates against its presence in the output. If markedness is dominant, the relevant feature would be neutralized. However, if faithfulness ranks higher instead, the designated feature would be preserved, and thus included in the segmental inventory of the language under study. Either way, one of the competing constraints would be violated by the optimal candidate, and that should be the one ranked lower by the grammar.

The consonantal inventory of MA consists of singletons and geminates alike. In fact, every singleton seems to unrestrictedly have a geminate counterpart. In English, as in many other languages, geminates are unattested, in the sense that consonantal length is not distinctive in these languages. The markedness constraint that penalizes the appearance of geminates is named *GEM. If this constraint is dominating faithfulness, geminates will be absent from the grammar, English is a case in point. In MA, however, *GEM should be dominated by the faithfulness constraint MAX-GEM, which demands the input specification for consonantal length to be preserved in the output. In this way, whenever a geminate is posited in the input, it would find its way to the output form thanks to the ranking MAX-GEM >> *GEM.

(10) Lexical geminates

/tt/	MAX-GEM	*GEM
a. t	*!	
☞b. tt		*

(11) Constraints responsible for underlying geminates in MA

a. **MAX-GEM**: an input Length specification and its output correspondent must be identical.

b. ***GEM**: geminate consonants are prohibited.

Geminate consonants are generally marked structures, but some geminates are more marked than others. Based on typological and perceptual evidence, it was proposed that the constraint *GEM should be broken down into a number of sub-family constraints that target specific instantiations of geminates. As has been mentioned earlier, languages tend to prefer some geminates more than others, considering their segmental composition and lexical position (see Podesva, 2002; Kawahara, 2007; Pajak, 2010). For instance, Kawahara (2007) and Pajak (2009a, b) suggested the following universal rankings to account for the cross-linguistic variations in geminate occurrences:

(12) Segmental markedness hierarchy (Kawahara, 2007)

*GemGlide >> *GemLiquid >> *GemNasal >> *GemObstruent

(13) Contextual markedness hierarchy (Pajak, 2009a, b)

non-vowel adjacent	>> single vowel-adjacent	>> intervocalic
*#GC, *CG#, *CGC;	*#GV, *VG#, *VGC, *CGV;	*VGV

According to the segmental markedness hierarchy, geminate glides appear to be the least favored type of geminates whereas obstruent geminates are more tolerated by virtue of being down in the hierarchy. Note that this hierarchy is based on a sonority scale. The more sonorous a geminate gets, the less favored it becomes. According to the contextual markedness hierarchy, however, geminates are mostly found intervocalically, while consonant-adjacent geminates are less common. However, since MA allows all sorts of geminates, irrespective of their segmental content or place in the word, the only relevant constraint for us is the general constraint *GEM. On the contrary, the vocalic inventory of MA does not contain long vowels as they never phonemically contrast with short vowels in the language. The vocalic inventory of MA is limited to three underlying short vowels /i, u, a/ and an epenthetic schwa /ə/. For the sake of illustration, the vocalic inventory of MA is schematized below:

(14) The vocalic inventory of MA

High	i		u
Mid		(ə)	
Low	a		

For this reason, the markedness constraint that sanctions the presence of long vowels, written as *V:, is set to dominate faithfulness (*V: >> IDENT-IO-Long-V).

(15) Long vowels are not attested in MA

/V:/	*V:	IDENT-IO [Long-V]
☞ a. V		*
b. V:	*!	

Basically, there should be no restrictions imposed on underlying structures. In other words, the grammar is free to posit any type of input. This is a central property of the lexicon in OT, dubbed the Richness of the Base. The basic idea is that structural well-formedness should be determined solely through constraint interaction without any stipulations at the level of underlying representations. It is then the role of constraint hierarchy (i.e. Eval) to determine which forms would surface and which would not.

4.4. Phonological geminates

4.4.1. Geminates through assimilation

Not only do geminates function as distinctive segments that contribute lexical contrast in MA, but they also happen to be phonologically derived. The bulk of initial geminates in MA are derived by a process of total assimilation that occurs between the definite article /l+/ and nouns whose first consonant is a coronal sound. The affected segment is the definite article morpheme, whose feature specifications completely change to match the feature specifications of whatever coronal sound the designated noun begins with. This case of assimilation is illustrated by the following items:

(16) Phonologically derived geminates

a. Assimilating nouns

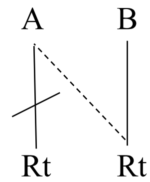
/l-ʃəms/	ʃʃəms	‘the sun’
/l-ḍar/	ḍḍar	‘the house’
/l-suq/	ssuq	‘the market’
/l-tuma/	ttuma	‘garlic’
/l-ʒlbana/	ʒʒəlbana	‘green peas’

b. Non-assimilating nouns

/l-qamaṛ/	lqamaṛ	‘the moon’
/l-ħiṭ/	lħit	‘the wall’
/l-bab/	lbab	‘the door’
/l-qamiʒa/	lqamiʒa	‘the shirt’
/l-kina/	lkina	‘the pill’

In the theory of feature geometry (Clements, 1985; Sagey, 1986), which assumes an internal hierarchical structure of features, total assimilation is believed to involve the spreading of the root node of the triggering segment to replace the root node of the targeted segment, the replaced root node is then deleted.

(17) Total assimilation in feature geometry:



This representation implies that all features of the trigger, Segment B, spread to replace the already specified features of the affected segment A. This process gives rise to a double-linked root node that spans two skeletal positions. The theory assumes a multiple association analysis of length, whereby long vowels and geminates are linked to two slots at the skeletal tier:

(18) Underlying long segments:

As a result, geminates derived by assimilation and those that are lexical appear to be identically represented as segments multiply associated to two slots on the timing tier. Therefore, they are posited to be distinguished from sequences of identical consonants across morpheme boundaries. These are normally represented as two independent root nodes, each linked to a timing slot.

(19) Identical heteromorphemic consonants

$$\begin{array}{cc} \mathrm{C} & \mathrm{C} \\ | & | \\ \mathrm{Rt} & \mathrm{Rt} \end{array}$$

The fact that geminates derived by total assimilation are represented the same as underlying geminates, with one melodic element associated to two skeletal slots, is supported by phonological evidence. Despite being derived from a sequence of two heteromorphemic consonants, these geminates behave the same as underlying geminates with respect to rules of epenthesis, in that they resist the application of the latter by remaining unbroken, unlike consonant clusters. For instance, in MA, geminates created by total assimilation behave as a monolithic segment, rather than a cluster of consonants:

(20)

a.	/l-sma/	ssma	*səsma	‘the sky’
	/l-ḍṛa/	ḍḍṛa	*ḍəḍṛa	‘the corn’
b.	/ktṛa/	kətṛa		“plenty”
	/ħʃma/	ħəʃma		“shyness”

Were they clusters of identical consonants, the derived geminates in (20a) would get split by a schwa in the same way the clusters in (20b) are. Nevertheless, schwa epenthesis is blocked in those items.

With this in mind, we claim that the assimilation process affecting the definite article is motivated by a dictate of the constraint AGREE-Coronal, which requires identical feature specifications of adjacent coronal sounds. The demand of this constraint takes effect when it dominates the faithfulness constraint IDENT-IO (F), militating against any change of input features. The ranking relations of these constraints are illustrated in the tableaux below:

(21) AGREE-Cor >> IDENT-IO (F): assimilation forces violation of faithfulness

/l-suq/	AGREE-Cor	IDENT-IO (F)
☞ a. ssuq		*
b. l-suq	*!	

This tableau shows that satisfaction of AGREE-Cor is more important than satisfaction of IDENT-IO (F). Therefore, the feature specifications of the root consonant and the affix should identically agree at the expense of violating faithfulness.

So far, the directionality of the assimilation process would be assumed to be governed by a faithfulness constraint which demands the preservation of the featural identity of the initial root consonant at the expense of the featural identity of the definite article affix. Such a constraint is formally written as IDENT-RtC_1 –where Rt and C_1 refer to the root and the first radical element, respectively.

In a previous analysis (see Rguibi, 2001:104), the whole root was seen as a privileged position that triggers total assimilation. For us, however, it is precisely the root-initial segment which triggers assimilation with the definite article, especially that the shared features are contributed by this particular position. Our view follows from the fact that roots are not fully immune to phonological change. More specifically, elsewhere, root-final and root-medial segments are prone to phonological alternation like in the case of causative derivation, whereby root-medial segments get geminated (e.g. √ktb → kəttəb) or in the case of the 1st person singular morpheme suffixation in which root-final segments get assimilated (e.g. /mṛəḍ-t/ → mṛətt). It remains that the only position which is consistently privileged over the others is the root-initial segment.

(22)

/l-suq/	AGREE-Cor	IDENT-IO (F)	IDENT-RtC1
☞a. ssuq		*	
b. l-suq	W*!	L	
c. lluq		*	W*!

Candidate (22b) is out early in the competition as it violates the high ranked constraint AGREE-Cor. The optimal candidate (22a) ties with candidate (22c) since they equally satisfy AGREE-Cor and violate IDENT-IO (F). Candidate (22c) satisfies AGREE-Cor differently, however. It spreads the features of the definite article and deletes those of the root consonant. Here comes the role of the constraint IDENT-RtC1 as a tie breaker. In OT, the ranking of tie-breakers, in connection with the other constraints, is irrelevant. This constraint privileges the maintenance of the feature specifications of the root consonant over the maintenance of the features of the affix. Candidate (22c) violates this constraint, and hence loses in the competition. Candidate (22a) wins out.

If a noun begins with a non-coronal sound however, the structural description of the constraint AGREE-Coronal would not be complete. In this case, assimilation would not take place. A candidate where the definite article morpheme would assimilate to a non-coronal sound would be penalized, and thus ruled out by IDENT-IO (F). The non-assimilating candidate fares better on the ranking since it satisfies both AGREE-Cor and IDENT-IO (F).

(23) AGREE-Cor has no effect: No assimilation

/l-ħiṭ/	AGREE-Cor	IDENT-IO (F)
a. ħħiṭ		*!
☞b. lħiṭ		

(24) The constraints responsible for the assimilation of the definite Article in MA:

a. **AGREE-Coronal**: adjacent coronal segments must be identical feature-wise.

b. **IDENT-IO (F)**: Corresponding segments in input and output must have identical feature values.

c. **IDENT-RtC$_1$**: an input feature specification in the first segment of the root and its output correspondent must be identical.

Despite the fact that the emergent structure of this assimilation process is a word initial geminate, in violation of the markedness constraint *GEM, the latter does not seem to have any active role in this assimilatory process. Geminates happen to follow naturally from the interaction between AGREE-Coronal and IDENT-IO (F). However, the created geminate structure is set to comprise the morphological exponence of both the assimilated affix and the initial root consonant. As a result, the grammar requires the faithfulness constraint MAX-Affix to dominate *GEM in order to ensure the realization of the definite article in the output form, and hence prevent *GEM from cancelling out the derived geminate structure. Without this dominance relation, the definite article affix could be compromised in favor of obviating a geminate structure.

(25) MAX-Affix >> *GEM: the definite article affix should have a correspondent in the output

/l-suq/	MAX-Affix	*GEM
☞a. ssuq		*
b. suq	*!	

This ranking ensures that the definite article affix has a correspondent in the output form, instead of being deleted for the sake of avoiding the creation of geminates. In this case, the violation of *GEM is less costly than the violation of MAX-Affix.

There is another case of total assimilation that gives rise to geminates in MA. This takes place between the final radical element of verbs and the 1st person past tense suffix [-t]. The assimilation is consistently triggered by coronal stops and exceptionally occurs with other sounds. The data in (26a) shows that if a verb ends in a coronal stop, it then totally assimilates to the suffix [-t], deriving the geminate /tt/ in every case. The items in (26b) are exceptions to this generalization. The relevant verbs therein end in a fricative and labial stop, respectively. Yet, the data in (26c) includes verbs ending in the same sounds without triggering any assimilation. On this basis, we render the items in (26b) as exceptions. (26c) also emphasizes the fact that only coronal stops trigger assimilation and not just any coronal. (26d) is provided to show that non-coronals never get involved in this assimilation interaction.

(26)

(i) Assimilating verbs

a.	/ʒbəd-t/	ʒbə**tt**	‘I pulled out’
	/ṣrəṭ-t/	ṣrə**tt**	‘I swallowed’
	/ɦbəṭ-t/	ɦbə**tt**	‘I came down’
	/ħfəḍ-t/	ħfə**tt**	‘I memorized’
	/mrəḍ-t/	mrə**tt**	‘I got sick’
b.	/ʃəf-t/	ʃə**tt**	‘I saw’
	/ʒəb-t/	ʒə**tt**	‘I brought’

(ii) Non-assimilating verbs

c.	/ħləf-t/	ħləft	‘I swore’
	/ʕṛəf-t/	ʕrəft	‘I knew’
	/ħməl-t/	ħməlt	‘I got pregnant’
	/nʒəṛ-t/	nʒəṛt	‘I sharpened’
	/ħbəs-t/	ħbəst	‘I stopped’
d.	/rkəb-t/	rkəbt	‘I mounted’
	/ħləm-t/	ħləmt	‘I dreamed’
	/dfəʕ-t/	dfəʕt	‘I pushed’
	/dəq-t/	dəqt	‘I tasted’
	/dəx-t/	dəxt	‘I got confused’

The constraints and interactions responsible for this effect are akin to the ones involved in the definite article assimilation. The basic interaction is: AGREE-Coronal-Stop outranking IDENT-IO (F). This ranking selects assimilated outputs over non-assimilated ones. The Agree constraint demands that adjacent coronal-stops should have identical feature specifications. The directionality of the assimilation process in question seems to be governed by IDENT-Affix. Such a constraint penalizes any candidate that sacrifices the feature specifications of the affix.

(27)

/ʒbəd-t/	AGREE-Cor-stop	IDENT-IO (F)	IDENT-Affix
☞a. ʒbətt		*	
b. ʒbədt	W*!	L	
c. ʒbədd		*	W*!

As we go over the competing candidates, Candidate (27a) emerges as the winner as it satisfies the AGREE constraint as well as the faithfulness constraint preserving the identity of the suffix, IDENT-Affix. Candidates (27b) loses for their violating AGREE. Despite obeying AGREE, candidate (27c) is ruled out due to its violation of IDENT-Affix.

By way of comparison, the assimilatory processes discussed above point out to an anomalous situation, whereby the directionality of assimilation is determined by means of root faithfulness in the first case and affix faithfulness in the second. The problem is that this implies that root faithfulness outranks affix faithfulness in the context of the definite article assimilation, while affix faithfulness outranks root faithfulness in the context of 1st person assimilation. In light of the inadequacy of positional faithfulness to consistently derive the directionality of assimilation in MA, we follow Pater and Werle (2001, 2003) in encoding directionality into the standard AGREE constraint in the following way:

(28) **AGREE-Left [Coronal]**:

A coronal consonant preceding another coronal must be identical with it.

This will allow us to unify both assimilatory processes under one analysis as demonstrated by the tableaux below:

(29) Definite article assimilation

/l-suq/	AGREE-L-Cor	IDENT-IO (F)
☞ a. ssuq		*
b. l-suq	*!	
c. lluq	*!	

(30) 1st person pronoun assimilation

/ʒbəd-t/	AGREE-L-Cor-Stop	IDENT-IO (F)
☞ a. ʒbətt		*
b. ʒbədt	*!	
c. ʒbədd	*!	

This means that the new version of the AGREE constraint can be violated in two ways. First, it can be violated by candidates where assimilation does not take effect. Second, it can also be violated by candidates whose assimilation is progressive (i.e. left-to-right) instead of being regressive (i.e. right-to-left).

4.4.2. Gemination as prosodic structure improvement

Assimilation is not the only phonological process that gives rise to geminates in MA. In fact, geminates can also be created by means of a process of consonant lengthening that serves the purpose of improving the prosodic structure of forms involved. In the data below, it could be argued that the suffixation of the /u/, representing the 3rd person masculine pronoun, triggers the lengthening of the consonant /t/ to supply an onset for the newly created syllable. Or else, it could be said that gemination is motivated by the need to compensate for the deletion of the vowel /a/ of the affix /at/. In either case, the outcome of this process consists in the creation of the geminated affix [tt] in the relevant forms. Some illustrative examples are shown below:

(31) Gemination by syllabification

a. Vowel deletion

/ḍəṛbat/	ḍəṛbət	‘she hit’
/ʃəṛbat/	ʃəṛbət	‘she drank’
/qətlat/	qətlət	‘she killed’
/nəʃṛat/	nəʃṛət	‘she hung’

b. Gemination

/ḍəṛbət-u/	ḍəṛbə**tt**u	‘she hit him’
/ʃəṛbət-u/	ʃəṛbə**tt**u	‘she drank it
/qətlət-u/	qətlə**tt**u	‘she killed him’
/nəʃṛət-u/	nəʃṛə**tt**u	‘she hung it’

Throughout this dissertation, we will revisit this set of data on multiple occasions. In chapter 4, a formal optimality-theoretic analysis will be provided for the observed alternation, in compliance with the phonological representation of geminates that will be argued for in chapter 3.

4.5. Morphological gemination

Besides being lexical or phonological, geminates in MA come in a third flavor, in that they can be morphologically induced. In particular, geminates in MA can work as morphological markers in derived forms like morphological causatives, agent nouns and instrument nouns. These forms appear with a medial geminate that corresponds to a singleton in the base form. In the case of causatives, the sole morphological marker is the geminate itself. As for agent and instrument nouns, additional morphological material is present, namely a vowel /a/ marking nominality in both structures and an extra /a/ specific to instrument nouns, behaving as a feminine marker. Examples of these derived forms are provided below:

(32) Morphologically derived geminates:

(i) The causative verb

ktəb	'to write'	kəttəb	'to make write'
ɦṛəb	'to run away'	ɦəṛṛəb	'to make run away'
xsər	'lose'	xəssər	'to cause lose'
zləq	'slip'	zəlləq	'to cause slip'
ʃṛəb	'drink'	ʃəṛṛəb	'to make drink'

(ii) The agent noun

fləħ	'to farm'	fəllaħ	'farmer'
nʒəṛ	'to sharpen'	nəʒʒaṛ	'carpenter'
gzər	'to butcher'	gəzzar	'butcher'
bni	'to build'	bənnaj	'mason'
ʃuf	'to see'	ʃəwwaf	'fortuneteller'

(iii) The instrument noun

ɣsəl	'to clean'	ɣəssala	'a washing machine'
sməʕ	'to listen'	səmmaʕa	'headset'
skət	'to be quite'	səkkata	'pacifier'
kwi	'to weld'	kəwwaja	'a welding machine'
sqi	'to water'	səqqaja	'fountain'

MA is not alone in using gemination for morphological purposes. Several other patterns of morphological gemination occur in languages like Tashlhit (Bensoukas, 2001) Choctaw (Lombardi and McCarthy, 1991) and Keley-i (Samek-Lodovici, 1992). We will be devoting a whole chapter to morphological gemination in MA. Our treatment of this category of geminates will be underlain by the assumption that morphologically-derived geminates should have the same phonological structure as underlying geminates. As it stands, this issue lies right at the intersection between phonology and morphology, in that morphological information is conveyed by means of phonological variation.

5. True vs. fake geminates

A distinction is in order between true geminates and false (or apparent) ones. True geminates are those that are base-generated (i.e. lexical) or derived via assimilation or morphological lengthening. In other words, a language is said to have true geminates if and only if one of the following is true: (i) they contrast phonemically with singletons (ii) they appear morpheme internally (iii) they get derived by total assimilation or (iv) they mark morphological exponence. MA is an example of a language where all the above conditions hold true. As a result, it is very safe to say that the language contains true geminates.

Fake geminates, on the other hand, are identical consonants that accidentally occur across morpheme or phrase boundaries. These may even appear in languages that lack true geminates. For instance, the examples below from English and MA alike represent cases of fake geminates which span different morphemes in the absence of any obvious assimilation. In English, this is very common in compounds. The examples from MA are past participle forms where the [m-] of the past participle morpheme coincides with the /m/ of the verb stems [mədd], [mnəʕ] and [mlək].

(33) Fake geminates in English

un-natural	/ʌ**n**ˈ**n**ætʃəɹəl/
rat-tail	/ˈræ**t**ˌ**t**eɪl/
book-case	/ˌbʊ**k**ˈ**k**eɪs/

(34) Fake geminates in MA

mmdud		məmdud	'stretched
mmnuʕ	[hypercorrected forms of]	məmnuʕ	'forbidden'
mmluk		məmluk	'owned'

Schwa epenthesis is an effective diagnosis to tell fake and true geminates apart in MA. True geminates normally block schwa epenthesis while fake ones allow it. This fact can be attributed to the different phonological representations that underlie each one. In autosegmental CV phonology (Goldsmith, 1976; McCarthy, 1979), true geminates are represented as one melodic element associated to two slots in the skeletal tier. Hence, given the non-crossing association lines constraint that regulates autosegmental representations, epenthesis is not permitted. (35) below provides a visual illustration. On the other hand, fake geminates are represented as two independent melodic elements linked to distinct positions. Under this representation, schwa epenthesis can take place without violating the non-crossing association lines constraint. This is exemplified by (36).

(35) True geminates block epenthesis

(36) Fake geminates allow epenthesis

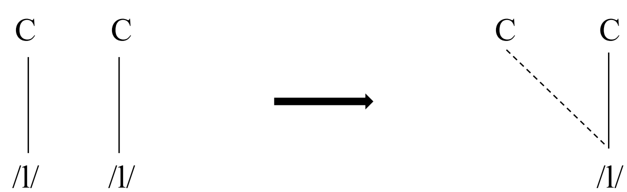
As a result, when two identical heteromorphic consonants occur side-by-side by sheer coincidence, they should be distinguished from geminate consonants derived by assimilation. However, the examples in (37) from MA could be a source of confusion.

(37)

UR	*Derived forms*	
/l-lubja/	llubja	'the beans'
/l-limun/	llimun	'the orange'
/l-luʕba/	lluʕba	'the toy'
/l-lil/	llil	'the night'
/l-lun/	llun	'the color'

What is of interest here is that we have a case of two identical consonants appearing across morpheme boundaries. In this particular case, the definite article /l/ coincides with nouns whose initial segment is the sound /l/. The question now is whether to consider this sequence a case of fake geminates or a case of true geminates that are derived. We have seen earlier that MA has a process of total assimilation between the definite article and any initial coronal sound belonging to the noun it attaches to, creating derived geminates. When a noun begins with a non-coronal sound, the definite article and the relevant non-coronal preserve their status as two independent segments. In (37), the structural description of the constraint AGREE-Coronal is satisfied, hence assimilation is expected to take place. The items in (37) could illude us in believing that we have a case of fake geminates, while in fact the assimilation process changes the structure of the underlying forms as follows:

(38)



This can be adequately corroborated by the behavior of the derived geminates in the following items:

(39)

/l-lɦa/	llɦa	*ləlɦa	‘hard palate’
/l-lɣa/	llɣa	*ləlɣa	‘gossip’

It has been shown earlier that schwa epenthesis represents a firm diagnosis that can mark off true geminates from fake ones. By the same standard, the geminates in (39) are entitled to the true-geminate status since schwa epenthesis is blocked from applying to them. Were they mere sequences of heteromorphemic identical segments, schwa epenthesis would have split them in the same way it does with fake geminates in the following examples:

(40)

/m-mdud/	məmdud	‘stretched
/m-mnuʕ/	məmnuʕ	‘forbidden’

While Heselwood and Watson (2013) concedes that geminates like those in (16) and (37) satisfy the description of true geminates by virtue of observing geminate integrity, it does not assume that they are derived through total assimilation, at least synchronically speaking. The authors believe that those geminates behave more like lexical (i.e. non-derived) geminates, and hence should be depicted as such. This view was based on the conviction that assimilation should typically be optional, meaning that non-assimilating forms should be allowed to occur in the grammar in question. The process of total assimilation that affects the definite article in Arabic does not allow optionality, however, in that only assimilated forms are possible. Heselwood and Watson (2013) suggests that these geminates should be treated as diachronically fossilized geminates that have become lexicalized though their morphological trace is still known. Whether that is true or not, for us, even in the case of a synchronic assimilation process, which is how we treat it, the emergent phonological structure will be the same as that of lexical geminates (i.e. a multiply associated melodic element). This suffices to explain why any geminate, be it derived or lexical, blocks epenthesis.

Nevertheless, there is no phonetic clear-cut between true and fake geminates since both can be “characterized by a single articulatory gesture, or by two distinct articulatory gestures” (Blevins, 2004:170). Phonologically speaking, true geminates are most likely to be represented differently from fake ones, given their distinct behavior with regards to phonological rules in many languages. However, there is the view that geminates that can appear morpheme-internally and those that arise across morpheme boundaries are no different, in that they behave similarly regarding certain phonological processes, hence both should be represented as normal consonant clusters. For example, Saib (1977) argues for a sequential analysis of geminates in Berber based on their behavior vis-à-vis consonant clusters. Accordingly, it would be unnecessary to have a distinction between fake and true geminates in such a case.

In this work, we are concerned with true geminates. In particular, we are interested in the phonological representation of true geminates and their behavior in contrast with singletons and consonant clusters. In addition, we endeavor to examine the syllabic properties of geminates as well as their role as morphological markers. Thus, throughout this work, the term ‘geminate’ will be used to refer exclusively to true geminates, unless otherwise specified.

6. The lexical distribution of geminates

Geminate consonants in MA occur in most of the major lexical categories of MA, namely nouns, verbs and adjectives. However, these grammatical categories differ in terms of the extent to which geminates are manifested by their phonological and morphological content. In particular, it has been observed that the existence of geminates is maximized in nouns and verbs while being minimized in adjectives.

In nouns, for example, geminates can be underlying, phonological as well as morphological. In terms of position, they can be initial, medial or final. Morphologically-derived geminates in nouns are typically word-medial, characterizing the derivation of agent and instrument nouns. Phonologically-derived geminates in nouns are the result of total assimilation between the definite article and the first radical element, provided the latter is a coronal sound. Geminates rising from this process happen to regularly appear word-initially.

In verbs, with the exception of the verb [ddi] ‘take’, initial geminates are very uncommon. Final geminates in verbs can be both phonological or underlying. As for medial geminates in verbs, they are mostly morphological, distinguishing a large class of causative verbs. In adjectives, however, more precisely in some comparatives, only final underlying geminates are attested.

This makes of nouns the lexical category where geminates are most pronounced, followed by verbs and adjectives, in that order. By making this observation, we aim to draw the attention of other researchers to this type of lexical asymmetry, hoping that it would be addressed in the future.

7. Conclusion

In this chapter, we have provided a taxonomy of geminate patterns in MA. We have shown that, in MA, geminates can be lexically motivated or emerge as the result of specific phonological and morphological processes. Phonologically, it was demonstrated that geminates can be derived through a process of total assimilation on two occasions: (i) total assimilation between the definite article and the initial coronal consonant of nouns or (ii) total assimilation between the first-person past pronoun and the final coronal-stop consonant of verbs. Prosodic lengthening was also pointed out to give rise to phonologically derived geminates. Morphologically, it has been shown that geminates can play the role of a morphological marker, leading to the formation of morphological causatives, agent nouns and instrument nouns. In terms of position, geminates were shown to occur word-initially, word-medially and word-finally. Also, it has been pointed out that geminates may come from every major class feature that is present in the phonological system of MA.

What's more, this chapter has made a distinction between true geminates and fake ones. In this regard, we have argued that true geminates are not necessarily underlying but can also be derived. This was supported by the fact that derived geminates in MA tend to resist schwa epenthesis in the same way underlying geminates do.

CHAPTER THREE

GEMINATES IN MOROCCAN ARABIC: BEHAVIOR AND REPRESENTATION

1. Introduction

A significant cross-linguistic characteristic concerning the phonological behavior of geminates is that they sometimes combine properties of sequence structures as well as singleton structures. In other words, within the same language, geminates may simultaneously pattern with consonant clusters with respect to certain rules and with short consonants with regards to other rules. This dual function of geminates has made it hard for phonologists to provide a consistent phonological representation that can account for this observed, and apparently erratic, behavior (Leben, 1980; Saib, 1977). Furthermore, the treatment of geminate consonants becomes even more challenging when phonologists try to put forward a cross-linguistic account of geminate tendencies. Every attempt in this direction is marred by major cross-linguistic irregularities and variations.

As has been previously shown for MA, geminate consonants permeate the MA language system. Not only do they phonemically contrast with short consonants (e.g. mṛa 'woman' vs. məṛṛa 'one time'), but they also happen to be phonologically and morphologically derived (e.g. l-suq vs. ṣṣuq 'the market'; ktəb vs. kə**tt**əb 'write'). MA is an exemplary language where geminates display characteristics of both consonant sequences and singletons. Evidence for this comes from the distribution of geminates and their behavior in response to a number of phonological and morphological processes. Our discussion of the patterning of geminates in MA will be substantiated by data from various other languages.

This chameleonic behavior of geminates raises the following pressing question: Should geminates be interpreted as a sequence of short consonants, as a single long consonant or as a segment that is similar to and, at the same time, different from both? In answering this question, different models of phonological representation will be evaluated for their ability to adequately capture and explain the behavior of geminates in MA and elsewhere.

2. The dual behavior of geminates in MA

2.1. Geminates as a sequence structure

To begin with, geminate consonants in MA have distributional properties that are identical to those of consonant clusters. As the language allows consonant clusters in all positions, they can be word-medial, word-initial as well as word-final.

(1) The distribution of clusters in MA

Word-initially		*Word-medially*		*Word-finally*	
ʕ.dəm	‘a bone’	ṭər.ʒəm	‘to translate’	mər.d	‘illness’
ʃ.rəb	‘to drink’	ʃər.ʒəm	‘a window’	kəl.b	‘dog’
ʒ.bəd	‘to pull’	kər.kəb	‘to roll’	fər.x	‘bird’

Word medial clusters extend across two syllables. The first member of the cluster is typically syllabified as the coda of the previous syllable and the second member is syllabified as the onset of the ensuing syllable. At the margins, however, complex onsets and codas are not tolerated in MA. Thus, when clusters appear at the margins, the spare consonant is syllabified as the head of its own syllable. Interestingly, geminates are distributed similarly.

(2) The distribution of geminates in MA

Word-initially		*Word-medially*		*Word-finally*	
<u>d.d</u>i	'to take'	sə<u>n.n</u>a	'tooth'	sə<u>d.d</u>	'to close'
<u>s.s</u>uq	'the market'	gə<u>f.f</u>a	'bag'	ʒə<u>r.r</u>	'to pull
<u>l.l</u>imun	'the orange'	sa<u>ʕ.ʕ</u>a	'clock'	ʕə<u>d.d</u>	'to bite'

For this reason, geminates tend to be syllabified in much the same way as clusters, hence arguably regarded as a cluster of two identical singletons. This is subject to a great deal of controversy as will be discussed in due time. The distribution properties of geminates and clusters in MA can be summed up as follows:

(3) The distribution and syllabification of geminates and consonant clusters in MA

	CC clusters	*Geminates*
Word-medially	$VC_x.C_yV$	$VC_x.C_xV$
Word-initially	$C_x.C_yV$	$C_x.C_xV$
Word-finally	$CVC_x.C_y$	$CVC_x.C_x$

In addition, the language has a common regular rule of schwa epenthesis which takes place for the purpose of breaking up impermissible consonant clusters. In consonant-only words, the domain of such a rule is restricted to a minimum sequence of three unparsed consonants (e.g. CCC). That is, whenever a sequence of more than two unsyllabified consonants occurs, an epenthetic schwa breaks it. Consider the following illustrative examples:

(4) Schwa epenthesis in MA

	Root	*Verb*	*Gloss*
a.	/ʃṛb/	ʃṛəb	‘to drink’
	/qtl/	qtəl	‘to kill’
	/ktf/	ktəf	‘shoulder’
b.	/tṛʒm/	təṛʒəm	‘to translate’
	/ʃṛʒm/	ʃəṛʒəm	‘window’
	/ʃṛml/	ʃəṛməl	‘to marinate’

These examples show that a cluster of three consonants triggers the epenthesis of only one schwa, while a sequence of four consonants triggers the epenthesis of two schwas. On this basis, it is believed that if geminates were single consonants, sequences like those in (5a) are not expected to trigger schwa epenthesis at all and those in (5b) are expected to trigger the epenthesis of only one schwa.

(5) Geminates in a sequence

a.	/s**dd**/	‘close’	b.	/k**tt**b/	‘write’
	/ħ**ll**/	‘open’		/ɦ**ṛṛ**b/	‘escape’
	/ʃ**dd**/	‘catch’		/t**ll**f/	‘lose’
	/d**qq**/	‘knock’		/ʃ**ṛṛ**b/	‘drink’
	/ʃ**qq**/	‘crack’		/k**mm**ʃ/	‘wrinkle’

That is, if the roots in (5a) were bi-consonantal, the domain of schwa epenthesis would be unsatisfied. Yet, schwa epenthesis does take place as far as these sequences are concerned. This entails that their geminates count as two consonants, which together with the first consonant make a three-consonant cluster. Also, if the structures in (5b) were tri-consonantal, counting their geminates as single long consonants, we would expect only one epenthetic schwa in their output forms. However, these structures trigger the epenthesis of two schwas in the same way the quadrisegmental roots in (4) do.

(6) Geminates and schwa triggering

a.	sədd	‘to close’	b.	kəttəb	‘to make write’
	ħəll	‘to open’		ɦəṛṛəb	‘to make escape’
	ʃədd	‘to catch’		təlləf	‘to make disappear’
	dəqq	‘to knock’		ʃəṛṛəb	‘to make drink’
	ʃəqq	‘to crack’		kəmməʃ	‘to wrinkle’

One could claim that schwa epenthesis breaks any number of consonant clusters, including bi-consonantal ones. This could explain the position of the schwa in (6a), where it occurs between the geminate and the first radical consonant. But, such a claim would mean that a schwa should break any sequence of two consonants, including those in (4), reiterated in (7) with the expected structure under this view.

(7) A schwa between each two consonants

a.	/ʃṛb/	*ʃəṛəb
	/qtl/	*qətəl
	/ktf/	*kətəf
b.	/tṛʒm/	*təṛəʒəm
	/ʃṛʒm/	*ʃəṛəʒəm
	/ʃṛml/	*ʃəṛəməl

Nevertheless, these structures are ungrammatical, which emphasizes the fact that the domain of schwa epenthesis is rather a sequence of three consonants or more. If (6a) is viewed as a case of schwa epenthesis in the context of a bi-consonantal root, then (7) should also be possible. By way of explanation, if schwa is capable of creating bimoraic prosodic words of the shape C_1əC_2, nothing should prevent it from occurring in open syllables as in (7).[6] Therefore, the geminates in (6a) cannot be counted as one single segment, but as two. As a result, we can maintain the generalization that words like C_1əC_2 and syllables like Cə are equally ungrammatical. With that said, geminates appear to behave as a sequence of consonant clusters with respect to the triggering of schwa epenthesis.

There is another case of epenthesis in MA that occurs exclusively in verbs ending in geminates. As the data in (8) illustrates, when these verbs are inflected for the past form, the vowel /i/ is inserted between the geminate and the perfective personal pronoun if the latter starts with a consonant. Were it a single segment, the geminate would be fully contained in the first

[6] C_1əC_2 word forms in MA do not belong to major lexical categories (e.g. mən 'from'). This is to say that word minimality condition applies primarily to content words.

syllable, functioning as a coda. However, it behaves as a cluster of two independent segments with the second member triggering epenthesis and forming a syllable of its own.

(8)

a.	/sdd+t/	s$_1$əd$_2$.d$_3$**i**t	'I closed'	*s$_1$ədd$_2$.t
	/sdd+na/	s$_1$əd$_2$.d$_3$**i**.na	'we closed'	*s$_1$ədd$_2$.na
	/sdd+ti/	s$_1$əd$_2$.d$_3$**i**.ti	'you closed'	*s$_1$ədd$_2$.ti
	/sdd+tu/	s$_1$əd$_2$.d$_3$**i**.tu	'you (pl.) closed'	*s$_1$ədd$_2$.tu
b.	/sdd/	səd.d	'he closed'	
	/sdd+at/	səd.dat	'she closed'	
	/sdd+u/	səd.du	'they closed'	

Note that epenthesis does not take place if the relevant personal pronoun starts with a vowel or if the verb stem does not take any pronoun, as illustrated by (8b).[7]

Another piece of evidence showing the patterning of geminates with consonant clusters comes from the derivation of the instrument noun from roots ending in a geminate. At this level of the discussion, it is assumed that the instrument noun is derived by doubling the second radical segment of the base form and affixing a vowel /a/ before and after the last consonant. A more detailed account of this data will be provided in chapter 8.

[7] We will have more to say about this pattern in the chapter 5. More specifically, this process will be analyzed in the light of our proposed phonological representation of geminates, using the constraint-based framework of OT.

(9)

	Root		Instrument noun	
a.	/ɣsl/	‘to wash’	ɣəssala	‘washer
	/kwa/	‘to weld’	kəwwaja	‘welding machine’
	/tlʒ/	‘ice’	təllaʒa	‘fridge’
b.	/$s_1d_2d_3$/	‘to close’	səddada	‘cover’
	/$ħ_1k_2k_3$/	‘to scratch’	ħəkkaka	‘mincer’
	/$ṛ_1ʃ_2ʃ_3$/	‘to spray’	ṛəʃʃaʃa	‘sprayer’

The point here is that the roots in (9b) are treated in the same way as the roots in (9a), in that the second segment in each category is doubled regardless of whether it is a member of a consonant cluster or ‘part’ of a geminate consonant. What one would expect if the geminates in (9b) were single long consonants is that the whole segment would be doubled and not just some part of it. If that were the case, we would have the following structures instead:

(10)

/s_1dd_2/	‘to close’	*sə**dd**a**dd**a
/$ħ_1kk_2$/	‘to scratch’	*ħə**kk**a**kk**a
/$ṛ_1ʃʃ_2$/	‘to spray’	*ṛə**ʃʃ**a**ʃʃ**a

To extend the range of situations where geminates act as a sequence of consonants in MA, we refer to cases where they get split by vowels in the exact same way as other consonant clusters do. This happens in the case of deriving the past participle from roots ending with geminates. For concrete examples, see the data below:

(11)

	Root	Past Participle	
	Root	*Past Participle*	
a.	/ʕḍḍ/	məʕ.ḍ**u**ḍ	‘bite’
	/ʃqq/	məʃ.q**u**q	‘crack’
	/ʃdd/	məʃ.d**u**d	‘catch’
	/ħdd/	məħ.d**u**d	‘limit’
b.	/ʃṛb/	məʃ.ṛ**u**b	‘drink’
	/ʕrḍ/	məʕ.r**u**ḍ	‘invite’
	/ħsd/	məħ.s**u**d	‘envy’
	/ʕṭb/	məʕ.ṭ**u**b	‘injure’

In (11), we notice that the derivation of the past participle form involves the prefixation of /m-/ to the root and the infixation of /-u-/ between the last two radical segments. The infixed vowel splits consonant clusters forming syllables of the shape CVC. Both root categories in (11) behave similarly with regards to this infixed vowel, that is as sequences of independent consonants whose two last members are breakable by vocalic material.

The patterning of geminates as sequences of consonants recurs across a variety of other languages. One such a language is Tamazight, for which Saib (1977) suggests a sequential analysis of its geminates. Tamazight has two rules of shwa epenthesis. One rule inserts a schwa before two consonants followed by a vowel. Another rule places a schwa between the last two consonants of tri-segmental sequences. These are reproduced below for easy reference:

(12) Schwa epenthesis in Tamazight

a. $\varnothing \longrightarrow$ ə / — CCV

b. $\varnothing \longrightarrow$ ə / CC — C #

These two rules treat geminates and consonant clusters equally, in that they apply in the presence of each. The first rule applies unrestrictedly before a geminate followed by a vowel and the second one consistently interprets a sequence of a geminate plus a singleton as a tri-consonantal cluster. The data below, from Saib (1977), demonstrates how this works:

(13)

a.	/xdm/	əxdəm	‘work’	
	/ṛʒm/	əṛʒəm	‘open’	
b.	/kks/	əkkəs	‘take off’	
	/qqn/	əqqən	‘close’	
c.	/fa/	fa	‘yawn’	*əfa
	/su/	su	‘drink’	*əsu

First, the rule in (12b) applies consistently to the roots in (13a) and (15b) by inserting a schwa before the last consonant. This creates an environment that triggers the application of the rule in (12a), which places a schwa before a sequence of two consonants followed by a vowel. These rules could not apply if geminates were single consonants since rule (12b) does not apply to a sequence of two consonants, while rule (12a) never applies before one consonant followed by a vowel as can be proved by considering the items in (13c).[8]

Similar properties are found in Leti, treated in Hume et al. (1997), where geminates and consonant clusters alike pattern together in connection with a number of rules, in opposition to singletons. Leti has a word final VC metathesis rule which applies before two-consonant clusters and geminates while it is blocked before single consonants. Also, in Leti, a word final /a/ is deleted only when it precedes a short consonant, but not when a consonant cluster or a geminate follows. These two rules are illustrated in (14a) and (14b), respectively.

(14)

a. Word final VC metathesis in Leti

/kunis+vnutan/	[kunsi**vn**utan]	'iron key'
/vuar+spou/	[vu:ra**sp**ou]	'schooner mountain'
/ukar+ppalu/	[ukra**pp**alu]	'index finger'
/maun+ppuna/	[ma:nu**pp**una]	'bird's nest'

b. Word final /a/ deletion in Leti

/samɛla+nura/	[samɛl**n**ura]	'tricoloured squirrel'
/samɛla+ttɛnan/	[samɛla**tt**ɛnan]	'mouse's spine'
/samɛla+tpunan/	[samɛla**tp**unan]	'mouse's throat'

[8] Other aspects of Tamazight phonology support a featural analysis of geminates. For instance, in the derivation of the intensive form of verbs, underlying geminates fail to further geminate, contradicting the sequential quality they display elsewhere. To solve this problem, Saib (1977) promotes the sequential analysis by suggesting a phonetically based universal convention whereby the two elements of a geminate are redundantly specified as [+long] to fortify each other. This way the structure description of each member of the geminate cluster would be different from that of normal consonants.

In addition to this, the distribution of geminates and consonant clusters in Leti supports their patterning since they are unified by the fact of being limited to word-initial position, for example **pp**ikan 'plate' vs. **pt**una 'star'.

Further distributional evidence for the patterning of geminates and consonant clusters is provided by Leben (1980), based on data from Hausa. Leben points out to the fact that the restrictions that govern the distribution of CC clusters are identical to the restrictions that regulate the distribution of geminates. These restrictions are borrowed here from Leben (1980):

(15) The distribution of geminates and CC clusters in Hausa

Geminates	*CC clusters*	
*GV…	*CCV…	
*…VG	…*VCC	
*…GC… and *…CG…	…*CCC…	
…VGV…	…VCCV…	(possible position)

The idea here is that the distributional parallel between geminates and clusters in Leti and Hausa reveals that they are structurally equivalent. Otherwise, the only other possible explanation is that their identical distribution is accidental. This issue will be the matter of further discussion. In the following section, we introduce the properties that geminate consonants share with unit structures.

2.2. Geminates as a unit structure

Considering all the above-mentioned commonalities between geminates and consonant clusters, it seems very convincing that they should be treated the same. That is, geminates ought to be interpreted as a cluster of two identical consonants. Nonetheless, the picture is not quite complete yet since the two structures were reported to diverge from each other in many ways. More specifically, geminates were shown to resist and block rules of epenthesis, metathesis, syncope and spirantization that normal consonant clusters undergo or trigger (Kenstowicz and Pyle, 1973; Guerssel, 1977, 1978; Schein and Steriade, 1986; Hayes, 1986). This property of geminates is commonly referred to as geminate integrity and inalterability.

This characteristic of geminate consonants was first made known by the seminal works of Kenstowicz and Pyle (1973) and Guerssel (1977). Yet, these treatments do not claim that geminates should be regarded as unit segments as a result. In fact, they advance a sequential analysis that regards geminates as a cluster of two alike consonants. In defense of this claim, they argue that geminates are just a special type of consonant clusters that behaves differently.

Kenstowicz and Pyle (1973) starts by emphasizing the sequential make-up of geminates drawing on data from Sierra Miwok. The language is reported to have a shortening rule that shortens long vowels before a geminate or a consonant cluster, stated in (16). Also, there is a stress rule which places stress on the first stem vowel if it is long or if it precedes a geminate or a consonant cluster.

(16) $V{:} \longrightarrow [\text{-long}] \;/\; \_\_\, CC$

Sierra Miwok has another rule whose effect geminates the word-final singleton of the past form:

(17) $C_i \longrightarrow C_iC_i \,/\, \text{—}\,]_{past}$

The third rule is a metathesis rule that swaps the order of any final vowel and the consonant that precedes it in the past form:

(18) $C_1C_2V]_{past} \longrightarrow C_1VC_2$

The data to which these rules apply is partially reproduced below where stress is placed above the relevant vowels:

(19)

	Present	*Past*		*Gloss*
Type I	patí:t	patítt		'to take'
	sakú:k	sakúkk		'to tear'
Type II	cé**lk**u	celúkk		'to quit'
	wí**mk**i	wimíkk		'to spear'
Type III	há**mm**e	hamé??	*hamemm	'to bury'
	mí**ll**i	milí??	*milill	'to sing'

The gemination rule applies first to type I stems turning their final consonants into geminates in the past form. This creates an environment for the shortening rule, hence the relevant long vowels get short. The metathesis rule, on the other hand, applies to type II stems which creates an environment for the gemination rule. However, metathesis does not apply to type III stems although they contain what is believed to be a consonant cluster based on their behavior with regards to the shortening and the stress rules. Yet, geminates appear to diverge from normal consonant clusters by blocking metathesis. Instead, type III stems undergo a degemination rule of their middle geminates followed by a glottal stop insertion stem-finally, which then gets geminated by rule (17).

Kenstowicz and Pyle proceed to show that this is not a problem for Sierra Miwok alone by citing a whole range of other languages, including Kolami, Kasem and Tunisian Arabic, where geminates diverge from clusters of unalike consonants. On this basis, the authors propose the Integrity Hypothesis as a universal condition on rules:

(20) The Integrity Hypothesis (Kenstowicz and Pyle, 1973)

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.

As it stands, this condition prevents rules from applying to geminates even though the latter meet the structural description of those rules. Thus, there is no need for any ad hoc restrictions on the application of individual rules.

In much the same spirit, Guerssel (1977) supports the sequential analysis of geminates relying on data from Amazigh, but still points out to the fact that geminates do behave differently from clusters of unalike consonants. One piece of evidence that Guerssel provides in favor of the geminate-cluster patterning has to do with a rule of syncope in Amazigh which deletes a schwa before CV (e.g. /ʒər-i/- [ʒri] 'throw me down'). Yet, the syncope rule gets blocked when followed by a consonant cluster or a geminate (e.g. /bəddəl/ - *[bddəl] 'to change').

Nevertheless, according to Guerssel (1977), geminates are not consistent in their patterning with consonant clusters. This is evidenced by a process of metathesis that affects normal clusters exclusively (e.g. /uʒl+əx/ - [uʒəlx] 'I hang'). As for geminates, the process is blocked (e.g. /sərrəg/ - *[sərərg]. To account for this observed behavior, Guerssel (1977) suggests a constraint on the application of phonological rules dubbed the Adjacency-identity Constraint. This constraint stipulates the following:

(21) The Adjacency-Identity Constraint (Guerssel, 1977)

Given two segments A1A2 where A1=A2, a phonological rule can alter the adjacency of A1A2 if and only if it alters the identity of A1 or A2.

The proposed constraint prevents rules of metathesis and epenthesis from splitting geminates unless the featural identity of one is altered. Again, this is devised to avoid any ad hoc restrictions on rules while justifying the special status of geminate clusters.

In this light, geminates in MA appear to obey the stipulation of both the Integrity Hypothesis and the Adjacency-Identity Constraint. More specifically, geminates in MA behave as unbroken segments, in that they resist schwa epenthesis that other normal consonant clusters undergo. In the normal case, a schwa occurs between the last two consonants of every tri-consonantal verb or adjective (e.g. CCəC), unless they end in a geminate, consider (22a) and (22b), for example. In the presence of a geminate, the schwa appears between the first consonant and the geminate of these verbs or adjectives (e.g. $CəC_iC_i$). Also, when a noun ends in a full vowel, see (22c), a schwa is inserted between the two initial consonants, except when it is a geminate.

(22) Geminate integrity and schwa epenthesis

a. Verbs:

(i)	sədd	'close'	*sdəd
	ħəll	'open'	*ħləl
	ʃədd	'catch'	*ʃdəd
	dəqq	'knock'	*dqəq
(ii)	ɦṛəb	'escape'	
	tləf	'disappear'	
	ʃṛəb	'drink'	
	kbər	'grow'	

b. Comparatives:

(i)	rəqq	'thin'	*rqəq
	xəff	'light'	*xfəf
	bənn	'tasty	*bnən
	ṣəħħ	'strong'	*ṣħəħ

(ii)	tqəl	'heavy'	
	ṭwəl	'tall'	
	ɣləḍ	'fat'	
	sɦəl	'easy'	

c. Nouns:

(i)	ssma	'the sky'	*səsma
	ḍḍṛa	'the corn'	*ḍəḍṛa
	ṣṣḍa	'the rust'	*ṣəṣḍa

(ii)	kətra	'plenty'	
	ħəʃma	'shyness'	
	ləkra	'the rent'	

On this ground, we could claim that the rule responsible for schwa epenthesis in MA is subject to the Integrity Hypothesis and the Adjacency-Identity Constraint. This blocks the rule from applying to geminates while it still applies to other consonant clusters. However, the fact that geminates cross-linguistically diverge from consonant clusters could indicate a difference in their structure. This view will be considered for further discussion later in this chapter.

To recapitulate, the behavior of geminate consonants has been shown to be marked by indeterminacy, in the sense that geminates exhibit properties of both singletons and consonant clusters. On the one hand, they act as sequences of two-like consonants that have the exact same distributional restrictions and the structural effects as sequences of unlike consonants. On the other hand, rules of epenthesis and metathesis, which typically split normal consonant clusters, are blocked from affecting them, in much the same way they cannot affect singletons. In what follows, we will be discussing the various phonological representations that were proposed to explain the special behavior of geminates.

3. The phonological representation of geminates

3.1. The linear approach to geminate representation

In the framework of linear phonology (Chomsky and Halle, 1968), phonological length was accounted for in terms of the binary feature [-/+long]. Under this view, geminates could be perceived either as single unit segments that are [+long] (i.e. featural analysis) or as a sequence of two-like consonants, C_yC_y (i.e. sequential analysis).

(23) Geminate representations in linear phonology

C	C_y	C_y
[+long]	[-long]	[-long]

The featural analysis of geminates can adequately explain the integrity that geminates display with regards to rules of epenthesis and metathesis. That is, if geminates are monolithic segments that are long, no phonological material can penetrate their featural make-up. However, the featural analysis fails to justify the structural parallel between geminates and other consonant clusters. If we choose to hold to a featural analysis, the distributional parallel between geminates and consonant clusters should be rendered accidental. Besides, the rules triggered by consonant clusters would need to be modified for their environment. Instead of only applying in the context of consonant clusters, they should apply in the presence of long consonants as well. This means that long consonants and consonant clusters will be made equivalent. According to Saib (1977), this "makes a simple rule complex by stating two environments instead of one".

The sequential analysis does a good job as far as capturing the observed parallel between geminates and normal clusters. Yet, it falls short of explanatory adequacy when it comes to explaining the ability of geminates to block rules that are otherwise applicable to other clusters. The works of Kenstowicz and Pyle (1973) and Guerssel (1977) fall under the purview of linear phonology. As mentioned before, they both argue for a sequential analysis of geminates. Such an analysis can account for geminate blockage only by resorting to additional stipulations on phonological rules. Here comes the role of the Integrity Hypothesis and the Adjacency-Identity Constraint. There are two problems related to such stipulations. First, they are stated in the form of universal inviolable constraints. This is seen as a problem because there are many exceptions that contradict these constraints. For example, MA is laden with productive processes that split geminates, the past participle as a case in point. Second, they indicate that the representational mechanism of the theory is insufficient.

As far as the MA data presented so far is concerned, neither of the above-mentioned linear analyses seem to adequately account for the observed patterns. The behavior of geminates in MA is quite complex. The fact that geminates in MA split on some occasions but hold their integrity in others is troublesome for both the featural analysis and the sequential one. If a geminate is a segment specified for the feature [+long], then the rule of schwa epenthesis interprets it as an indivisible unit. That is why the geminates in (23) above are not split by schwa.

In this case, we would be asking: why do geminates still count as two segments in triggering schwa epenthesis? Another question would be: why do the same geminates split elsewhere?

(24) Inconsistent geminate integrity in MA

Root	*Verb*	*Past Participle*	*Gloss*
/ʕḍḍ/	ʕə**ḍḍ**	məʕ**ḍ**u**ḍ**	'bite'
/ʃqq/	ʃə**qq**	məʃ**q**u**q**	'crack'
/ʃdd/	ʃədd	məʃdud	'catch'
/ħdd/	ħə**dd**	məħ**d**u**d**	'limit'

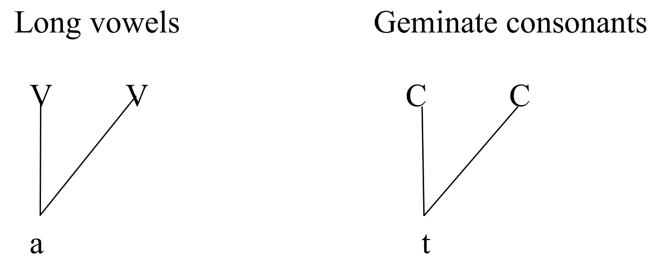
This data set is problematic not only for the featural analysis but also for the sequential one. The latter would make use of some constraint (i.e. Integrity Hypothesis or the Adjacency-Identity Constraint) on rules to account for geminate integrity. However, it is obvious that such a constraint would not be obeyed by all the rules of MA where geminates are commonly split (e.g. the past participle). Thus, it is clear that a linear analysis is struggling to explain the dual function of geminates.

3.2. The non-linear approach to geminate representation

3.2.1. The autosegmental approach

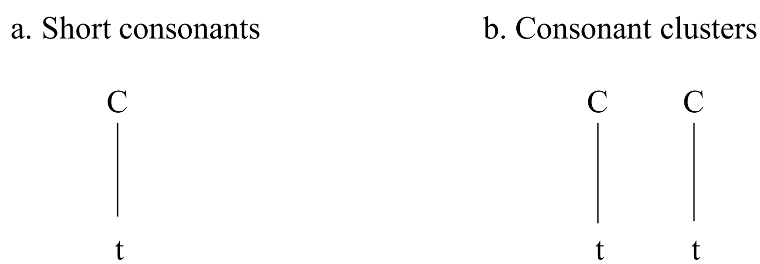
The advent of autosegmental phonology (Leben, 1973; Goldsmith, 1976; McCarthy, 1979) has opened new avenues for phonological representations. One of the major contributions of this framework is the autosegmentalization of tonal and melodic features, separating them from their linearly-ordered segmental positions. These components are then placed on distinct interacting tiers, namely the melodic tier and the timing (or skeletal) tier, allowing for multiple associations between these two levels of representation. In this theory, segmental length is represented as one melodic element multiply-linked to two segmental positions.

(25) Geminate representation in autosegmental phonology



This mechanism allows geminates to be represented differently from both single short consonants and consonant sequences:

(26) Singletons and consonant clusters in autosegmental phonology



The double association structure of long segments is governed by two independently motivated principles of autosegmental phonology: (1) the Obligatory Contour Principle (Leben, 1973; McCarthy, 1986) and (2) autosegmental spreading (Leben,1973; Goldsmith, 1976). On the one hand, the OCP prohibits the occurrence of adjacent identical melodies morpheme-internally.

(27) Obligatory Contour Principle (OCP; McCarthy, 1986:208)

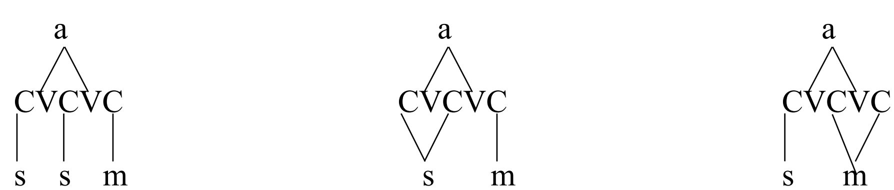
At the melodic level, adjacent identical elements are prohibited

Autosegmental spreading, on the other hand, associates melodic units with their corresponding segmental positions, giving rise to the structure in (25). However, the OCP does not restrict heteromorphemic identical adjacencies, permitting the structure in (26b).

In Arabic, lexical representations are argued to be governed by the OCP. The language is characterized by the absence of patterns of the shape $C_xVC_xVC_y$, where the two first consonants are identical, whereas patterns of the shape $C_yVC_xVC_x$ abound. According to McCarthy (1979, 1981, 1986), this right-left asymmetry is subject to the OCP effect and autosegmental spreading, which is rightward in Arabic. The following example is borrowed from McCarthy (1986) for easy reference:

(28)

a. *sasam b. *sasam c. samam



This shows that the structures in (28a) and (28b) are non-occurring due to their violation of the OCP and the rightward spreading rule, respectively. As a result, the only way long tautomorphemic segments can occur in Arabic is via the double linking of melodic segments to the available corresponding timing slots.

Armed with these representational tools, we can now return to our data. We will try to see how an autosegmental approach to geminates accounts for the observed geminate patterns in MA. To start with, it will be argued that, like in Classical Arabic, lexical geminates in MA follow from the restriction imposed by the OCP on melodic elements. This means that underlying geminates in MA are the result of rightward spreading of a consonantal root node, initially associated with a segmental position in the skeleton, to another slot in the same skeleton, as the following examples show:

(29) MA underlying geminates in autosegmental phonology

a. Final geminates	b. Medial geminates	c. Initial geminates
	a	a
CC C	CC CV	C CV
s d	s n	d
[sədd]	[sənna]	[dda]

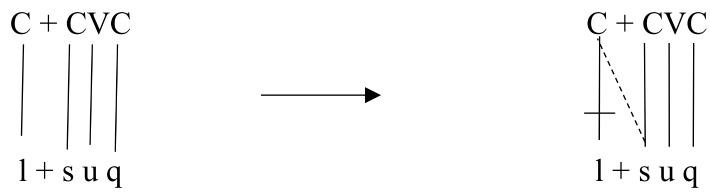
Word templates do not distinguish between consonant clusters and geminates. The latter arise when the number of melodic elements provided by the lexicon is less than the available slots in the template. The language then resorts to spreading the rightmost element to cover the remaining slots, resulting in multiple associations. The interesting point here is that even when lexical geminates in MA occur word-initially, they only do so in the case of mono-consonantal lexical items, as in (29c). Hence, being the rightmost melodic element and in the absence of another consonantal melody competing for the relevant slot, the designated melody can spread rightward. This provides us with a structural parallel between geminates in all word positions.

Next, we show how derived geminates in MA are taken care of under the purview of the autosegmental framework. In the previous chapter, we discussed in detail the derivation of geminates through total assimilation, assuming an autosegmental representation of geminates. In this chapter, we put the analysis in its context by highlighting the structural parallel between all categories of geminates in MA. The remarkable aspect of a theory of autosegmentalization is that every geminate structure is the outcome of the spreading mechanism.

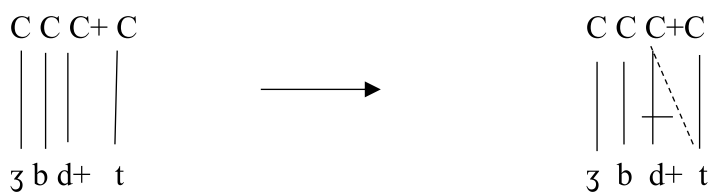
The default spreading direction in MA is rightward, but specific rules can make melodic elements spread in the other direction. In the case of total assimilation and compensatory lengthening, spreading goes from right to left. As far as the assimilation data is concerned, it has been shown that the direction of the spreading follows from the interaction of constraints á la OT. In the case of geminates created by compensatory lengthening, spreading is also leftward. No definite analysis of this process is provided yet as we wait to make our assumptions about the representation of geminates as clear as possible.

(30) MA phonologically-derived geminates in autosegmental phonology

a. Total assimilation

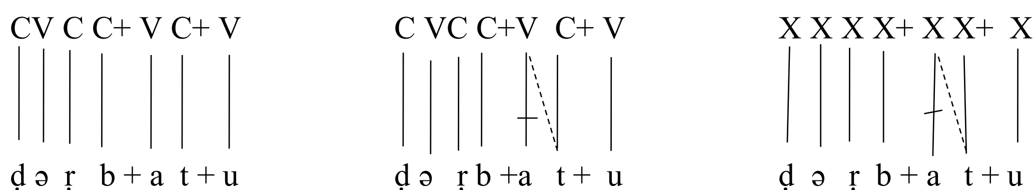


/l+suq/ [ssuq] ‘the market’



/ʒbəd+t/ [ʒbətt] ‘I pulled out’

b. Compensatory lengthening



[ḍəṛbatu] [ḍəṛbəttu] ‘she hit him’

In the theory of autosegmental phonology, deletion takes place when a melodic element dissociates from its position on the timing tier. This leaves behind an empty slot. Different languages resort to spreading of nearby melodies to occupy the lingering slot. If what spreads is a vocalic melody, long vowels are created. If what spreads is a consonantal melody, however, a geminate is derived. The case of gemination illustrated in (30b) above is depicted as a case of

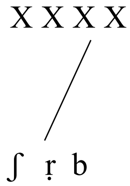
CL, whereby the delinking of the vowel /a/ of the past tense feminine pronoun /-at/ is believed to trigger the spreading of the consonant /t/ of the same pronoun. The problem that we face in (30b) is that a consonantal melody spreads to a V-position. For this reason, we replace the CV tier by the X-slot tier (Levin, 1985). This allows skeletal positions to freely associate with consonantal and vocalic melodies.

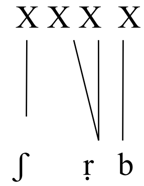
As we mentioned before, morphological gemination will be dealt with in a chapter in its own (chapter 8). As for now, we want to show how the theory of autosegmental phonology representationally handles morphologically-derived geminates. To do so, we will briefly review the work of Bennis and Iazzi (1995), whose analysis of morphological causatives in MA provides some insights into this matter. The basic assumption of this work is that the causative template has the shape of XXXX, where X is a position on the skeletal tier. The account rests on the idea that different morphological causatives are derived by associating different lexical items to the designated template. For instance, when the melodies of the root /ʃṛb/ are linked to the X-slots of the causative template, we get the causative verb [ʃəṛṛəb]. This analysis is faced with a challenge, however. The spreading direction of MA does not yield the expected form. Only the most rightward melody can spread.

(31) MA morphologically-derived geminates in autosegmental phonology

a.	*Causative template*	*Lexical entry*	*Associations*	*Output*
	XXXX	/ʃṛb/	X X X X	*[ʃṛəbb]
			ʃ ṛ b	

b. *Pre-association* *Association* *Output*



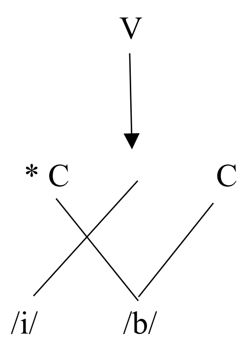


[ʃəṛṛəb]

To overcome this problem, the authors postulated a rule of pre-association, whereby the second root consonant is pre-linked to the third slot in the template (31b). Subsequently, normal association lines are added. This way, expected forms are derived.

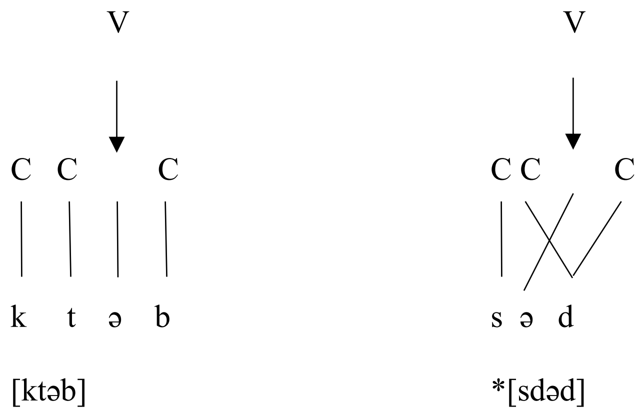
Now the question is: how does the theory of autosegmental phonology account for geminate integrity and geminate splitting in MA. To answer this question, we refer to the non-crossing association lines constraint of autosegmental phonology. According to this constraint, the application of an epenthesis rule is blocked if it would create an association line that crosses an existing association line:

(32) The non-crossing association line constraint



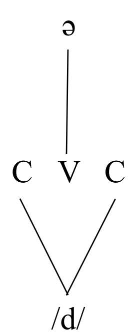
By the same token, we can explain why schwa in MA splits normal consonant clusters while it fails to break up geminates. The following examples illustrate how it works:

(33) Schwa epenthesis in autosegmental phonology



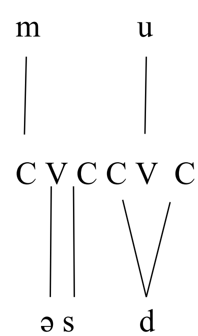
However, some nagging problems still remain. Most striking is that fact the melodic elements in autosegmental phonology can exist on different tiers without crossing each other's association lines. This means that we could imagine a schwa on its own independent tier associated to the inserted V-position without interrupting the association lines of the geminate:

(34)



It appears that only morphologically relevant melodies are capable of having their own tiers. Since schwa in MA is purely phonetic, it cannot be placed on an independent tier. This takes us to the case where geminates get split by full vowels in MA, especially in the case of the past participle (e.g. [sədd] – [məsdud] ‘close’). Our stand on this case of geminate splitting relies on the assumption we have just advanced, that morphemes have independent tiers. As a result, the vowel /u/ of the past participle can link to its V-position on the skeletal tier without violating the non-crossing association lines constraint.

(35)



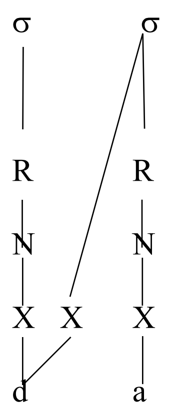
[məsdud]

As a consequence, the autosegmental representation of geminates proves quite successful in capturing the dual patterning of geminates. It does so by utilizing independently motivated principles to represent geminates in a consistent way and characterize their behavior without resorting to additional stipulations. Under this theory, all patterns of geminates, be they derived or underlying, emerge from the regulations of OCP, autosegmental spreading and non-crossing association lines constraint. The interaction between these rules produces a single melodic unit doubly associated to two positions on the timing tier. This way, rules that effect the featural specifications of segments would interpret geminates as single consonants while those that refer to structural information would construe geminates as clusters (Leben, 1980; Schein and Steriade, 1986; Hayes, 1986).

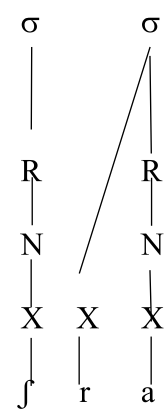
The predictions that an autosegmental approach to geminates make about syllable structure are also vital to our discussion. Being represented as two C-slots on the skeletal tier, geminates are expected to be distributed and syllabified in the same way as CC clusters (Curtis, 2003). That is, if a language has CCV and CVCC word shapes, nothing should prevent GV and CVG shapes from existing, and vice versa. Also, if a language counts a syllable with a branching nucleus or a branching rime (i.e. a syllable with one or more X-slots after the nucleus) as heavy, then CVV, CVC, CVCC and CVG should be equally heavy. However, if only branching nuclei are weight-sensitive, then CVC, CVCC and CVG should be all light. The point is that syllables closed with geminates and those closed with singletons should always have the same weight configurations.

Interestingly, these predictions are borne out by the data of MA. As we noted before, geminates share the same distributional properties of CC clusters. Therefore, like consonant clusters, geminates in MA can appear both intervocalically and at the margins (e.g. GV and CVG). A skeletal approach attributes this commonality of distribution to the structural parallel between the two. As far as their syllabification is concerned, geminates and clusters are expected to display the same syllabic behavior. Since branching rimes are weight-sensitive in MA, CVG and CVC syllables would equally count as heavy. In addition, since complex margins are not tolerated in MA, geminates and clusters would never be tautosyllabic. Instead, a syllabic consonant is created out of their constituents to satisfy this requirement.

(36) Initial geminates/clusters

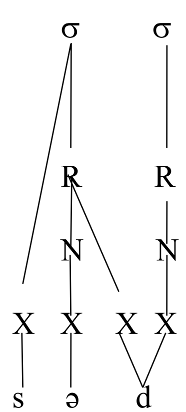


[dda]

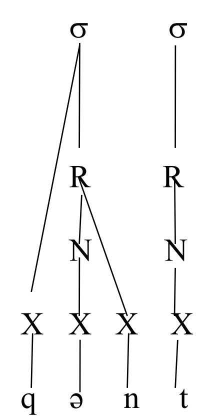


[ʃra]

(37) Final geminates/clusters



[sədd]



[qənt]

However, autosegmental theory falls short of explanatory power elsewhere. For example, in MA, content words of the shape CəC and syllables of the shape Cə are not possible, but CVC words and CV syllables are. In a skeletal approach, CəC and CVC words are equivalent, and so are Cə and CV syllables. This is because such a theory relies on the counting of skeletal positions. Therefore, both CəC and CVC have three positions, while both Cə and CV have two.

(38) The inadequacy of the skeleton:

$$\begin{array}{ccc} X & X & X \\ | & | & | \\ C & \text{ə} & C \end{array} \qquad\qquad \begin{array}{ccc} X & X & X \\ | & | & | \\ C & V & C \end{array}$$

$$\begin{array}{cc} X & X \\ | & | \\ C & \text{ə} \end{array} \qquad\qquad \begin{array}{cc} X & X \\ | & | \\ C & V \end{array}$$

This calls for the consideration of a better prosodic theory. Such a theory should be able to explain the difference between CəC/ Cə, on the one hand, and CVC/ CV, on the other. Moraic theory explains this asymmetry by claiming that schwa is nonmoraic in MA (Bensoukas and Boudlal, 2012a-b).

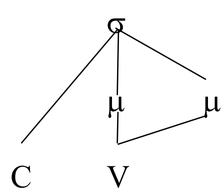
3.2.2. The moraic approach

Moraic phonology (Hyman, 1985; McCarthy and Prince, 1986; Hayes, 1989) abstracts away from skeletal prosodic representations by associating melodic elements with moras rather than CV-slots or X-slots. In this theory, the mora is an abstract prosodic unit for measuring syllable weight and representing segmental length. This development in the theory of phonological representation came as a response to the limitations posed by a theory of skeletal-tiers, which have to do mainly with typological observations about quantity-sensitive phenomena such as stress, word-minimality and Compensatory Lengthening.

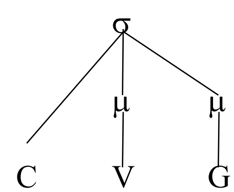
In languages with weight-sensitive stress systems, primary stress falls on heavy syllables. In this regard, an open syllable with a long vowel and a closed syllable with a geminate both attract stress in languages that have them. Hence, they pattern as heavy by virtue of being bimoraic.

(39)

a. CVV



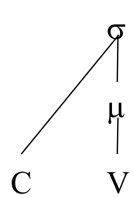
b. CVG



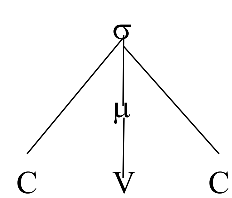
On the other hand, an open syllable with a short vowel or one closed with a singleton count as light since they comprise only one mora.

(40)

a. CV

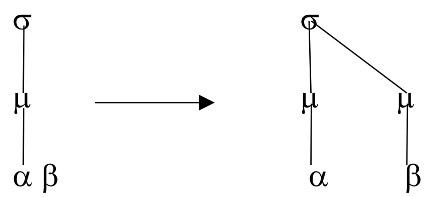


b. CVC



This typology holds true for languages where Weight-by-Position does not apply. In case it does, a CVC syllable would be assigned an additional mora to its coda, making it equivalent to CV: and CVG syllables.

(41) Weight by Position (Hayes, 1989: 258)



(where σ dominates only μ)

Accordingly, moraic theory makes the distinction between CVG and CVC syllables formally possible. In skeletal theory, both syllable shapes are predicted to be quantitatively equivalent under all circumstances. That is, they should both pattern either as heavy in languages with branching rimes or light in languages with no branching rimes. This fact, among many others, justifies the need for the mora and proves the inadequacy of the skeleton. More specifically, it shows that there is some unit (i.e. mora) above the segment that accounts for syllable weight.

On the basis of these observations, long vowels and geminates appear to consistently contribute additional weight to their syllables. For this reason, they are conceived of as being underlyingly moraic. In particular, a geminate is represented as underlyingly associated to one mora. Short consonants, however, are underlyingly nonmoraic. Their moraicity is subject to the language-particular application of Weight-by-Position.

(42) Moraic representation of segmental length

Geminate	Singleton
μ	
|	
C	C

As for vowels, a long vowel is represented as underlyingly linked to two moras while a short one projects only one mora.

(43)

Long vowel	short vowel
μ μ	μ
| /	|
V	V

In this sense, moraic theory characterizes segmental length and syllable weight by means of the same prosodic principles. In other words, length and weight are both expressed by reference to moras.

The predictions made by moraic theory about the underlying representation of geminates is corroborated by a plethora of weight-sensitive phenomena. Davis (1994) discusses the case of Hindi where stress is weight-sensitive. In Hindi, CVG syllables, but not CVC ones, pattern with CVV syllables as heavy. Likewise, in the Uto-Aztecan language Cahuilla (Hayes, 1995) and San'ani Arabic (Watson, 2002), CVG and CVV syllables consistently behave as bimoraic with regards to stress assignment.

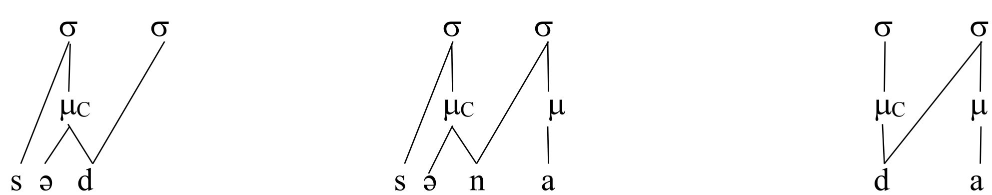
Further, while singletons in Hindi can follow long vowels (i.e. CVVC), geminates cannot (i.e. *CVVG) since the language does not permit trimoraic syllables. Similarly, Davis (2011) mentions the case of Swedish (Kiparsky, 2008) and Koya (Tyler, 1969) which have a vowel shortening process that applies before geminates, but not before singletons. This too is motivated by the avoidance of trimoraic syllables in Swedish and Koya alike.

Another piece of evidence reported by Davis (1999) comes from the prosodic morphology of Hausa and Sinhala where it is believed that, unlike CVC syllables, CVG syllables are treated as bimoraic, on a par with CVV syllables. In Hausa, class 3 plural nouns are derived via the suffixation of one of the following two allomorphs: -aaCee or -aayee. The first allomorph attaches to monomoraic nominal roots (e.g. dam-aamee 'monitor'). The second allomorph attaches to bimoraic nominal roots (e.g. zoom-aayee 'hare' and tukk-aayee 'bird crest'). Comparably, in Sinhala, there are two allomorphs for the genitive suffix, [ee] and [e]. The former is selected by monomoraic roots (e.g. mal-ee 'the flower's'), whereas the latter is selected by bimoraic roots (e.g. paar-e 'the street's' and pott-e 'the core's). This fact illustrates that CVG roots and CVV ones are equivalently bimoraic in Hausa and Sinhala.

In view of this compelling evidence, we proceed to represent the geminates of MA as underlyingly moraic. Under this conception, geminates are expected to inherently contribute prosodic weight to the words they belong to. This representational shift would change our perspective of explaining the observed geminate patterns in many ways. In fact, as appealing as it might seem to adopt a moraic approach to geminates, there are many challenges facing it as far as the geminates of MA are concerned. We decide to endure these challenges and make the moraic approach work for our data. This decision is spurred by our conviction that a unified representation of geminates is more desired in phonological theory.

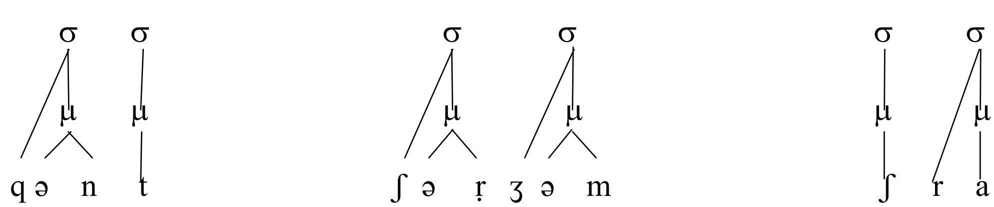
Under this moraic conception, the identical distributional properties between geminates and consonant clusters would be rendered accidental, instead of being motivated by their structural parallel as predicted by the skeletal approach. In practice, geminates in moraic theory are construed as unit (i.e. monolithic) segments. In line with this, the syllabic parallel between geminates and clusters becomes an inadvertent consequence of the fact that both geminates and consonant clusters are hetrosyllabic units, but each for its own special nature. While the members of a consonant cluster would independently belong to different syllables, a geminate would be linked to two syllabic nodes.

(44) The syllabification of geminates in moraic theory



[sədd] [sənna] [dda]

(45) The syllabification of consonant clusters in moraic theory



[qənt] [ʃəṛʒəm] [ʃra]

To elaborate, two major differences should be noted between the syllable structure of the items in (44) and those in (45). First, the geminates in (44) are single segments that are simultaneously linked to two syllabic positions, while the consonant clusters in (45) are two independent segments, each belonging to a different syllable. Second, the geminates in (44) bring their moras with them, whereas normal consonants are assigned moras depending on their position and function in the syllable (for example, as codas or syllabic consonants).

Additional evidence for the moraic structure can be obtained from the facts of Compensatory Lengthening patterns as delineated by Hayes (1989). According to Hayes, a skeletal approach to Compensatory Lengthening overpredicts that every segment deletion by dissociation can be compensated for by spreading the features of a nearby melody. This means that the skeletal theory does not distinguish between onset deletion and coda deletion since both would leave an empty slot. Also, a skeletal theory does not differentiate between the deletion of moraic codas and nonmoraic ones.

However, it is cross-linguistically observed that compensatory lengthening occurs only in the aftermath of coda deletion, in languages with W-b-P, or vowel deletion for that matter. Moraic theory captures this generalization very well, in that it assumes that only the deletion of moraic elements triggers compensatory lengthening. In this capacity, compensatory lengthening becomes motivated by the conservation of mora count.

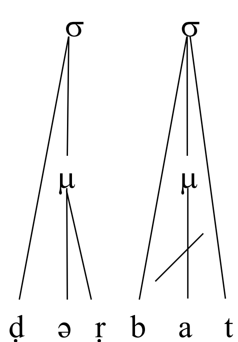
(46) Moraic conservation (Hayes, 1989:285)

Compensatory Lengthening processes conserve mora count

Since moraic theory assumes that onsets do not add to the mora count of syllables, it becomes evident why CL does not take place upon the deletion of onset segments or any other nonmoraic element.

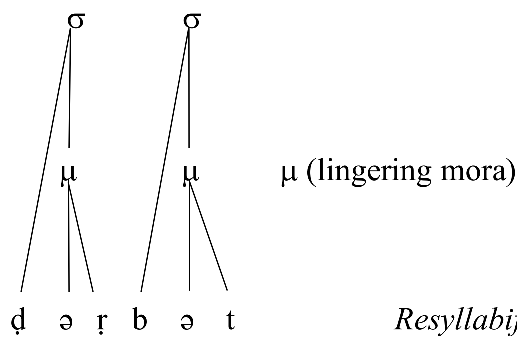
Under the conception of CL, the gemination characterizing forms such as [ḍəṛbəttu] can be interpreted along the same line, whereby geminating the consonant /t/ can be argued to be triggered by the deletion of the vowel /a/ in the affix /at/. This seems to be consistent with the thinking of moraic theory since a vowel is underlyingly associated with a mora. Thus, it could be argued that the deleted vowel leaves behind it a lingering mora. In order to preserve the mora count of the word, the onset of the ensuing syllable links to the empty mora, giving rise to a word internal geminate structure. A schwa is then inserted to replace the deleted vowel as the nucleus of the syllable. The details of this process are illustrated by the representations in (47) below. An elaborate OT analysis of this case of gemination will be provided in chapter 4.

(47) Compensatory lengthening in MA under the moraic approach



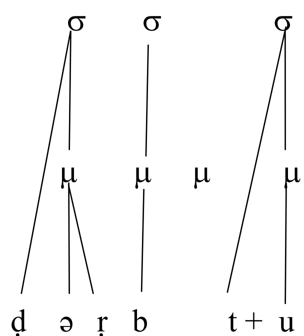
Deletion

[ḍəṛbat] ‘she hit’



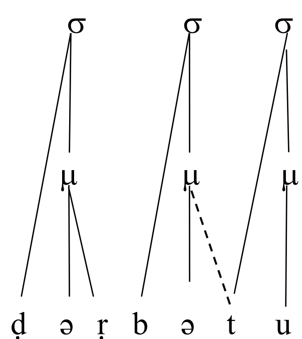
Resyllabification

[ḍəṛbət] ‘she hit’



Affixation and resyllabification

/ḍəṛ.b.tu/ ‘she hit+ him’



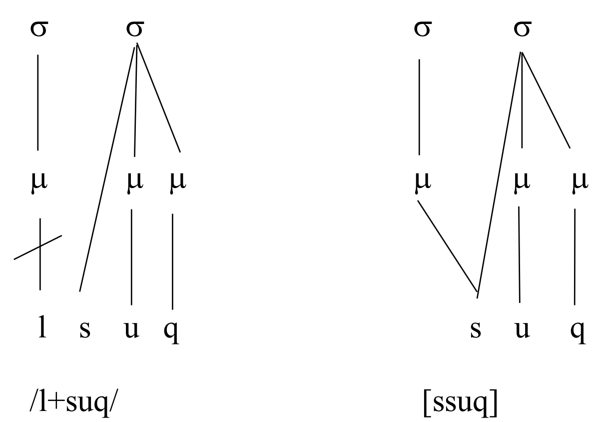
Compensatory lengthening and resyllabification

[ḍəṛbə**tt**u] ‘she hit him’

Unfortunately, MA does not contain other cases of deletion that could make the asymmetrical nature of CL clearer. What we have in mind is a case where the deletion of a nonmoraic element, say an onset, does not call for CL, while deleting moraic elements, like vowels and codas, does. Actually, it is very hard to establish this asymmetry in a single language. However, such an asymmetry is cross-linguistically adduced.

Next, we continue to show how the adopted moraic approach handles the geminates derived phonologically by assimilation as well as those derived morphologically by means of word derivation. Segments in moraic phonology are associated directly to moraic or syllabic nodes, in lieu of skeletal positions. With this in hand, we can think of assimilation as spreading some melodic material from one prosodic position to another. In our case of assimilation, spreading goes from the first radical segment associated directly to the syllable node, by virtue of being an onset, to the head mora left by the dissociated melody representing the definite article.

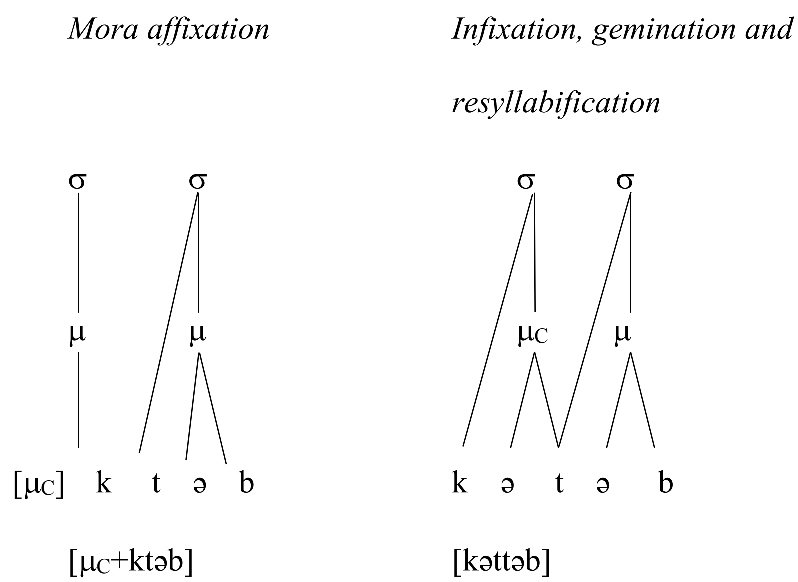
(48) Total assimilation in MA under the moraic approach



A moraic approach to geminates assumes that every geminate, be it lexical or derived, is moraic. To maintain this generalization, it is widely argued that a morphologically derived geminate emerges from the affixation of a mora. When it attaches to a singleton, the mora creates a geminate which is indistinguishable from other lexical or phonological geminates. Except that it has a morpheme status as it adds meaning by creating new lexical items (or word forms like in the case of the imperfective form in Tashlhit (Bensoukas, 2001)).

Along the same lines, the geminates that characterize the derivation of causative verbs, agent nouns and instrument nouns are assumed to be the outcome of affixing a mora to help derive those forms. Here, we will illustrate by referring to the case of causative verbs as analyzed by Noamane (2013/2014). In that work, we accounted for the causative formation in terms of stem modification, by lengthening the second radical element, as mora affixation to a root, following Lombardi and McCarthy (1991), Samek-Lodovici (1992) and Bensoukas (2001). An example is given below:

(49) Morphological gemination in MA under the moraic approach



The essence of this analysis can be extended to the cases of agent and instrument nouns. This will be the topic of the chapter about morphological gemination. There, we will present in detail the advantages a moraic approach provides in dealing with morphological geminates.

The remaining issue now is to understand how the moraic approach plays into the integrity displayed by geminate consonants with regards to schwa epenthesis in MA. It appears that representing geminates as a one root-node associated to a mora explains geminate-integrity straightforwardly since unit segments are inherently indivisible. This way, the integrity of geminates is minimally captured, unlike in the skeletal approach which calls upon the non-crossing association lines constraint to explain such a property.

This indivisible monolithic nature of geminates in moraic theory turns out to be a problem for part of our data. Precisely, it is well attested that geminates in MA can be split by means of inserting some vocalic material between what appears, given this vocalic insertion, to be a geminate cluster. By way of illustration, let's reconsider the case of the past participle mentioned earlier:

(50)

Root	*Verb*	*Past Participle*	*Gloss*
/ʕḍḍ/	ʕə**ḍḍ**	məʕ**ḍuḍ**	'bite'
/ʃqq/	ʃə**qq**	məʃ**quq**	'crack'
/ʃdd/	ʃə**dd**	məʃ**dud**	'catch'
/ħdd/	ħə**dd**	məħ**dud**	'limit'

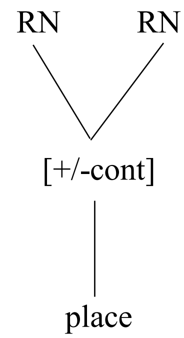
There is no way the vowel /u/ could break the relevant geminates if they were indivisible monolithic segments. The other closely related problem our data can pose to a monolithic conception of geminates has to do with the triggering of schwa epenthesis. As mentioned before, the context for schwa epenthesis should be at least a tri-consonantal sequence. Interestingly, roots that consist of a consonant plus a geminate (i.e. CG) qualify as an environment for schwa epenthesis, hence, contrary to the prediction made by the moraic representation, the relevant geminates count as two segments.

These two issues constitute very serious challenges that could work as strong arguments against the moraic approach of geminates, and thus jeopardize our goal to advance a moraic treatment of geminates in MA. Still, since the mora associated with the geminate is primarily a prosodic unit, these issues could be resolved at the segmental level without sacrificing the good contributions of the moraic approach. This can be done by representing geminates as two-root nodes linked to a mora, in lieu of the standard representation of a moraic one root node. In other words, the problem has to do with the strictly built-in monolithic nature of geminates and not their moraic status. By taking this step, it becomes possible to maintain a degree of structural parallel between geminates and clusters in accordance with the observed facts that our data reveals.

3.2.3. The two-root node theory of length

The two-root node theory of length was proposed by Selkirk (1990) as an alternative to both the skeletal approach and the moraic one. For Selkirk, geminates should be represented with two identical root nodes whose stricture and place features are shared. This representation is schematized below:

(51)



This model benefits from the development brought by the theory of feature geometry (Clements, 1985), whereby features are hierarchically organized. Under this conception, the root node functions as an anchor point for the other intermediate feature nodes. For a feature to be realized in the output form, it should be licensed by some root node. Since geminates can be represented with two root nodes, the relevant features get realized twice in the output.

The phonotactics of MA resonates well with the premise of the two-root node model, in the sense that other sequences of consonants with identical sonority levels are permitted. Particularly, sequences of stop-stop, fricative-fricative and nasal-nasal are easy to be found. Specific examples are given below:

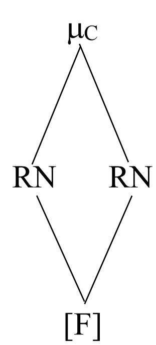
(52) Identical sonority sequences in MA

a. Stop-stop: dbəħ ‘to slaughter’, ktəb ‘to write’
b. Fricative-fricative: ħsəb ‘to cout’, fʃəl ‘to fail’
c. Nasal-nasal: nməl ‘ants’, mnəʕ ‘to forbid’

3.2.4. The proposed model: a moraic two-root node model

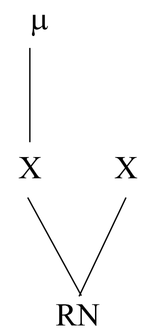
Though the motivation behind proposing a two-root node model was to replace the other models in the market, our goal is to sustain the moraic model and supplement it with the two-root node one. This can be easily done since the two models make predictions about different levels of representation. A moraic level is needed to capture the prosodic phenomena that involve geminates, while the root level is required to tackle the segmental issues. This is specifically necessary in the case of languages that distinguish between geminates and singletons in terms of weight, and at the same time maintain a degree of patterning between geminates and consonant clusters.

(53) The hybrid model: a moraic two-root node model



The idea of a hybrid model that combines both a prosodic level and a segmental one first came up in Hume et al. (1997) and was later developed in Muller (2001). The central premise underlying this model is that the weight specification of geminates is assigned on a language-specific basis rather than being inherent, whereas length is segmentally encoded via the multiple association to two slots on the skeletal tier. This is represented in the following way:

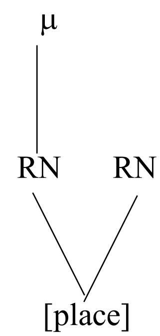
(54)



In this sense, the Composite Model reaffirms the assumptions of the skeletal approach about the structural patterning between geminates and consonant clusters with the main exception of incorporating a moraic level for the sake of eschewing the prosodic problems that a skeletal level alone fails to account for. Therefore, in the presence of a moraic level, the timing slots are stripped off their prosodic nature to be merely segmentally relevant. Under this conception, geminates and singletons are put on an equal footing weight-wise, in that WBP applies (or does not apply) to them equally. This way, a geminate can be assigned a mora only when it is in a coda position, which explains why the mora is restrictedly associated to the first member of the geminate sequence in the above representation.

This version of a hybrid model differs from ours in two ways. First, this model captures the segmental properties of geminates in terms of association to slots on the skeletal tier, whereas in our model the skeleton is dispensed with entirely. Instead, melodic elements are linked directly to prosodic nodes. In the presence of a moraic level, skeletal tiers become unnecessary and redundant. Thus, a hybrid model that replaces timing slots with root nodes is preferable (Stuart Davis, personal communication).

(55)

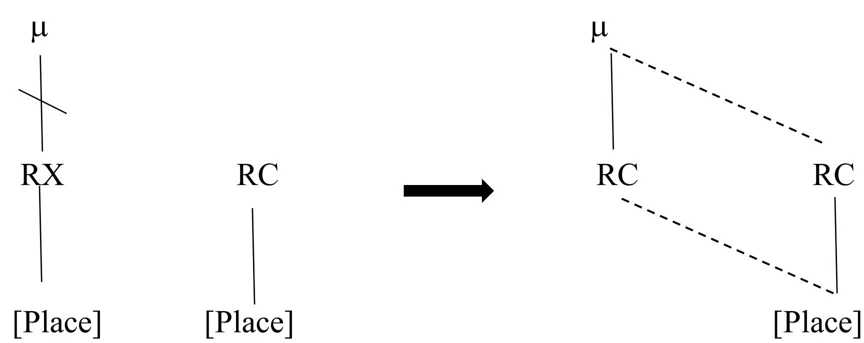


Second, the mora is not part of the underlying structure of geminates. Under the assumption that singletons and geminates are equally weightless, the latter get assigned weight on a language-specific basis. In our model, the mora is underlying. Hence, geminates are expected to be moraic wherever they occur. Also, in our model, the mora is linked to both root nodes at the same time. It cannot pick one root node over the other since the information about their syllabic positions is not available yet.

Injecting a two-root node level into our model requires a revision of some of the accounts we suggested earlier under an exclusively moraic model. This primarily concerns the cases of total assimilation, compensatory lengthening and morphological gemination. The geminates involved in these processes were thought to be derived as consisting of a single root node linked to a mora, resembling the structure of lexical geminates. In order to maintain a unified structure between all types of geminates (i.e. derived or underlying), we should find a way to derive the two-root structure in derived geminates as well.

To achieve a unified representation of geminates in MA, we suggest a structure motivated root node epenthesis process, whereby an empty root node is inserted to provide a segmental interpretation to the new weight of the spreading consonant. Then, the newly inserted root node shares the mora and the features of the spreading root node. This way the derived structure would exactly match the structure proposed for underlying geminates. This epenthesis process is needed to derive the right structure of geminates generated by assimilation, CL and morphological lengthening.

(56) Deriving the moraic two-root node representation



This solution is based on Selkirk's (1990) mora motivated epenthesis process devised to provide a two-root account of CL. However, the difference is that, for Selkirk, weight is not crucial to representing length. Rather, the latter is purely segmentally determined via the two-root node structure.

The central assumption embedded into claiming that a geminate underlyingly consists of a mora that is linked to two root nodes is that weight and length are in a corresponding relationship. That is, it is believed that both prosodic weight and segmental length are essential properties of making a geminate. Phrased differently, it is claimed that a consonant can be long only in the existence of both a moraic component and a segmental one.

4. Conclusion

In this chapter, we have tried to describe the phonological behavior of geminates in MA and, on that basis, advance a representational model that can best capture the observed tendencies without falling behind in terms of keeping up with theoretical development in phonological theory. First, we have shown that geminates in MA display a dual behavior by combining properties of both unit structures (i.e. singletons) and cluster structures (i.e. consonant sequences). On the one hand, it has been shown that geminates pattern with consonant clusters with regards to lexical distribution, schwa triggering, hetrosyllabicity, and breaking by means of full vowels. On the other hand, it has been noted that geminates maintain their integrity in relation to schwa epenthesis, behaving more like monolithic structures.

In response, we have attempted to identify the most appropriate representational framework that could naturally account for the inconsistency characterizing geminate behavior in MA. To this end, it has been shown that, in nonlinear phonology, a geminate could be represented as a single melodic element associated either to multiple skeletal positions, a mora or two root nodes. While the skeletal approach came very close to fully characterizing the patterning of geminates in MA, it has been criticized for failing to explain the cross-linguistic asymmetries related to weight sensitive phenomena, such as compensatory lengthening, stress placement and word size minimality. Alternatively, moraic theory has been successful in explaining the asymmetrical variations displayed by weight sensitive patterns, and was accordingly considered a better representational framework. However, a moraic approach to geminate representation fails miserably in accounting for the bi-positional properties of geminates in MA and a slew of other languages. As a reaction, a hybrid model was suggested whereby geminates are represented as underlyingly moraic and bi-positional by virtue of consisting of a mora shared by two root nodes. It was argued that this could provide a universal representation of geminates that can explain their varying patterning across languages.

CHAPTER FOUR

EVIDENCE FROM MOROCCAN ARABIC FOR THE MORAIC REPRESENTATION OF GEMINATES

1. Introduction

In the previous chapter, we have shown that the behavior of geminates in MA seems to tip the balance more in favor of the skeletal representation than the moraic representation. This is mainly manifested by the analogous lexical distributional properties and phonological patterning of geminates and consonant clusters. However, upon further scrutiny, it has been demonstrated that autosegmental skeletal theory alone fails miserably when it comes to explaining other independent issues such as the status of schwa in MA and the asymmetrical nature of compensatory lengthening cross-linguistically. Therefore, and in light of the overwhelming available evidence from a whole host of languages in support of the moraic status of geminates, we reacted to these inadequacies by putting forward a moraic analysis of consonantal length in MA. In this chapter, it will be shown that MA can provide additional evidence for the underlying moraicity of geminates as many other Arabic dialects do.

Almost every Arabic dialect contains geminate consonants of some sort (Davis, 2014). Many of these dialects constitute a good source for data in support of the moraic conception of geminates. For example, Hadhrami Arabic, spoken near the southern coast of Yemen, provides an interesting argument against the patterning of geminates with consonant clusters. In this dialect, final consonant clusters are not allowed while final geminates are. This fact is indicated by the following items:

(1) Hadhrami Arabic (Davis and Ragheb, 2014:8)

a. /qird/ - [qirid]	'monkey'
b. /bint/ - [binit]	'girl'
c. [rabb]	'Lord'
d. [ʔaxaff]	'lighter'

The items in (1a)-(1b) show that Hadhrami resorts to vowel epenthesis to break up underlying consonant clusters. The items in (1c)-(1d) illustrate that epenthesis does not take place in the case of underlying geminates. This clearly demonstrates that geminate consonants exhibit distinct properties from consonant clusters. Contrary to Hadhrami Arabic, both final geminates and final clusters are allowed in MA (e.g. səmm 'poison' and fərx 'bird'). However, these distributional asymmetries/symmetries between geminates and clusters do not directly argue for or against the inherent moraicity of geminates.

Weight sensitive phenomena such as stress, word minimality and compensatory lengthening can tell us more about the underlying nature of geminates. For instance, in San'ani Arabic, Weight-by-Position is limited to one of the last three syllables of the word. That is, a CVC syllable in pre-antepenultimate position acts as monomoraic, failing to attract stress (item (2a)). Interestingly, stress does fall on CVG and CVV syllables even when in pre-antepenultimate position (items (2b)-(2c)). This can be attributed to the inherent weight of the geminate, making a CVG syllable heavy whatever its position in the word is.

(2) San’ani Arabic (Watson, 2002:82)

a. mak.ˈta.ba.ti:	‘my library’
b. ˈha:.ka.ða.ha:	‘like this’
c. mu.ˈsadʒ.dʒi.la.ti	‘my recorder’

In MA, evidence for the effect of geminates on stress placement is not easy to come up with because all closed syllables in MA have equal weight due to the application of Weight-by-Position.

Another equally strong piece of evidence in support of the underlying moraic nature of geminates from an Arabic dialect is found in Lebanese Arabic, which imposes a bimoraic minimality condition on its content words. As a result of this condition, Lebanese Arabic has no words of the shape CV and CVC due to their monomoraic status. Yet, CVG words are frequently attested thanks to their bimoraic size, in which the geminate and the vowel contribute a mora each. Examples of words that prove this point are provided below:

(3) CVG words in Lebanese Arabic

bajj	‘father’
sitt	‘grandmother’
rabb	‘god’
ʒidd	‘grandfather’
ʔimm	‘mother’

In the same spirit, the central aim of this chapter is to provide evidence, from within MA, that could support the moraic conception of geminates. In particular, this chapter is set to present evidence from three morpho-phonological phenomena that motivate the moraic representation of geminates. These include: word-minimality, compensatory lengthening and morphological gemination. For each phenomenon, we will show that the moraic analysis of geminates delivers better insights and allows for cross-linguistically generalizable accounts.

2. Evidence from word-minimality

Many languages place a restriction on the prosodic size of their content words. Particularly, a language may demand that words should consist of at least a binary foot, under a syllabic or a moraic analysis (McCarthy and Prince, 1986). As a result, light syllables would not be permitted to constitute words on their own in these languages. In order to ensure this word minimality requirement, languages of this type make use of various repair strategies to keep their words in check, either via entirely banning light monosyllabic words or enhancing their structure by augmenting their size (e.g. vowel lengthening) or making them prosodically heavier (e.g. WBP). In OT, the requirement of a minimal size on words is ensured by the stipulation of the constraint FT-BIN (short for Foot-Binarity):

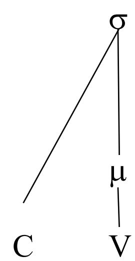
(4) **FT-BIN** (Prince and Smolensky, 1993):

Feet should be binary under moraic or syllabic analysis.

This constraint penalizes every word with less than two moras or two syllables for that matter. By doing so, such a constraint ensures that input forms meet the minimal word threshold required by the language in question. The details as to how this requirement is satisfied are defined by the interaction with other constraints.

As has been noted elsewhere (Al Ghadi, 1994; Boudlal, 2001), MA enforces the word minimality condition on its content words. This is manifested by the nonexistence of words of the shape CV for being monomoraic.[9] Thus, every CV candidate would lose in any possible competition for it is always going to violate the undominated FT-BIN.

(5) Monomoraic syllables in MA



Despite qualifying as bimoraic by virtue of containing long vowels, CVV words do not emerge in MA for the simple reason that long vowels are not part of its phonological system. In OT terms, the constraint against long vowels (i.e. *V:) happens to be high-ranked in the grammar of MA.

However, it is observed that MA allows words consisting of a geminate plus a short vowel (i.e. GV), see (6) for examples, which begs the question: what difference does it make to have a geminate in a word? The answer to this question lies in Morén (1999, 2003) distinction, which maintains that weight comes in two flavors, namely distinctive and coerced. Distinctive weight is the result of a lexical specification of moraicity, whereby segments (i.e. vowels and

[9] Exceptions to this restriction include the words: ʒa ‘come’, ma ‘water’, xu ‘brother’. These words rarely occur in isolation, hence their sub-minimality is often compensated for through combining with other morphemes, for example ʒa-w ‘they came’, l-ma ‘the water’, xu-h ‘his brother’.

geminates) are underlyingly moraic. Coerced weight, on the other hand, stems from a restriction on surface moraicity in some phonological context (e.g. Weight-by-Position).

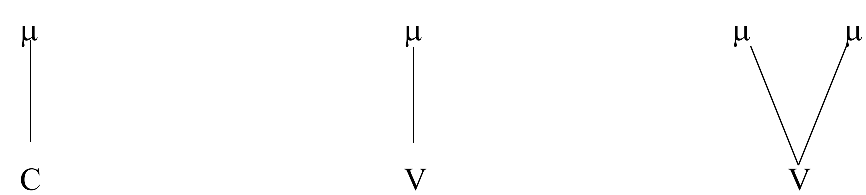
(6) Moraic word-initial geminates in MA

ddi	*di	‘to take’
mm^wi	*m^wi	‘my mother’
bb^wa	*b^wa	‘my father’
lla	*la	‘No’
ssi	*si	‘Mr.’

Within the realm of moraic theory, it is argued that geminates and vowels have distinctive weight by dint of being inherently moraic. This means that they would tend to contribute weight to their words. In support of this claim, the presence of GV words in MA, as opposed to CV words, could constitute evidence for the moraic nature of geminates in MA. Therefore, GV words can be said to be possible because they are inherently bimoraic, with each segment contributing a mora, unlike CV shapes where singletons are nonmoraic.

(7) Distinctive weight in moraic phonology

Geminate Short vowel Long vowel



We posit that the distinctive weight of GV words is ensured by the faithfulness constraint IDENT-Weight, which requires underlying weight specifications of segments to be preserved.

(8) **IDENT-Weight**:

An output segment should bear the same weight as an input one.

This constraint has to be ranked over the markedness constraints *μ/C and *μ/V, militating against moraic consonants and moraic vowels, respectively.

(9) Constraints against moraic elements (Sherer, 1994)

***μ/C**: Moraic consonants are not allowed

***μ/V**: Moraic vowels are not allowed

The interaction between these constraints is illustrated in the following tableau:

(10)

/CμVμ/	IDENT-W	*μ/C	*μ/V
☞ a. CμVμ		*	*
b. CVμ	*!		*
c. CμV	*!	*	

By violating the high-ranking IDENT-W in different ways, candidates (10b) and (10c) are excluded. The optimal candidate, however, obeys IDENT-W by preserving the weight specification of the input. Therefore, GV words inherently satisfy the word minimality condition. Another way to be inherently bimoraic is for a word to include two full vowels, each with a mora: CVCV, CCVCV, CVCVC, CVCCV. Examples of these word shapes are provided in (11) below:

(11) Inherently bimoraic word shapes

CVCV		CCVCV	
ʃita	‘broom’	mdina	‘city’
fuṭa	‘towel’	ħdida	‘iron bar’
mika	‘plastic bag’	brika	‘lighter’
biḍa	‘egg’	rwina	‘mess’
CVCVC		**CVCCV**	
ʒiran	‘neighbors’	saqṭa	‘lock’
sarut	‘key’	ʒumla	‘sentence’
ħanut	‘grocery shop’	manṭa	‘blanket’
kamun	‘cumin’	ʃanṭa	‘suitcase’

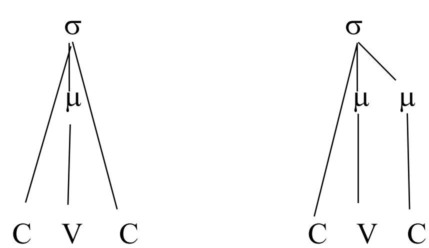
However, words of the shape CVC and CCV, like the ones in (12), also abound in MA despite being monovocalic. This raises the question: what is that makes CVC and CCV words bimoraic despite being monovocalic? To answer this question, we refer to Morén’s second type of prosodic weight, i.e. coerced weight. On this basis, we show that the prosodic minimality of these forms is equally obtained through the interaction of constraints that are independently needed for regulating syllabic well-formedness in MA.

(12) Bimoraicity by coercion

CCV		CVC	
ʃra	‘buy’	ʃaf	‘see’
ʕja	‘get tired’	gal	‘say’
ʒra	‘run’	ħiṭ	‘wall’
bka	‘cry’	ʕid	‘feast’

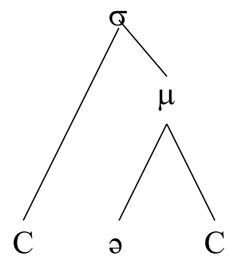
In the case of CVC words, we postulate that bimoraicity is achieved thanks to the language specific preference of Weight-by-Position, which assigns a mora to every consonant in coda position, making every closed syllable with a full vowel bimoraic.

(13) Moraic codas in MA



Therefore, closed syllables in MA are always heavy, except when headed by a schwa, which, despite being a vowel, is nonmoraic. This discrepancy stems from the fact that, unlike full vowels, a schwa can never occur in an open syllable *Cə and cannot be a head of monosyllabic words of the shape *CəC. In multisyllabic words, however, a schwa heading a syllable shares the mora of its coda.

(14) Closed schwa syllables are monomoraic (Bensoukas, 2006; Bensoukas & Boudlal, 2012a-b)



Turning back to CVC words, we posit that the word minimality of this pattern is met by assigning a second mora to the consonant in coda position, CVμCμ. We believe that this is drivable from the constraint *APPENDIX, which prohibits weightless codas.

(15) ***APPENDIX** (Sherer, 1994):

Weightless codas are prohibited

This constraint must dominate the constraint against moraic codas (i.e. *μ/CODA) in order to coerce the consonant in the coda position into acquiring weight (See Rosenthall and Hulst (1999) for similar ranking). The following tableau demonstrates this interaction:

(16)

/CVμC/	*APPENDIX	*μ/CODA
☞ a. .CVμCμ.		*
b. .CVμC.	*!	

In previous accounts, namely Al Ghadi (1994) and Boudlal (2001), it is assumed that the output form CVμCμ is derived through the following constraint interaction: FT-BIN >> H-NUC

>> *CODA. This illustrated by the tableau below, where coda consonants are believed to be coerced into acquiring weight mainly to satisfy the word minimality threshold, represented here by the constraint FT-BIN.

(17)

/CVμC/	FT-BIN	H-NUC	*CODA
☞ a. .CVμCμ.			*
b. CVμ.C		*!	
c. .CVμC.	*!		

Yet, when we consider larger patterns, our analysis appears to be more consistent in deriving moraic codas. For example, the pattern CVCVC is canonically bimoraic, hence it would be unnecessary to make its coda moraic for FT-BIN reasons. Thus, both CVμ.CVμCμ and CVμ.CVμC would satisfy FT-BIN and both would violate *CODA, leaving us with a tie.

(18) Tie between candidates (a) and (c)

/CVμCVμC/	FT-BIN	H-NUC	*CODA
☜ a. CVμ.CVμCμ			*
b. CVμ.CVμ.Cμ		*!	
☜ c. CVμ.CVμC			*

This specific pattern proves that the moraic status of the coda is not ensured by Ft-Bin, but rather it is the result of an independent requisite against moraless codas, represented by *APPENDIX.

(19)

/CVμCVμC/	*APPENDIX	*μ_h/C	*μ/CODA
☞ a. CVμ.CVμCμ			*
b. CVμ.CVμ.Cμ		*!	
c. CVμ.CVμC	*!		

Words of the shape CCV are also abundant in MA. In this regard, we purport that the bimoraicity of such words is achieved by means of creating a syllabic consonant out of the first member of the consonant cluster, C.CV.

(20) Syllabic consonants in MA

$$\begin{array}{c}\sigma\\ | \\ \mu_h \\ | \\ C\end{array}$$

We argue that the prosodic minimality of this pattern derives from the interaction between the constraints *COMPLEX and *μ_h/C, where the former is the dominant constraint: *COMPLEX >> *μ_h/C. Thus, candidate (21b) is ruled out primarily for violating *COMPLEX. Candidate (21a) wins out for satisfying *COMPLEX and satisfying word minimality.

(21)

/CCVμ/	*COMPLEX	*μ_h/C
☞a. Cμ.CVμ		*
b. .CCVμ.	*!	

(22) **$*\mu_h$/C**:

Consonants cannot not be moraic syllable heads.

Al Ghadi (1994:7) claims that the structure C.CV is derivable from either FT-BIN or *COMPLEX. We maintain that such a claim poses a serious problem for the grammar since it leads to a constraint disjunction whereby both FT-BIN and *COMPLEX can yields the winner C.CV, which undermines the ranking arguments *COMPLEX >> $*\mu_h$/C or Ft-Bin >> $*\mu_h$/C. In OT, "the ranking argument is secure only if there is no third constraint that could also be responsible for the winner beating the loser" (McCarthy, 2008:42).

(23) Constraint disjunction: both FT-BIN and *COMPLEX can do the job

/CCVμ/	FT-BIN	*COMPLEX	$*\mu_h$/C
☞ a. Cμ.CVμ			*
b. .CCVμ.	*W	*W	L

This tableau shows that a solid ranking argument cannot be guaranteed since both FT-BIN and *COMPLEX do the same job. They both disfavor the loser CCV and prefer the winner C.CV.

To resolve this problem, we look at larger patterns. The pattern CCVCV is canonically bimoraic by virtue of containing two full vowels. Hence, the structure C.CV.CV could not derive from the ranking FT-BIN >> $*\mu_h$/C since the prosodic minimality requirement is already satisfied. As a result, candidate (24b) loses only for violating *COMPLEX, while it ties with the winner for FT-BIN.

(24)

/CCVμCVμ/	*COMPLEX	*μ_h/C
☞ a. C.CV.CV		*
b. CCV.CV	*!	

In this case, the only way to account for the syllabic consonant at the beginning of the pattern is by positing the ranking *COMPLEX >> *μ_h/C. As a result, *COMPLEX appears to be consistent in deriving syllabic consonants everywhere in the grammar, while FT-BIN seems to do so only when the canonical pattern fails to meet the prosodic threshold. Based on this, we conclude that the creation of syllabic consonants in MA results from the constraint *COMPLEX. Such a claim has two advantages: first, it resolves the issue of constraint disjunction through securing the ranking argument *COMPLEX >> *μ_h/C. Second, it maintains a consistent analysis of the creation of syllabic consonants, instead of positing that they derive from FT-BIN in some instances and *COMPLEX in others.

Boudlal (2001:66/7) approaches the problem differently. It proposes to divide the labor of creating syllabic consonants between *COMPLEX and FT-BIN, whereby the former creates a minor syllable out of the designated consonant cluster without assigning it a mora, while the latter forces it to be moraic. For this, Boudlal distinguishes between two constraints: *Minor-σ and H-NUC. When *Minor-σ is dominated by *COMPLEX, a member of the consonant cluster becomes a syllable of its own. When H-NUC is dominated by FT-BIN, a consonant becomes moraic. According to Boudlal, both *COMPLEX >> *Minor-σ and FT-BIN >> H-NUC are simultaneously active. The structures CCVμ, C.CVμ and Cμ.CVμ then compete.

(25) Ft-Bin wrongly rules out candidate (b)

/CCVμ/	FT-BIN	*COMPLEX	*Minor-σ	H-NUC
☞ a. Cμ.CVμ			*	*
b. C.CVμ	(*!)		*	
c. CCVμ	*!	*		

However, Boudlal (2001) misses the point that the interaction *COMPLEX >> *Minor-σ suffices to satisfy word minimality since the form C.CVμ is disyllabic even if the minor syllable is not assigned a mora, meaning that it satisfies FT-BIN under syllabic analysis. Hence, the interaction FT-BIN >> H-NUC becomes redundant and unnecessary. In other words, the constraint FT-BIN would not be able to distinguish between the output structures Cμ.CVμ and C.CVμ as both meet word minimality. The tableau in (26) illustrates this point:

(26) Candidates (a) and (b) cannot be distinguished by Ft-Bin

/CCVμ/	FT-BIN	*COMPLEX	*Minor-σ	H-NUC
☜a. Cμ.CVμ			*	*!
☜b. C.CVμ			*	
c. CCVμ	*!	*		

In the current analysis, the constraint $*\mu_h/C$ carries the effects of both *Minor-σ and H-NUC. Hence, when dominated, the emergent structure would be a syllabic consonant that is moraic. Besides, our analysis is consistent in using different instantiations of the constraint *μ/C to derive all cases of moraic consonants in MA.

Since FT-BIN cannot always dominate *μ/Coda and *μ_h/C, it is safe to assume that it has no active role in triggering the creation of moraic consonants. Rather, *μ/Coda and *μ_h/C are dominated by *APPENDIX and *COMPLEX, respectively. As we consider words of the shape CCVC, it becomes even clearer that FT-BIN has no direct interaction with *μ/Coda and *μ_h/C. If creating a syllabic consonant or assigning a mora to a coda consonant are indeed done for the purposes of meeting the word minimality threshold, then only one of these options would suffice in the case of CCVC words. If so, Cμ.CVμC, where the coda is nonmoraic, and CCVμCμ, where the onset is complex, could be two possible ways to satisfy FT-BIN. This clearly goes against the fact that closed syllables in MA are independently heavy and margins are not complex.

(27)

CCVC	
slaħ	'weapon'
ɣlaf	'cover'
draʕ	'arm'
ṣdaʕ	'noise'

(28)

/CCVμC/	*COMPLEX	*APPENDIX	$*\mu_h$/C	*μ/C
☞ a. Cμ.CVμCμ			*	*
b. .CCVμCμ.	*!			*
c. Cμ.CVμ.Cμ			**!	*
d. .CCVμC.	*!	*		
e. Cμ.CVμC		*!	*	

Thus, it is better to think of FT-BIN as an undominated constraint whose only role is to evaluate the prosodic minimality of words without being able to compensate for it. That is, when subminimal inputs such CV and CəC are evaluated, FT-BIN would prohibit them from turning into words for being monomoraic. When subminimal inputs such as CVC or CCV are evaluated, *COMPLEX and *APPENDIX would turn them into bimoraic structures, satisfying FT-BIN as a result.

To reiterate, three instances of moraic consonants could be identified in MA. These come in two main flavors: underlying (i.e. distinctive) or derived (i.e. coerced). Despite the apparent differences, our account unifies the emergence of moraic consonants in MA by positing different instantiations of the markedness constraint prohibiting consonants from being moraic. The basic constraint interactions that govern the occurrence of moraic consonants in MA are restated below:

(29) Moraic Consonants and word-minimality in MA

a. Geminates:

/CμVμ/	IDENT-W-C	*μ/C
☞ a. CμVμ		*
b. CVμ	*!	

b. Syllabic consonants:

/CCVμ/	*COMPLEX	*μ_h/C
☞a. Cμ.CVμ		*
b. .CCVμ.	*!	

c. Moraic codas:

/CVμC/	*APPENDIX	*μ/CODA
☞ a. .CVμCμ.		*
b. .CVμC.	*!	

When considered in the context of monosyllabic words, moraic consonants give rise to two basic types of word minimality:

(30) Types of moraic consonants in MA

a. Inherent word minimality: CVG IDENT-W >> *μ/C

GV IDENT-W >> *μ/C

b. Derived word minimality: CVC *APPENDIX >> *μ/CODA

CCV *COMPLEX >> *μ_h/C

This distinction rests on the assumption that the distributional equivalence between geminates, on the one hand, and clusters or singletons, on the other, is not a determinant of their prosodic nature. Despite the apparent similarities prosody-wise, each structure is the result of a unique constraint interaction.

3. Evidence from syllable structure and moraic structure

The next piece of evidence to support the moraic conception of geminates in MA will be drawn from data pertaining to syllable and moraic structure. This section will be divided into two subsections. While both subsections would share the purpose of introducing further evidence from MA in favor of geminate moraicity, they will be offering two different analyses of the same phonological phenomenon. The data in question involves a case of consonant lengthening that seems to be triggered by the affixation of the vocalic 3rd person pronoun /u/ to past tense triliteral feminine verbs. Some examples are given below:

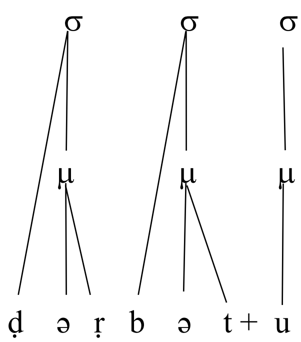
(31) Gemination in the context of prosodic structure

/ḍəṛbət-u/	ḍəṛbə**tt**u	'she hit him'
/ʃəṛbət-u/	ʃəṛbə**tt**u	'she drank it
/qətlət-u/	qətlə**tt**u	'she killed him'
/nəʃṛət-u/	nəʃṛə**tt**u	'she hung it'

3.1. Gemination by syllabification analysis

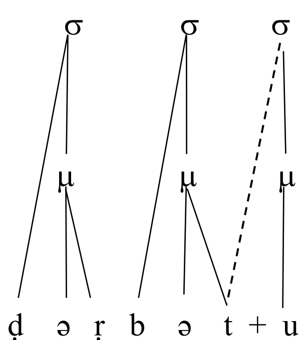
The gemination in the items on the right in (31) could be thought of as following from syllable well-formedness, in that the affix [t], in forms such as *ḍəṛbə**t*** 'she hit', lengthens after attaching to the third person pronoun [u] for the sake of supplying the ensuing syllable with an onset. This process can be illustrated as follows:

(32)



Affixation

/ḍəṛbə**t-**u/ 'she hit- him'



Lengthening

[ḍəṛbə**tt**u] 'she hit him'

In dealing with this case of gemination, we choose to refer to the constraints shown in (33) below:

(33) Constraints responsible for gemination by syllabification

a. ONSET: syllables must have onsets.

b. *GEM: geminate consonants are banned.

The constraint in (33a) prohibits the existence of geminate consonants in the language. (33d) stipulates that syllables must have an onset. The tableau in (35) demonstrates how these two constraints interact to yield the expected structure. It shows that in the wake of affixing the vocalic 3rd person personal pronoun, the constraint ONSET demands that the subsequent syllable must have an onset. If the mora is linked to a consonant, a geminate structure is derived. For this to be possible, ONSET should outrank the constraint against long consonants, *GEM.

(34)

/ḍərbət-u/	ONSET	*GEM
☞a. ḍəṛ.bət.tu		*
b. ḍəṛ.bət.u	*!	

The optimal candidate lengthens the consonant /t/ to create an onset for the subsequent syllable. Creating an onset this way happens at the expense of violating the low-ranking *GEM. The candidate in (34b) loses for violating ONSET. The way we think this supports a moraic conception of geminates lies in the fact that the created geminate ends up being associated with a mora, whose role changed from being purely prosodic to being segmentally relevant.

Recall that, in this work, it is assumed that derived geminates in MA acquire the same phonological structure that characterizes underlying geminates. In particular, a true geminate, be it derived or lexical, should be marked for length both moraically and segmentally. In OT, derived geminates can acquire a moraic two-root node structure thanks to the interaction between the constraint representing the condition on consonantal length to be moraic and segmental, written as W-to-L Corr (short for weight-to-length correlation) and the faithfulness constraint DEP-RN, militating against the insertion of empty root nodes:

(35)

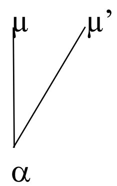
μ RN	W-to-L Corr	DEP-RN
a. μ RN	*!	
☞b. μ RN RN		*

3.2. Compensatory lengthening analysis

This subsection attempts to offer a quite different perspective on the same case of gemination analyzed above. What is new is that we are going to think of the case of gemination illustrated in (31) as emerging from a process of compensatory lengthening. Thinking of this as a case of CL is done for the purpose of pointing out to an even stronger evidence for the moraic nature of geminates.

CL generally involves the deletion of the melodic features of a moraic segment, leading into the lengthening of a nearby segment in order to preserve the mora left behind by the deleted melodic element (Hayes, 1989). The most typical case of CL consists in the lengthening of a vowel upon the deletion of a weight carrying coda.

(36) Compensatory Lengthening in Moraic Phonology (Hayes, 1989):



where μ’ is a segmentally unaffiliated mora

Arguably, we will assume that the deletion of the vowel /a/ of the feminine suffix /-at/ triggers a word internal process of consonant lengthening. Such an analysis requires reference to what is believed to be the canonical form of the derived structure. This can be exemplified by the relation between the following forms:

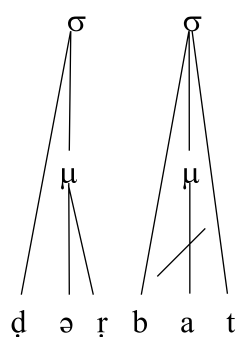
(37) Gemination by compensatory lengthening:

ḍəṛb**at**	ḍəṛbət-u	ḍəṛbə**tt**u	‘she hit him’
ʃəṛb**at**	ʃəṛbət-u	ʃəṛbə**tt**u	‘she drank it
qətl**at**	qətlət-u	qətlə**tt**u	‘she killed him’
nəʃṛ**at**	nəʃṛət-u	nəʃṛə**tt**u	‘she hung it’

Specifically, it is being assumed that in the wake of the deletion of the /a/ of the feminine suffix /-at/, the consonant /t/ turns into a geminate by spreading to the mora left behind. The deletion of the /a/ first yields a structure with a lingering mora (38b). The lingering mora stays unparsed and unfilled until the affixation of the pronoun /u/, which, first, triggers a resyllabification of the form it attaches to (38c). The affixation of the pronoun /u/ also provides the right conditions for filling and parsing the lingering mora (38d). The derived structure then creates an environment that requires an additional schwa epenthesis. These processes are illustrated by the following representations, borrowed from chapter 3 for easy reference:

(38) Gemination as compensatory lengthening in MA

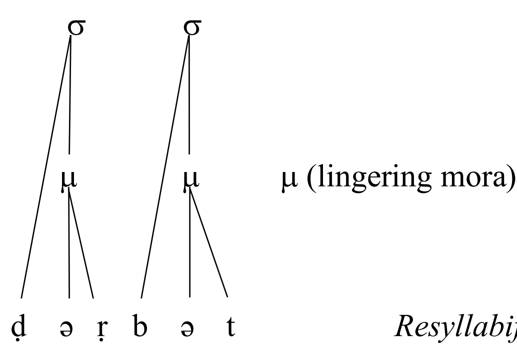
a.



Deletion

[ḍəṛbat] ‘she hit’

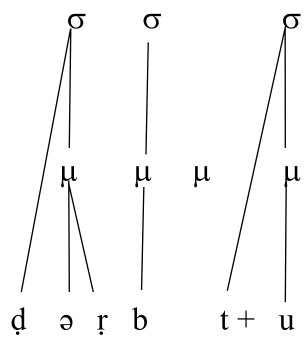
b.



Resyllabification

[ḍəṛbət] ‘she hit’

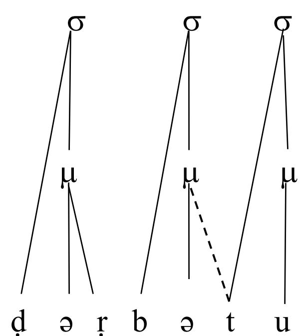
c.



Affixation and resyllabification

/ḍəṛ.b.tu/ ‘she hit+ him’

d.



Compensatory lengthening and resyllabification

[ḍəṛbə**tt**u] ‘she hit him’

At the representational level of the created geminate, a structure-motivated empty root node is inserted to attain segmental length, reflecting the new weight of the spreading consonant. One caveat to emphasize is that CL as seen here takes place only world internally. The reason why this is the case will be explained later in this subsection. With this background in mind, we believe that in order to entertain a constraint-based CL analysis, we first have to account for the deletion of the vowel /a/. To do so, we will make use of the constraints shown in (39) below:

(39) Constraints responsible for CL in MA

a. *a: avoid the vowel /a/ in the output.

b. MAX-V_{-at}: The input vowel of the affix /-at/ must have a correspondent in the output. (No deletion)

(39a) is a markedness constraint that simply disfavors candidates with the vowel /a/. The constraint in (39b) is a faithfulness constraint prohibiting the deletion of input vowel in the affix /-at/. The tableau in (40) showcases the interaction responsible for the deletion of the vowel /a/ subsequent to the suffixation of the affix /-at/ to a verb stem:

(40) Vowel deletion

/ḍəṛbat/	*a	MAX-V_{-at}
☞ a. ḍəṛ.bət.		*
b. ḍəṛ.bat	*!	

The faithful candidate in (40b) incurs a fatal violation of *a. The latter happens to be the dominant constraint in the hierarchy, favoring the unfaithful candidate in (40a). Thus, (40a) emerges as the winner.

At this stage, the syllabification shown in (38c) should be justified. In particular, we have to clarify why /ḍəṛ.bət-u/ has been intermediately syllabified as /ḍəṛ.b.tu/ instead of directly syllabified as [ḍəṛ.bət.tu]. We argue that the reason behind this is to maintain a level of consistency in the syllable structure of MA. One known fact about the 3rd person pronoun /u/ is that it triggers the resyllabification of trisegmental verbs as well as quadrisegmental ones for the sake of providing the /u/ with an onset (Boudlal, 2001:171). This is illustrated by the following examples:

(41) Resyllabification of quadrisegmental verbs in MA

/ṭəṛ.ʒəm+u/	ṭəṛ.ʒ.mu	*ṭəṛ.ʒəm.mu	'he translated it'
/kər.kəb+u/	kər.k.bu	*kər.kəb.bu	'he rotated it'
/ʒəṛ.ʒəṛ+u/	ʒəṛ.ʒ.ṛu	*ʒəṛ.ʒəṛ.ṛu	'he dragged it'
/dər.dəb+u/	dər.d.bu	*dər.dəb.bu	'he rolled it'

Therefore, we believe that the same should be done with forms such [ḍəṛ.bət], which happen to have the same phonological shape as the quadrisegmental verbs above.

(42)

/ḍəṛ.bət+u/ Base:[ḍəṛ.bət]	ONSET	FAITH-OOσ
☞a. ḍəṛ.b.tu		*
b. ḍəṛ.bət.u	*!	

In order to complete the picture, we will have to introduce two additional constraints: LINK-μ and *GEM. The interaction between these two constraints is what captures the CL process. The constraint in (43a) prohibits the existence of geminate consonants in the language. (43b) stipulates that input moras should be preserved in the output.

(43)

a. *GEM: geminate consonants are banned.

b. LINK-μ: an input mora must be realized in the output.

To elaborate, the constraint LINK-μ demands that the lingering mora should associate to some other melodic element. Since the mora gets linked to a consonant, a geminate structure is derived. For this to be possible, LINK-μ should outrank the constraint against long consonants, *GEM.

(44)

/ḍəṛ.b.tu/	LINK-μ	*GEM
☞a. ḍəṛ.bət.tu		*
b. ḍəṛ.b.tu	*!	

The optimal candidate lengthens the consonant /t/ to fill the lingering mora. This leads to the violation of the low-ranking *GEM. The candidate in (44b) loses for being unable to preserve the mora left over upon the deletion of /a/. Note that the lengthening of the /t/ creates a new environment for the application of schwa epenthesis. This conventionally ensues from the ranking of PARSE-Seg >> DEP-ə. The lengthening of the affix consonant in lieu of the stem consonant is decided by independent constraints on syllabic well-formedness. That is, if the /b/ in [ḍəṛbatu] lengthens instead, any possible candidate would be harmonically bounded by the winner for incurring extra violations. For instance, the candidates [ḍəṛ.b.b.tu] [ḍəṛ.bəb.tu], [ḍəṛ.b.bə.tu] and [ḍəṛ.b.bət.u] would violate *μ/C, GEM-Integrity, *μ/ə and ONSET, respectively.

Now, we move on to explain why compensatory lengthening does not take place when the suffix /-at/ is not followed by the vocalic 3rd person pronoun marker (i.e. [-u]). Consider the following examples:

(45)

ḍəṛbat	ḍəṛbət	*ḍəṛbətt	'she hit'
ʃəṛbat	ʃəṛbət	*ʃəṛbətt	'she drank'
qətlat	qətlət	*qətlətt	'she killed'
nəʃṛat	nəʃṛət	*nəʃṛətt	'she cleaned'

what this data shows is that even when the /a/ of the affix deletes, the /t/ does not undergo lengthening. Our account for this relies on a markedness distinction between internal geminates and peripheral geminates, whereby the latter happen to be more marked, hence more disfavored even in languages that allow geminates. We posit a constraint against the occurrence of edge geminates, which we believe has to dominate LINK-μ.

(46) ***EDGE-GEM**: peripheral geminates are not allowed.

(47) **Antigemination effect:** edge geminates are disfavored

ḍəṛbat	*EDGE-GEM	LINK-μ	*GEM
☞ a. ḍəṛ.bət		*	
b. ḍəṛ.bət.t	*!		*

The ranking of LINK-μ between *EDGE-GEM and *GEM means that the phonology of MA prefers leaving the mora unlinked over creating a geminate at the edge. The following tableau summarizes all the interactions at play:

(48)

ḍəṛbat	*EDGE-GEM	LINK-μ	*a	*GEM	MAX-V_{at}
☞ a. ḍəṛ.bət		*			*
b. ḍəṛ.bat			*!		
c. ḍəṛ.bət.t	*!			*	*

The view that this CL process stands as evidence to the moraic nature of geminates in MA is based on the idea that CL patterns are best described in moraic terms. Therefore, since we have a case of consonant lengthening that is moraically motivated, it is more reasonable to assume that all cases of consonant length in MA, be they underlying or derived, are moraic in nature.

4. Evidence from word-formation

The third type of evidence for the moraic analysis of geminates comes from word formation. Specifically, we argue that the process of morphological gemination, which consists in utilizing consonant lengthening as a morphological marker, is best accounted for if length is expressed in moraic terms. The forms that are derived through this process include: the causative, the agent and the instrument. For convenience, some examples are repeated below:

(49) Morphologically derived geminates:

(i) The causative

ktəb	‘to write’	kəttəb	‘to make write’
ɦṛəb	‘to run away’	ɦəṛṛəb	‘to cause to run away’
xsər	‘lose’	xəssər	‘to cause to lose’
zləq	‘slip’	zəlləq	‘to cause to slip’

(ii) The agent noun

fləħ	‘to farm’	fəllaħ	‘farmer’
nʒəṛ	‘to sharpen’	nəʒʒaṛ	‘carpenter’
gzər	‘to butcher’	gəzzar	‘butcher’
bni	‘to build’	bənnaj	‘mason’

(iii) The instrument noun

ɣsəl	‘to clean’	ɣəssala	‘a washing machine
sməʕ	‘to listen’	səmmaʕa	‘headset’
kwi	‘to weld’	kəwwaja	‘a welding machine’
sqi	‘to water’	səqqaja	‘fountain’

There are three possible ways to account for morphological gemination in these forms. We will take the causative form as our illustrative case. The agent and the instrument rely on other vocalic material in their derivation. To start with, the causative can be analyzed as a case of mapping a tri-segmental root (e.g. rkb ‘ride’) to a quadrisegmental template (e.g. CCCC) (Bennis and Iazzi, 1995). Since templatic slots outnumber the segments of the root, four to three, the second segment of the root associates to two positions, creating a geminate structure. The mapping analysis is undermined by the fact that the spreading direction of Arabic, which operates from left to right, fails to derive the right form (e.g. *rkbb). Actually, neither can right to left spreading yield the desired forms (e.g. *rrkb). The analysis, as presented in Bennis and Iazzi (1995), ad-hocly resorts to the pre-association of the second root segment to the second slot in the template to resolve this issue.

Second, it was suggested that morphological causatives in MA could be better accounted for through a process of partial reduplication of the base (Boudlal, 2001). The reduplicative morpheme is set to target the left edge of the base by an alignment constraint. The reduplicant consists of one single segment, which then gets infixed under the pressure of higher-ranking constraints, causing the doubling of the second segment of the root (see chapter 8 for more on this). Based on the nature of these two above-mentioned analyses, they will be jointly dubbed the segmental analysis.

The reduplication analysis of morphological gemination in MA seems to be unrestricted, in the sense that the reduplicant is not specified in terms of a well-defined phonological entity. Within current prosodic morphology, namely templatic morphology (McCarthy and Prince, 1986/1996, 1990a, 1990b), restrictions on the size and shape of reduplicants were found to be cross-linguistically defined in prosodic terms, viz. phonological word, foot or syllable. On this view, the analysis of morphological gemination as a case of segment reduplication appears to be inconsistent with this generalization since it does not refer to a constituent of the prosodic hierarchy. In other words, it is not possible to consider morphological gemination a result of segment duplication since a segment is obviously not a prosodic constituent.

The inadequacies that characterize the segmental accounts of mapping and reduplication can be interpreted as arguments for a moraic theory of length. A moraic analysis can account for morphological gemination more straightforwardly, namely through the affixation of a morphologically induced mora (Lombardi and McCarthy, 1991; Samek-Lodovici, 1992; Bensoukas, 2001; Noamane, 2013). The two analyses can be schematically contrasted as follows:

(50)	a. Moraic analysis	b. Segmental analysis
	ALIGN-L μc	Mapping/RED
	$\downarrow$	$\downarrow$
	μ	
	$\vert$	Ci+Ci
	C	

Unlike the segmental analysis, the moraic analysis of morphological gemination has the advantage of using a well-defined phonological unit that has been theoretically and typologically established. This allows for a generalized crosslinguistic account of morphological gemination. The fact that morphologically derived geminates favor a moraic analysis may work as suggestive evidence that lexical geminates are moraic too. An elaborate treatment of morphological gemination using a moraic model will be provided in chapter eight.

5. Defusing the nonmoraic view of geminates

Moraic phonology predicts that CVV and CVG syllables should always pattern as heavy thanks to their bimoraic nature. Therefore, both of these syllable shapes are expected to attract stress in languages where the latter is weight sensitive. Contrary to this prediction, it has been reported that there exist languages where CVV syllables alone qualify as bimoraic for stress.

This means that CVG syllables, in this case, pattern as monomoraic as they get ignored by stress rules, suggesting that the geminate is nonmoraic.

One such a language, where stress ignores CVG syllables in a weight sensitive system, is Selkup (Tranel, 1991). In Selkup, stress typically falls on the rightmost heavy syllable. If there is no heavy syllable, stress falls on the first syllable instead. The problem is that CVG syllables are systematically ignored by the stress rule in the same way monomoraic CV and CVC syllables are. This is demonstrated by the data below, where the last item shows that a syllable closed by a geminate is not recognized as bimoraic, hence skipped by stress:

(51) Stress in Selkup

quˈmo:qi	‘two human beings’
ˈu:cɨqo	‘to work’
qumo:qliˈlɪ:	‘your two friends’
ˈu:cɨkkak	‘I am working’

To explain this, Tranel (1991) proposes the Principle of Equal Weight for codas. This principle states that in a language like Selkup, where Weight-by-Position is not applicable, all coda consonants should be equally weightless regardless of being singletons or geminates. Davis (1994, 1999a) offers a different view from that of Tranel (1991). Davis defends the moraic conception of geminates by showing that Selkup may just be a language where only a subset of moras are relevant for stress. To elaborate, for Selkup, it may be that only moras dominating vowels are relevant for stress placement while those linked to geminates are not. In OT, this effect can be easily obtained through the work of some high-ranked constraints. Therefore, languages in which geminates pattern as nonmoraic with regards to their stress systems do not constitute conclusive evidence against the underlying moraic nature of geminates.

6. Conclusion

This chapter has been concerned with providing pieces of evidence from MA that can support the moraic conception of geminates, and hence further justify our decision to represent geminates as underlyingly moraic. The pieces of evidence that we have presented in the course of this chapter consist of arguments from word minimality, compensatory lengthening and word formation. First, it has been shown that, despite being monovocalic, GV words in MA are possible. This was argued to follow from the underlying moraic nature of their initial geminates, meaning that GV words meet the word minimality condition by virtue of being underlyingly bimoraic. Second, we have used a case of dialectal compensatory lengthening to demonstrate that a moraic approach to geminate representation is needed. We have argued that the mere fact that compensatory lengthening exists in MA, leading to the creation of geminate instances, constitutes a strong argument in favor of the moraic analysis since CL is a mora preserving phenomenon par excellence. Third, a moraic approach to gemination was shown to be preferable in accounting for morphologically geminated forms such as MCs, ANs and INs. The moraic analysis of morphological gemination in MA was compared to the segmental analyses of reduplication and template mapping, which were revealed to rely on ad-hoc assumptions. The moraic analysis has the advantage of being cross-linguistically generalizable and theoretically supported.

CHAPTER FIVE

THE SYLLABIFICATION OF GEMINATES IN MOROCCAN ARABIC

1. Introduction

This chapter is devoted to answering the following questions: How are geminates syllabified in MA? And in what way do they contribute to the prosodic weight of their words? Answering these two questions is subject to a great deal of controversy and debate since each of the previously discussed models of geminate representation makes different predictions about the syllabic distribution and prosodic status of geminates.

The skeletal model (Goldsmith, 1976; McCarthy, 1979; Clements and Keyser, 1983; Levin, 1985), for example, predicts geminates to behave in much the same way as any sequence of two unlike consonants. In this sense, a geminate would have no special underlying prosodic status at all to distinguish it from other consonants. In addition, its syllabic distribution would be no different from those of consonant clusters. For instance, when they occur at the edges, both structures would be rendered as complex syllable margins.

The moraic model (McCarthy and Prince, 1986; Hayes, 1989; Davis, 1994, 1999, 2003, 2011; Davis and Ragheb, 2014; Davis and Topintzi, 2017), however, predicts geminates to be segmentally equivalent to singletons by virtue of their single root node structure, but prosodically different from them, in that they are underlyingly moraic. Under this view, geminates are expected to inherently and consistently add weight to their syllables. Therefore, moraic theory predicts geminates to occur only in a position where their moras can be calculated (e.g. coda position). In such a model, the syllabic distribution of geminates is believed to follow from independent requirements related to syllabic well-formedness.

The two-root model (Selkirk, 1990; Curtis, 2003) dispenses with the idea of a skeletal tier to which melodies should associate in order to be prosodically licensed. Instead, root nodes become the sole mediator between melodic structure and prosodic structure. Aside from this, such a model does not differ much from the skeletal one in terms of the predictions it makes about the distribution of geminates. Particularly, it too equates geminates and consonant clusters structurally. Quite similarly to the CM, the two-root node theory of geminates presumes that consonant weight assignment is language specific for both geminates and singletons. That is, geminates may or may not be moraic, without any effect on their phonetic length.

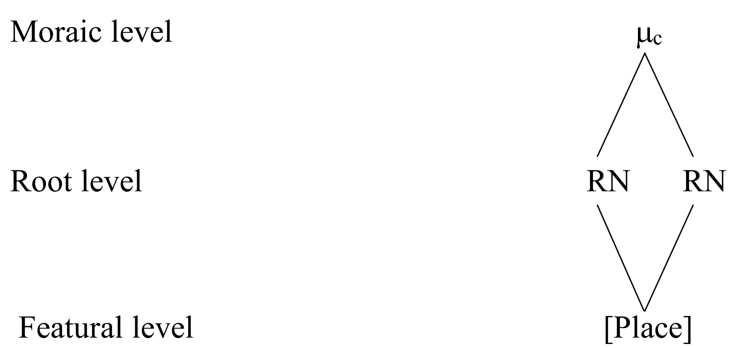
The moraic two-root-node model being defended in this dissertation is set to reiterate the predictions put forward by the standard moraic model and the two-root-node model. It too has at its center the claim that geminates are inherently moraic and, other things being equal, should stay so irrespective of their position in a word. It also advances the claim that geminates are bi-positional by dint of involving two root nodes. This means that such a model combines the predictive powers of both the moraic theory and the two-root-node one. Upon further investigation, some of the predictions made by this model appear to stand in conflict with each other. To resolve this, the predictions made by the two-root node aspect of the model would need to be restricted in favor of the moraic aspect. We will address this issue in the subsequent section.

In the course of this chapter, two fundamental claims would be shaping our treatment of the prosodic properties of geminates in MA. These are as follows: (i) geminates are moraic in all positions and (ii) geminates do not constitute complex syllable margins. Each claim will be justified and defended in due time.

2. Geminates as non-complex margins

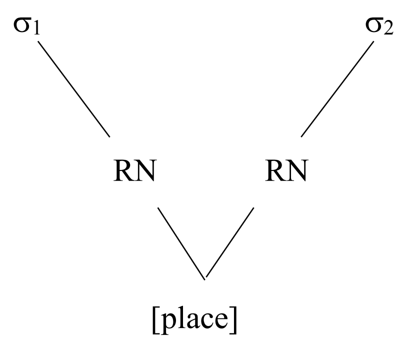
When our proposed model of geminate representation is put into action, a redundancy problem emerges. This can be attributed to the premise that geminates in MA are both underlyingly moraic and bi-positional, and could hence be redundantly syllabified according to the terms of both aspects.

(1) The moraic two-root node model of geminates



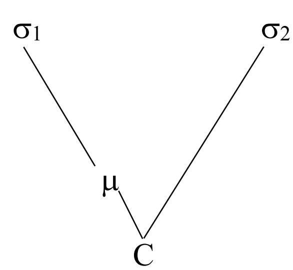
In Particular, when based on their root level structure, geminates should be syllabified the same way normal consonant clusters are. For instance, peripheral geminates would be equated with peripheral clusters in that, in the case of both, complex edges would be at stake, necessitating the creation of a syllabic consonant out of the spare member of the cluster (e.g. m.ʃa ‘to go’ and d.da ‘to take’). Medial intervocalic geminates, on the other hand, would split between syllables supplying a coda for one and an onset for the other.

(2) Segmental-based syllabification



Considering the moraic structure, however, geminate syllabification should normally be subject to the gemination-by-syllabification algorithm (Selkirk, 1990), whereby geminates straddle two syllables for the purpose of securing a mora-bearing position to carry their lexical weight. That is, under a moraic analysis, the hetrosyllabicity of geminates is the consequence of associating the mora to a position where it can be calculated for weight.

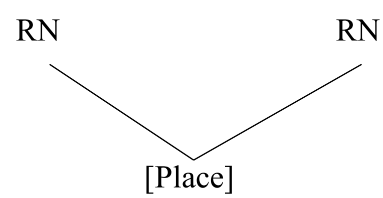
(3) Moraic-based syllabification



The hybrid representation of geminates in question has consequences on word minimality as well. The issue is whether words of the forms GV and CVG meet the word minimality threshold by virtue of creating a syllabic consonant to avoid complex edges or due to the inherent moraic nature of the geminate. Given this state of affairs, two questions arise: (1) What structural level of our model should be the primary basis for the syllabification and prosodic representation of geminates in MA? (2) How can this overlap between the moraic structure and the root structure be resolved?

As a means to resolve this redundancy problem, we contend that despite being represented with two root nodes at the segmental level, geminates are conceived of as semi-unit consonants, unified by virtue of sharing a mora. We entertain the idea that a consonant cluster refers to a pair of totally independent segments with independent feature sets. Hence, since geminates have two root nodes with a shared feature set and a shared mora, they do not qualify as clusters per se.[10] We can think of geminates as fake clusters:

(4) Geminates as fake clusters:



Put simply, geminates do not create complex margins in MA. This line of reasoning is coherent with the claim that it takes multiple place gestures in syllable margins to produce complex edges, as has been argued in Jones (1999). By arguing so, we intend to limit the

[10] When geminates get broken, each root node takes its own copy of the feature specifications to become an independent segment and the underlying mora gets erased as a result (see the next chapter for more on geminate breaking).

predictive power of the two-root part of the representation, hence tip the balance in favor of the moraic part.

(5) True clusters:

$$\begin{array}{cc} \text{RN} & \text{RN} \\ | & | \\ \text{[place]} & \text{[place]} \end{array}$$

Having said that, this is not to deny the fact that geminates and clusters bear some resemblance owing to their two-root node nature. Framed differently, despite the fact that geminates are characterized as bi-positional, their unified prosodic and featural make-up grants them some degree of unity, and hence distinguishes them from normal consonant sequences. At first blush, this seems to be a counterintuitive claim, but we believe that there are specific externally-motivated observations that can support it.

In order to substantiate this claim, we refer to cases of languages that make a distributional distinction between geminates and clusters by allowing one structure but not the other. Especially, we are interested in languages that happen to allow geminates while disallowing clusters of unlike consonants. This is interesting since geminates should represent the most marked case of consonant clusters. A fact like this could mean that geminates and clusters are not completely structurally equivalent. To illustrate this point, we use the case of Trukese and Luganda geminates as examples (Muller, 2001). Both Trukese and Luganda do not allow consonant clusters entirely. Trukese geminates may contrast with singletons in word-medial and word-initial positions. Final geminates, however, are not attested. Consider the following data from (Muller, 2001):

(6) Trukese geminates

a. Word-medially:

ikkair	‘here they are’	ika	‘if’
affat	‘be made clear	afat	‘reveal’
aččawa	‘make slow’	ačawa	‘sp. fish’

b. Word-medially:

kkak	‘ring (noun)’	kak	‘ring (verb)’
ffaat	‘be strung’	faat	shallow water’
čču	‘wooden comb’	ču	‘extracted’

Similarly, geminates in Luganda are observed word-medially and word-initially, while missed word-finally.

(7) Luganda geminates

a. Word-medially

kùggúlá	‘to open’	kùgúlá	‘to buy’
kùssá	‘to breathe’	kùsá	‘to grind’
mùggó	‘stick’	mùgó	‘rim of a pot’

b. Word-initially

ggúlá	‘open (imp.)’	gúlá	‘buy (imp.)’
ssá	‘breathe (imp.)’	sá	‘grind (imp.)’
mmánjú	‘at the back	mánjó	‘at the front’

With this as a background, it is assumed that languages only allow or disallow both structures at the same time by mere accident, but not because of their supposed structural equivalence. Were it due to the latter, it would be very unlikely for languages such as Trukese and Luganda to allow the existence of the more marked structure (i.e. geminates) in the absence of the less marked one (i.e. normal clusters).

Accordingly, it appears that these languages lean toward considering geminates as structurally different from sequences of unlike consonants. The distributional asymmetry between geminates and consonant clusters in Trukese and Luganda could be used to generally argue against a bi-positional representation of geminates. Interestingly, geminates in Trukese do not pattern with singletons either, in that words may end in a singleton but may not end in a geminate (Davis, 2011). Our purpose is to only suggest that the bi-positionality of geminates in our model is different from the bi-positionality of normal clusters.[11] Therefore, we claim that even when geminates exhibit some of the properties of consonant sequences, they still hold certain properties of unit segments. In short, a geminate is not equivalent to a pure sequence of two unlike consonants.

To recap, given the facts displayed by the distribution of Trukese and Luganda geminates, the purely segmental interpretation of geminates assumed by the standard two-root node theory appears to be inadequate as it falls short of explanation when it comes to this positional asymmetry between geminates and clusters. As suggested by our model, a constrained

[11] In fact, the bounded bi-positionality that characterizes our model can explain what standard bi-positional models cannot. For example, the standard two-root nodes model of Selkirk (1990) cannot explain why clusters are totally prohibited in Trukese and Luganda while geminates are totally accepted both initially and medially. The only way to explain this asymmetry, while maintaining a bi-positional view, is to claim that the bi-positionality of geminates is distinct from that of clusters.

segmental approach is possible, where geminates are not totally equivalent to a sequence of two independent consonants despite their being represented as two root nodes. This could explain why initial and medial geminates are possible in Trukese and Luganda whereas initial and medial consonant clusters are not.

Hence, it is safe to go on assuming that geminates in MA do not qualify as complex margins despite their bi-positionality, which clears the way for the moraic aspect of geminates to be the reference structure for the syllabification algorithm. For the sake of convenience, only the moraic structure of geminates will be shown, while the two-root node structure will be referred to only when necessary.

3. The syllabication of geminates

3.1. Word-internal geminates

It has been securely established in the literature that word-medial geminates are cross-linguistically more common (Davis and Topintzi, 2017). On the contrary, peripheral geminates tend to be less so. As has been previously demonstrated for MA, geminates may occur in all positions, including word-finally and word-initially. Word-medial geminates in MA can be either lexical or morphologically derived. In terms of their environment, they are mostly intervocalic, appearing between either two full vowels, two schwas or a schwa and a full vowel. Examples for each situation are provided below:

(8) Medial geminates in MA

a. Lexical medial geminates:

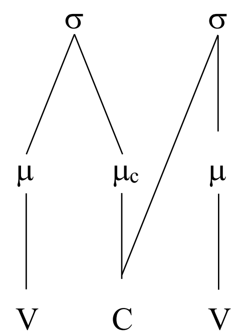
ṣakka	‘tobacco shop’
saʕʕa	‘clock’
kaʃʃa	‘blanket’
madda	‘substance’
ħənna	‘henna’
sənna	‘tooth’
səkka	‘rail’
rəẓẓa	‘turban’

b. Derived medial geminates:

nəʕʕəs	‘to make sleep’
zəlləq	‘to make slide
fəllaħ	‘farmer’
nəʒʒaṛ	‘carpenter’
ɣəssala	‘washing machine’

Under the assumption of being underlyingly moraic, geminates are expected to preserve their moraicity wherever they occur. In accordance with this prediction, the syllabification of geminates in moraic theory is straightforwardly accounted for intervocalically. In this location, a geminate is typically divided between two syllables, supplying a coda for one and an onset for the other. This gives rise to the following flopped structure:

(9) The syllabification of medial geminates



In this context, the geminate closing the first syllable satisfies *APPENDIX by means of its inherent weight. In fact, the only reason we have a coda in the first syllable is to avoid a moraic onset in the second. This means that hetrosyllabicity does not necessarily imply bi-positionality. Even in a purely moraic representation, where geminates are conceived of as one root-node structures, the hetrosyllabicity of geminates is still possible.

To account for the hetrosyllabicity of medial geminates in MA, the following constraints will be needed:

(10) Constraints responsible for the syllabification of geminates in MA

a. **ONSET**: syllables must have onsets

b. **IDENT-Weight**: input and output weight configurations must be identical

c. ***μ/ONSET**: moraic onsets are not allowed

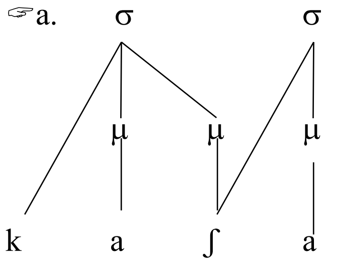
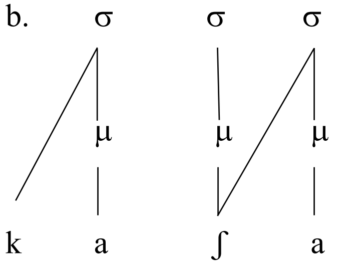
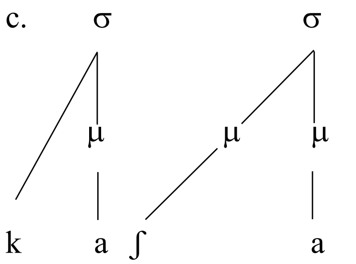
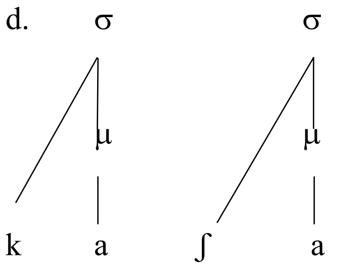
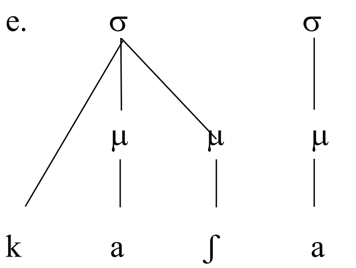
d. ***μ_h/C**: syllabic consonants are not allowed

e. ***μ/CODA**: moraic codas are not allowed

By ranking ONSET and IDENT-W high in the hierarchy, the medial geminate will have to both serve as an onset and preserve its weight specification. The markedness constraints *μ/ONSET, *μ_h/C and *μ/CODA represent different instantiations of the more general markedness constraint *μ/C, militating against moraic consonants. These constraints should be ranked below IDENT-W so that the mora of the geminate can be maintained. Each one of them stands for a possible scenario of how the mora of the geminate can be preserved. The ranking of these constraints among themselves decides on the position of the mora. If, for example, *μ_h/C is the lowest ranking constraint, then the geminate would have to be the head of a syllabic consonant to hold its mora, in addition to serving as an onset.

While syllabic consonants in MA are allowed, they do not represent the least marked option of moraic consonants. Rather, the least marked scenario is for the mora to be in a coda position. This is the option that we believe the MA syllabification algorithm opts for. This can be achieved by ranking *μ/CODA the lowest in the hierarchy. For now, it will be assumed that *μ_h/C is ranked below *μ/ONSET, and hence represents the second least marked option. This ranking will be justified when we deal with the syllabification of initial geminates. Before that, let's see how our hierarchy evaluates the possible candidates and eventually chooses an optimum.

(11) The syllabification of medial geminates in MA

μ /kaʃa/	ONSET	IDENT-W	*μ/ONSET	*μ_h/C	*μ/CODA
a. 					*
b. 				*!	
c. 			*!		
d. 		*!			
e. 	*!				

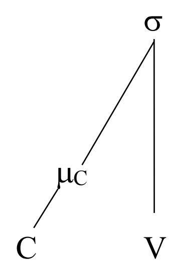
Considering an input such as *kaffa* 'blanket', we obtain the following results. Candidates (11d) and (11e) violate the high-ranking constraints IDENT-W and ONSET, respectively. Candidate (11d) satisfies ONSET, but it fails to preserve the weight specification of the input geminate. Candidate (11e) does preserve the mora in the coda of the first syllable, but without extending the geminate to the onset position of the next syllable. Candidates (11a), (11b) and (11c) represent different options to satisfying both ONSET and IDENT-W. Candidate (11c) has a moraic onset, which is penalized by *μ/ONSET. The ranking of *μ/ONSET above *μ_h/C and *μ/CODA makes it the least favored option. The geminate in candidate (11b) spans two syllables for the sake of avoiding moraic onsets. Its mora serves as the head of its own syllable. Yet, this option is disfavored by our hierarchy, in that *μ_h/C is ranked above *μ/CODA. Candidate (11a) also resorts to extending the geminate over two syllable nodes, but it differs from candidate (11b) in that it puts the mora of its geminate in a coda position, which is the least marked option according to our hierarchy.

The question now is: Is heterosyllabicity a characteristic of medial geminates only? A follow up question is: Where do peripheral geminates place their lexical moras? The answers are provided in the subsequent sections.

3.2. Word-initial geminates

Initial geminates pose a challenge to the moraic conception of consonantal length. Based on an overwhelming array of evidence from stress patterns, it has been firmly concluded that onsets are weightless. In this regard, the inherent weight of geminates appears to conflict with the presence of geminates in an initial position where, unlike in the case of intervocalic geminates discussed above, there is no possibility of coda association, and the only way out seems to have the geminate fully contained in the onset.

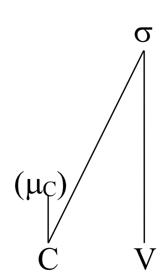
(12) The moraic-onset syllabification of initial geminates



Therefore, while this appears to be the most intuitive way to syllabify initial moraic geminates, the theory of moraic phonology does not approve such structure, given the irrelevance of onsets in weight related phenomena.

As an alternative, Hayes (1989) has suggested that an initial geminate could be doubly linked to the onset of the first syllable and to a word-initial mora which is extraprosodically licensed. This way, initial geminates can be equally moraic in compliance with the weightless status of onsets.

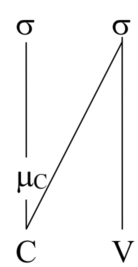
(13) The extrasyllabic syllabification of initial geminates



Nevertheless, this representation can be criticized for the obvious reason that an extrasyllabic mora would be unlicensed by an immediate prosodic node. This means that geminates would be rendered uninterpretable, giving rise to phonetically short consonants (Hume et al., 1997). This is akin to the nonrealization of the underlying weight of the geminate.

For our data, however, we adopt a syllabification that was also suggested by Hayes (1989). It was proposed that initial geminates could be alternatively syllabified in terms of double association to two syllable nodes, just as medial geminates are. An initial geminate in this case will be both the nucleus of the first syllable, projected by its underlying mora, and the onset of the following syllable.

(14) The syllabic-consonant syllabification of initial geminates



This way of syllabifying initial geminates in MA is supported by two facts. First, syllabic consonants are possible in MA. They are normally created to get rid of complex syllable margins. In the case of geminates, however, they are created mainly to bear the mora of the geminate. Second, words made up of a geminate and a short vowel are attested, which hints at their disyllabic/bimoraic nature since they satisfy the minimal word condition of MA.

Initial geminates in MA can be of two types: underlying or derived. Under the assumption that all types of geminates share the same phonological representation, the syllabification process is expected to treat each type equally. Some examples of initial geminates in MA are repeated below for convenient reference:

(15) Initial geminates in MA

a. Underlying initial geminates

ddi	‘to take’
mm^wi	‘my mother’
bb^wa	‘my father’
lla	‘no’
ssi	‘Mr.’
mmaq	‘the sound of slapping’
ṭtaq	‘the sound of falling’
bbaq	‘the sound of exploding’
ṛṛaq	‘the sound of hammering’

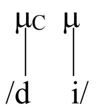
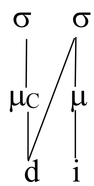
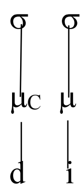
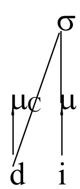
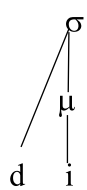
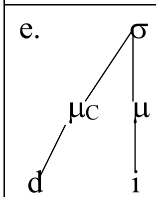
b. Derived initial geminates

ssuq	‘the market’
ṣṣif	‘the summer’
llil	‘the night’
ʃʃəṛʒəm	‘the window’
ttuma	‘the garlic’

To reiterate, standard moraic theory handles the syllabification of intervocalic geminates perfectly. However, the theory faces a great deal of awkwardness in relation to peripheral geminates. Our proposal to syllabify initial geminates as in (14) helps to establish some equivalence between initial geminates and intervocalic ones. OT provides the right tools to achieve this goal. Looking at the representations alone may not be enough to demonstrate this fact. The main claim we make is that initial geminates in MA are equally heterosyllabic and that their heterosyllabicity follows from the same constraint hierarchy that gives rise to heterosyllabic medial geminates.

Therefore, since initial geminates lack the option of parsing their moras in a coda position, their moras get parsed as heads of syllabic consonants for the sake of maintaining their underlying weight specification. The constraint hierarchy ONSET, IDENT-W >> $*\mu$/ONSET >> $*\mu_h$/C >> $*\mu$/Coda predicts this structure by ranking the constraint against syllabic consonants as the second lowest constraint in the hierarchy, making it the second least marked structure to opt for when a coda position is unavailable. The tableau below demonstrates how this structure gets chosen as the optimal one:

(16) The syllabification of initial geminates in MA

	ONSET	IDENT-W	*μ/ONSET	*μ_h/C	*μ/CODA
☞a. 				*	
b. 	*!				
c. 		*!			
d. 		*!			
e. 			*!		

This tableau illustrates how the input *ddi* 'to take' is syllabified. Candidate (16b) represents a suboptimal syllabification that preserves the weight of the geminate but fails to supply an onset for the subsequent onsetless syllable. Candidate (16c) and (16d) are both ruled out by IDENT-W due to their inability to parse or preserve the mora of the initial geminate, respectively. Candidate (16e) ties with the optimal candidate in satisfying the requirements of ONSET and IDENT-W. However, such a candidate contains a moraic onset, which happens to be the most marked option of moraic consonants. The optimum succeeds in preserving the weight of the geminates by means of projecting a syllabic consonant, which represents the least possible marked consonantal moraic structure for initial geminates.

Thus far, the two main assumptions of moraic theory that have been underlying our analysis of initial geminate syllabification in MA are as follows: (i) initial geminates are underlyingly moraic (ii) onsets are always nonmoraic. While they seem to resonate well with our data, it's worth noting that these two assumptions have been challenged by data from other languages.

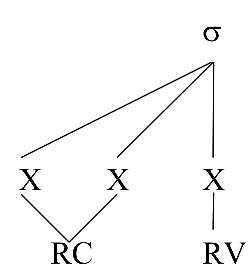
The first assumption is defied by Leti (Hume et al. 1997; Muller, 2001), where geminates do not participate in any weight sensitive phenomenon. To explain, Leti requires content words to be minimally bimoraic. If initial geminates were indeed underlyingly moraic, words of the shape GV would have been possible in the language. Yet, no words of this nature are attested. Instead, initial geminates in Leti occur only in words that are bimoraic by virtue of containing two vowels or a diphthong, as shown by the data below, taken from Muller (2001).

(17) Initial geminates in Leti:

ppikan	‘plate’
ttui	‘genre of literature’
kkusal	‘to be small’
ssoran	‘cough’
mmanan	‘food’
llai	‘beach’
rraa	‘again’

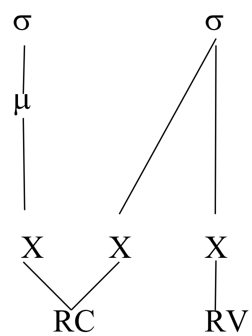
Stress provides further evidence against the moraic nature of initial geminates in Leti. While syllables that are heavy by dint of containing a long vowel do attract stress, syllables consisting of a geminate plus a short vowel fail to do so, behaving as monomoraic syllables. With these facts in mind, it was concluded that geminates in Leti are nonmoraic. Rather, geminates in Leti are assumed to be represented in terms of melodic multiple association to two timing slots. In this concern, the syllabification suggested by Hume et al. (1997) for Leti initial geminates is for the two slots constituting the geminate to occupy the onset position.

(18) Initial geminates as complex onsets



If this segmental view of initial geminates in Leti is generalized to cross-linguistically represent all cases of initial geminates, MA initial geminates would be subject to the stipulation of the ranking of *COMPLEX over *μ_h/C, producing a syllabic consonant as a result. However, it is not plausible to posit different representations for medial and initial geminates, especially that initial geminates have been reported to be moraic elsewhere.

(19) Initial geminates as heterosyllabic segments



In this concern, it has been shown that the facts about Leti initial geminates are inconclusive (Davis, 1999b; 2011; Curtis, 2003). In his 1999 squib, Davis argued that it is not even clear if Leti contains geminates in the first place. The argument goes as follows: since Leti places no featural restrictions on the type of consonant clusters it could have, it is very likely that those geminates are just sequences of identical consonants.

(20) Initial clusters in Leti:

ptuna	‘star’
pninu	‘fool’
vroan	‘axe’
sraki	‘gong’
vlira	‘weaving rod’
snuran	‘thread’
mniniru	‘soft’

The fact that Leti geminates may just be a sequence of identical consonants is also supported by the observation that geminates in Leti do not occur anywhere else besides word-initially, which is the case with consonant clusters as well. According to Davis (1999b, 2003, 2011), the inconclusive nature of Leti geminates becomes more evident thanks to the different patterning of Trukese initial geminates. In Trukese, not only do geminates occur both initially and medially, but they also occur to the total exclusion of clusters of unlike consonants.

Most importantly, Trukese initial geminates stand the test of the bimoraic restriction on content words. This is evidenced by the fact that words of the shape GV do exist in the language, suggesting that initial geminates are moraic. Examples of such words include: [čča] 'blood' and [ssɔ] 'thwart'. Moreover, it has been shown that Trukese has a general rule that shortens word-final long vowels. Interestingly, this rule is blocked from applying when the input consists of a singleton and a long vowel due to the bimoraic minimality condition. However, the shortening rule does apply to GVV inputs since the output would still be bimoraic thanks to the moraic nature of the initial geminate.

Suggestions as to how to syllabify initial geminates in Trukese remain controversial. The syllabic consonant syllabification (i.e. linking the mora to a syllable of its own) suggested for MA would be an oddity for the phonology of Trukese. The extraprosodity syllabification (i.e. linking the mora directly to the foot or word node) seems to raise questions about the phonetic realization of the relevant geminates. The incorporation of the whole geminate into the onset of the syllable would give rise to moraic onsets, challenging the strong cross-linguistic observation that onsets are never moraic. Our analysis captures this universal tendency by the following ranking: *μ/ONSET >> *μ_h/C >> *μ/CODA. This means that this ranking is not subject to typological permutations.

However, the assumption that onsets never carry weight has been contested. It has been shown that, in some languages, onsets do contribute to the weight of their syllables (Hajek and Goedemans, 2003; Topintzi, 2006, 2008). For example, in Pattani Malay, initial geminates clearly influence the placement of stress, such that the syllables in which they work as onsets appear to be heavier than syllables with singletons as onsets. This is substantiated by the following examples:

(21) Stress and initial geminates in Pattani Malay (Hajek and Goedemans, 2003):

sɨˌdaːˈduː		ˈddaːˌduː	'police'
pɨˌmaːˈtɔː		ˈmmaːˌtɔː	'jewelry'
ˌɟaːˈlɛː	'path'	ˈɟɟaːˌlɛː	'to walk'
ˌbuːˈwɔh	'fruit'	ˈbbuːˌwɔh	'to bear fruit'

In Pattani Malay, the default position of the primary stress is the final syllable of the word. However, when a word begins with a geminate, primary stress falls on the first syllable due to the Weight-to-Stress Principle (WSP). This clearly suggests that the onset of the first syllable is moraic by virtue of being a geminate. Our constraints can predict such a language if the ranking *μ/ONSET >> *μ_h/C >> *μ/CODA is relaxed to allow for typological permutations. This way, the constraint against moraic onsets will have to be ranked below *μ_h/C, permitting initial geminates to be fully contained in the onset. However, Davis (1999a) maintains that cases where onsets seem to affect stress placement are rare. Therefore, it is better to sustain the generalization about the nonmoraic status of onsets. Those rare cases like Pattani Malay can be interpreted as cases of stress shift.

3.3. Word-final geminates

While peripheral geminates are generally less common than medial ones, inside the peripheral categorization, final geminates have been found to be more common than initial geminates (Thurgood, 1993). When it comes to their syllabification, word final geminates are automatically syllabified as codas. Therefore, since the coda is a typical weight bearer, a final geminate's mora would fit perfectly in this position. In MA, final geminates abound, such that they occur in many verbs, nouns and adjectives. Some examples are given below:

(22) Final geminates in MA

sədd	'to close'
ħəll	'to open'
fəʃʃ	'to deflate'
dəmm	'blood'
səmm	'poison'
ʒədd	'grandfather'
xəff	'lighter'
rəqq	'thinner'

Since the input of all the above-cited examples consist of consonants only, a schwa is needed to aid in the parsing process. This, of course, is taken care of by the ranking of PARSE-Seg >> DEP-ə. With the exception of the constraint ONSET, it will be shown that the syllabification of final geminates in MA is subject to the stipulation of the same constraint hierarchy responsible for the parsing of medial and initial geminates: IDENT-W >> *μ/ONSET >> *μ_h/C >> *μ/CODA. This way, the primary demand would be to preserve the input specification of the final geminate. Being on the right periphery, the geminate finds itself in a coda position, which is no strange to consonantal moraic structures. The only difference is that the mora, in this case, is base-generated, instead of being phonologically coerced.

(23) The syllabification of final geminates in MA

μ_c /s d /	IDENT-W	*μ/ONSET	*μ_h/C	*μ/CODA
☞a. σ μ_c s ə d				*
c. σ μ s ə d	*!			*

Nevertheless, being on the edge, final geminates pose other challenges for standard moraic theory (SMT). The first challenge to consider recurs across many languages, including MA. Especially, SMT fails to distinguish between CVC and CVG syllables in languages where codas are assigned moras by Weight-by-Position.

(24) The weight equivalence of CVC and CVG syllables

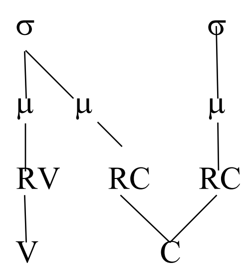
One crucial merit of our model is that it offers a way to distinguish between word final CVC syllables and CVG syllables, even when they have equal weight configurations. This is achieved by referring to the two-root node aspect of geminates, which phonologically embodies the longer phonetic durations of geminates. Said differently, in our model, the moras of geminates can be differentiated from purely prosodic moras associated with single consonantal root nodes. Such a distinction is clearly not possible in a purely moraic model. As a result, the competing candidates in (23) can be distinguished by our constraint hierarchy. Candidate (23b) is found to violate IDENT-W since its mora is not the same mora of the input. Rather, it is a coerced type of mora that follows from the independent interaction between *APPENDIX and *μ/CODA. The winner, candidate (23a), satisfies IDENT-W by maintaining the mora of its input, which is underlyingly associated to two root nodes.

Closely relatedly, the second challenge that SMT faces in relation to final geminates is rather specific to MA, where schwa is reported to be nonmoraic. This means that a monosyllabic word of the shape CəG runs the risk of being judged as subminimal. Yet, we know that MA imposes a restriction on its words to be minimally bimoraic or disyllabic. This is substantiated by the absence of monosyllabic words of the shape CV and CəC. The question now is how do CəG words satisfy the minimality condition of MA?

To answer this question, a brief recap is required. In the course of this chapter, the syllabification of geminates has been argued to follow exclusively from their underlying moraic structure. The two-root node structure, however, has been restricted and rendered irrelevant for the syllabification process. In particular, we maintained that the two root nodes that constitute geminates in our model do not fully pattern with sequences of independent root nodes (i.e. clusters), such that they do not create complex syllable margins. For the sake of consistency, we intend to endure all the challenges and maintain this line of argumentation.

So far, restricting the bi-positionality of geminates in favor of their moraicity has worked well for medial and initial geminates since word minimality has not been an issue for words with those types of geminates. However, in order to explain the bimoraic status of words where final geminates are the only moraic segments (i.e. CəG), one might jump to the conclusion that these words actually pattern with CəC.C words.

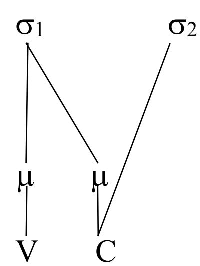
(25) Final geminates as true bi-positional segments



Such a conclusion would mean that the bi-positionality of geminates should be the deciding factor in the syllabification of geminates in MA, thus weakening the moraic conception of MA geminates. In other words, if we allow the root nodes to have any role in the syllabification of final geminates, the argument that geminates are underlyingly moraic will be at stake.

One way out of this dilemma is to argue that, in addition to serving as a coda, a potentially contained geminate, be it final or internal, extends to an ensuing empty-headed syllable, created mainly to avoid contained geminates.[12] Here, a contained geminate refers to a geminate that is associated to a single syllabic node. The derived structure is as follows:

(26) Final-geminate hetrosyllabicity



This way, CəG words can satisfy the word minimality condition under syllabic analysis. This also means that final geminates in MA get to be heterosyllabic, though for some different reasons.[13] So how do final geminates get to become heterosyllabic? To account for this in terms of constraint interaction, we postulate the following constraints:

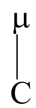
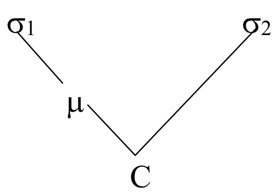
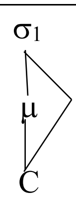
a. ***CONTAINED-GEM**: geminates should not be fully contained in a coda

b. ***EMPTY-headedσ**: syllables without a head nucleus are prohibited.

[12] Many thanks to Joe Pater for pointing this out to me.
[13] We believe that the heterosyllabicity of geminates is deeply rooted into their longer phonetic duration. In this sense, geminates are essentially conceived of as long segments that are entitled to filling in two consecutive syllabic positions, reflecting their intrinsic increased phonetic length.

The markedness constraint *CONTAINED-GEM prohibits geminates that are linked to a single syllabic node. The constraint *EMPTY-headed-σ, on the other hand, disallows syllables without nuclei. An interaction between these two constraints, in which *CONTIANED-GEM dominates *EMPTY-headed-σ, makes the violation of the latter less serious, giving rise to empty-headed syllables. This interaction is illustrated by the following tableau:

(27) Final geminates heterosyllabicity

	*CONTAINED-GEM	*EMPTY-headed-σ
☞a. 		*
b. 	*!	

For a while, it may look that this structure is at odds with the phonology of MA. In order to refute this claim, the next section will show that there is a couple of phonological processes that seem to directly interact with the structure of empty-headed syllables. Specifically, we will look at one case of epenthesis and another case of degemination, which, we believe, are triggered by the empty-headed syllable structure.

4. Empty-headed syllable avoidance strategies

4.1. [i]-Epenthesis

[i]-epenthesis in MA occurs exclusively in base-generated final-geminated verbs when inflected for the past tense. Hence, this type of epenthesis represents a very restricted case of vowel epenthesis, which, to the best of our knowledge, does not occur anywhere else in the language. The following paradigm for the verb *sədd* 'to close' illustrates this case of epenthesis:

(28)

a.	səd.d**i**t	'I closed'	*səddt	*sdədt
	səd.d**i**.na	'we closed'	*səddna	*sdədna
	səd.d**i**.ti	'you closed'	*səddti	*sdədti
	səd.d**i**.tu	'you (pl) closed'	*səddtu	*sdədtu

b.	səd.d	'he closed'
	səd.dat	'she closed'
	səd.du	'they closed'

The restricted status of [i]-epenthesis to verbs ending in geminates becomes more conspicuous when we consider, for comparison, triliteral verbs with the same inflectional properties. Consider the following portion from paradigm of the verb *ḍrəb* 'to hit' for illustration:

(29)

ḍrəbt	‘I hit’	*ḍrəb**i**t
ḍrəbti	‘you hit’	*ḍrəb**i**ti
ḍrəbtu	‘you (pl) hit’	*ḍrəb**i**tu

The lack of [i]-epenthesis in this type of verbs proves that the vowel [i] has no morphological function whatsoever. In other words, since we only find the epenthetic [i] with geminated verbs, it is obviously not part of the past tense inflectional morphology. This means that a purely phonological analysis is desired.

Our account for this case of epenthesis rests on the assumption that final geminates produce empty-headed syllables to avoid being fully contained in the coda (i.e.*CONTAINED-GEM >> *EMPTY-σ). Therefore, the role of the epenthesized [i] is to improve syllable structure by providing a nucleus for the created empty-headed syllables. Such an analysis is motivated by the fact that [i]-epenthesis is specific to geminates and excludes consonant clusters, implying that geminates yield a different syllable structure (i.e. empty syllables) from consonant clusters. We believe that this state of affairs explains why [i]-epenthesis is activated only in the environment of geminates. To prove this point, let’s consider the following nouns:

(30)

fəlfla	‘pepper’	*fəlfila
sənsla	‘necklace’	*sənsila
bərdʕa	‘a donkey’s saddle’	*bərdiʕa
sərtla	‘bracelet’	*sərtila

These nouns seem to have the exact same syllable structure as the input forms for the [i]-epenthesis. However, their inability to trigger [i]-epenthesis indicates that, in fact, the two structures are different, such that geminates create empty syllables, which happens to be the only phonological context entitled to triggering [i]-epenthesis.

By now, it has become a known fact that the default epenthetic vowel in MA is the schwa. Also, it goes without saying that schwa epenthesis is more general than [i]-epenthesis, in that it applies to break up consonant clusters regardless of the grammatical category they belong to or the inflectional process they are undergoing. We conjecture that the defaultness of schwa epenthesis follows from the following constraint hierarchy: DEP-V >> DEP-ə.

Given this constraint hierarchy, schwa is expected to always surface as the only epenthetic vowel in MA, either to break up consonant clusters (i.e. PARSE-C >> DEP-V >> DEP-ə) or to fill in empty-headed syllables (i.e. *EMPTY-σ >> DEP-i >> DEP-ə).

(31)

/sədd+ti/	*CON-GEM	*EMPTYσ	DEP-i	DEP-ə
a. səd.di.ti			*!	*
b. səd.d.ti		*!		*
c. sədd.ti	*!			*
☹ d. səd.də.ti				**

The question to ask now is how does the vowel [i] get to replace the schwa in filling the empty-headed syllables of final-geminated verbs? A cursory look at the verb forms of the past tense paradigm gives us the answer right away. It is observed that inserting a schwa would put the latter in an open position (i.e. Cə). This clearly goes against the established fact that schwa in MA cannot be the head of an open syllable thanks to its nonmoraic nature. This restriction is formally expressed by the markedness constraint *μ/ə. Being ranked above DEP-ə, this constraint blocks schwa epenthesis, clearing the way for [i]-epenthesis to take over.

(32)

/sədd+ti/	*μ/ə	*CON-GEM	*EMPTYσ	DEP-i	DEP-ə
☞a. səd.di.ti				*	*
b. səd.d.ti			*!		*
c. sədd.ti		*!			*
d. səd.də.ti	*!				**

However, inserting a schwa seems to be under the realm of possibility in the case of the first person singular form since a schwa-headed open syllable is not at stake (i.e. *səd.dət 'I closed'). This means that [i]-epenthesis is not justified in this case. We believe that given the fact that the majority of the forms in the paradigm require [i]-epenthesis, the 1st person singular form follows suit for the sake of paradigm uniformity (McCarthy, 2005)[14]. Paradigm uniformity demands tend to compel members of a given morphological paradigm to maintain some degree of phonological similarity. To represent this idea formally, we posit a constraint that makes the following demand:

(33) **Paradigm-Uniformity** (PRD-Uni):

Paradigm members should be phonologically similar.

[14] I would like to thank John McCarthy for turning my attention to this solution.

(34) PRD-Uniformity, *EMPTY-headedσ >> DEP-i >> DEP-ə

/sədd+t/	PRD-Uni	*μ/ə	*CON-GEM	*EMPTYσ	DEP-i	DEP-ə
☞ a. səd.dit					*	*
b. səd.d.t				*!		*
c. sədd.t			*!			*
d. səd.dət	*!					**

A second look at the paradigm shows that [i]-epenthesis does not affect all members of the paradigm. For example, the third person feminine singular and the third person plural forms do not require [i]-epenthesis because their suffixes begin with a vowel. However, the third person masculine singular form is problematic for our analysis since it does not allow [i]-epenthesis despite its supposed empty-headed syllable. In reaction to this, we assume that [i]-epenthesis is blocked by a high-ranking constraint that prohibits epenthesis in a word final position, written as DEP_#.

(35)

/sədd/	DEP_#	*μ/ə	*CON-GEM	*EMPTYσ	DEP-i	DEP-ə
☞a. səd.d				*		*
b. səd.di	*!				*	*
c. sədd			*!			*
d. səd.də	*!	*				**

The other case where [i]-epenthesis is blocked from filling in empty-headed syllables is found in inflected causative verbs whose inflectional suffixes begin with a vowel.

(36)

kət.təb	‘to make write’	kət.t.bu	‘he made it write’
fəz.zəg	‘to wet’	fəz.z.gu	‘he wet it’
nəʃ.ʃəf	‘to dry’	nəʃ.ʃ.fu	‘he dried it’
dəw.wəb	‘to melt’	dəw.w.bu	‘he melt it’

In the infinitive form of causatives, medial geminates behave as a sequence of two consonants thanks to their two-root node nature, triggering the epenthesis of two schwas together with the first and final consonants. However, when the causative is followed by a vowel, the last consonant gets syllabified as an onset, and hence becomes unavailable for schwa epenthesis. As a result, in addition to being the coda of the first syllable, the medial geminate of the causative creates an empty-headed syllable structure to obviate the less favored contained geminate structure.

We have seen that, unlike word-final empty-headed syllables, word-internal empty syllables tend to trigger [i]-epenthesis. However, this is not born out by the internal geminates of the causative verbs. To explain this, we should try to identify some differences between the two cases. One difference is that the geminates in the verbs above are underlying, while those in causative verbs are derived. Yet, this does not seem to be a convincing reason as to why they behave differently with regards to [i]-epenthesis since all geminates are assumed to have the same phonological make-up. Another difference is that when [i]-epenthesis applies to final geminated verbs, it still does not apply inside the root material, whereas in the case of the inflected causatives in (36), [i]-epenthesis would have to disrupt the contiguity of the root material.

(37) **ROOT-Contiguity**:

The portion of the root in the output standing in correspondence forms a contiguous string (no intrusion).

(38)

kəttbu	ROOT-Contiguity	*CON-GEM	*EMPTY-headedσ	DEP-i
☞ a. kət.t.bu			*	
b. kət.ti.bu	*!			*
c. kətt.bu		*!		

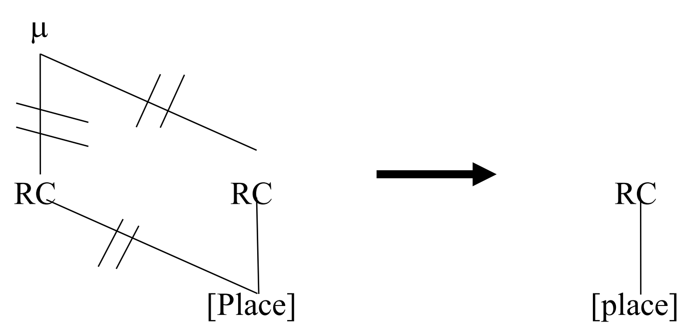
The disruption of root contiguity in MA is not uncommon. Arabic, in general, is known to have a nonconcatenative morphology, whereby consonantal roots are interleaved with vowel patterns to form words. We believe that disrupting root contiguity in MA is reserved to morphologically motivated vowels. The epenthetic [i] put in use here to fix empty-headed syllables is purely phonological, and hence does not fit the description. Schwa, on the other hand, does not disrupt root contiguity thanks to its unique phonological nature. Unlike full vowels, schwa has been found to be nonmoraic.

This disparity between schwa and the epenthetic full vowel [i] can be best described in terms of the distinction that Hall (2006) makes between epenthetic vowels and intrusive vowels. According to Hall, intrusive vowels are not fully-fledged phonological units as they merely involve the "retiming of existing articulatory gestures to produce a vowel-like transition between consonants". However, while schwa does exhibit distinct qualities from other vowels in the sound system of MA, it does not possess all the properties of intrusive vowels that Hall (2006) has identified, namely the properties of optionality and high variability. With this in mind, we can safely claim that schwa does not disrupt root contiguity, not necessarily because it is an intrusive vowel per se, but simply due to its distinct phonological nature.

4.2. Degemination

Degemination is a phonological process that simplifies geminate segments into singletons. Phonetically, this means cutting the closure duration of the relevant geminate by half. Phonologically, degemination changes the underlyingly structure of geminates, turning them into simplex structures. In our case, a degeminated geminate is believed to give up its underlying moraic and bi-positional structure. In this regard, by adhering to the weight-to-length correlation principle, the deletion (i.e. dissociation) of one root node should automatically lead to the erasure of the underlying mora.

(39) Degemination in our model



Degemination represents another strategy to avert the empty-headed syllable structure since, unlike geminates, singletons can be associated to only one single syllabic node at a time.

In MA, the passive form of verbs is basically derived via prefixing the geminated affix /tt-/ to the base verb. This is put into evidence by the following forms:

(40)

ktəb	ttəktəb	'write'
mnəʕ	ttəmnəʕ	'prevent'
ʃṛəb	ttəʃṛəb	'drink'
ḍrəb	ttəḍrəb	'hit'
xdəm	ttəxdəm	'work'
nʃəṛ	ttənʃəṛ	'spread'
ħfəḍ	ttəħfəḍ	'recite'
ʃra	ttəʃra	'buy'
kra	ttəkra	'rent'
nsa	ttənsa	'forget'
bla	ttəbla	'get addicted'
xda	ttəxda	'take'
kla	ttəkla	'eat'

As these items demonstrate, in the case of tri-consonantal verbs and final weak verbs, the passive prefix maintains its geminated form. However, the passive prefix shortens (i.e. degeminates) and surfaces as [t-] with the following base verbs:

(41)

sədd	tsədd	'close'
ħəll	tħəll	'open'
ʃədd	tʃədd	'catch'
ʃəqq	tʃəqq	'break'

kəttəb	tkəttəb	'write'
ɣəṭṭa	tɣəṭṭa	'cover'
ʃəṛṛəb	tʃəṛṛəb	'drink'
nəʃʃəf	tnəʃʃəf	'dry'
laħ	tlaħ	'throw away'
ʃaf	tʃaf	'see'
gal	tgal	'say'
dar	tdar	'make'
zad	tzad	'expand'
xad	txad	'take'
kal	tkal	'eat'

The base verbs that trigger the degemination of the passive affix include: (i) underlyingly final-geminated verbs (ii) morphologically derived medial-geminated verbs (iii) and tri-segmental medial weak verbs. The latter class eliminates the possible hypothesis of avoiding two geminate consonants in the same domain. So, what do these base verbs have in common that makes them trigger the degemination of the passive prefix?

We argue that the defining context of this allomorphy has to do with the phonological status of the consonant that follows the passive affix. It is believed that the degemination of the passive affix takes place when the latter is followed by a consonant in the onset position (i.e. GCV), like in (41), preventing empty-headed syllables from being formed. The question now is: why is [i]-epenthesis not used here as well to fix the empty-headed syllables emerging from the passive affix?

The answer seems to depend on the morphemic status of the geminate involved in the empty-headed syllable structure, such that different strategies are allocated to stem geminates versus affix geminates. We entertain the idea that [i]-epenthesis is reserved for stem geminates while degemination is assigned to affix geminates. This could be the result of ranking the faithfulness constraint against the identity change of affix geminates, IDENT-W-Affix, below DEP-i. Therefore, having IDENT-W-Affix and DEP-i dominated by *EMPTY-σ in that way, degemination will take precedence over [i]-epenthesis in the presence of affix geminates. However, the faithfulness constraint preserving the weight specification of stem geminates should be ranked above *EMPTY-σ, preventing stem geminates from shortening and paving the way to [i]-epenthesis instead.

(42) Positional faithfulness

/tt+sədd/	IDENT-W-STEM	*EMPTY-σ	IDENT-W-AFFIX
☞ a. t.səd.d		*	*
b. t.səd	*!		*
c. t.t.səd.d		**!	

This tableau shows that there is a positional faithfulness asymmetry between stem geminates and affix geminates. Only affix geminates may get degeminated under this constraint ranking. The degemination of the stem geminate, as illustrated by candidate (42b), is punished by the undominated faithfulness constraint IDENT-W-STEM. The non-degemination of both the stem geminate and the affix geminate, like in (42c), leads to multiple violations of *EMPTY-headedσ. The optimal candidate in (42a) incurs a single violation of *EMPTY-headedσ, hence qualifies as the most harmonic form. The following tableau zooms out on the issue of empty-headed syllables by showing all the relevant constraints:

(43)

/tt+sədd/	*CON-GEM	DEP_#	IDENT-W-STEM	*EMPTY-headedσ	DEP-i	IDENT-W-AFFIX
☞ a. t.səd.d				*		*
b. t.səd			*!			*
c. t.t.səd.d				**!		
d. tt.sədd	**!					
f. t.ti.səd.d				*	*!	
g. t.səd.di		*!			*	*

In this tableau, we show how [i]-epenthesis and degemination are distributed between the geminate of the passive affix and that of the base verb. We also show how [i]-epenthesis is blocked in a word final position, leaving the empty-headed syllable hanging in limbo.

However, if the consonant that comes after the passive affix is a syllabic one (i.e. G.C.CV), it then provides the possibility of re-syllabification by working together with the passive affix to trigger schwa epenthesis. As a result, degemination would be unnecessary in such a case. This is illustrated by the tableau in (44):

(44)

/tt+xdəm/	PARSE-Seg	*EMPTY-headedσ	IDENT-W-AFFIX	DEP-ə
☞ a. t.təx.dəm				**
b. təx.dəm			*!	**

The optimal form in (44a) satisfies *EMPTY-headedσ without having to violate IDENT-W-AFFIX. Candidate (44b), however, violates IDENT-W-AFFIX gratuitously, and hence loses in the competition.

5. Conclusion

In this chapter, we have provided a constraint-based account of geminate syllabification in MA. The chapter has kicked off with solving the redundancy problem that comes with combining a moraic level and a segmental level to represent geminates in MA. The solution consisted in restricting the segmental aspect of the model. In particular, it was argued that, despite being bi-positional, geminates do not constitute complex syllable margins, and thus have to be distinguished from consonant clusters which are characterized by having independent place nodes. As a result, it was maintained that the syllabification of geminates in MA relies exclusively on their moraic structure.

On the basis of these assumptions, we have proceeded to show that geminates in MA are characteristically heterosyllabic, not only intervocalically but also peripherally. It has been indicated that a word medial geminate is subject to two main syllabification demands: (i) supply an onset and (ii) preserve the mora somewhere other than the onset. This way, an intervocalic geminate becomes associated to two syllabic nodes: the coda of the preceding syllable and the onset of the ensuing one. Similarly, initial geminates have also been analyzed as heterosyllabic. However, an initial geminate has been shown to parse its underlying mora as the head of its own syllable, which is also motivated by the demand of avoiding moraic onsets. This was formally accounted for by means of the following constraint hierarchy: ONSET, IDENT-W >> $*\mu$/ONSET >> $*\mu_h$/C >> $*\mu$/CODA.

As for final geminates, it was shown that their syllabification is subject to the same constraint hierarchy outlined above, except that a final geminate has no chance of becoming an onset, hence the constraint ONSET becomes irrelevant in this context. Therefore, it has been suggested that being in a coda position is conducive to preserving the underlying weight of final geminates. However, the twist was that final geminates in MA also tend to be heterosyllabic, though not for the same reasons as medial and initial geminates. In this regard, we have argued that a final geminate gets heterosyllabic mainly to avoid being contained in a single syllabic node, creating a structure of an empty-headed syllable as a result. This was shown to follow from the ranking of the markedness constraint against contained geminates, *CON-GEM, over the markedness constraint against empty-headed syllables, *EMPTY-headedσ. We elaborated on this by showing that MA involves two phonological processes that seem to prove the existence of empty-headed syllables in the context of final geminates, namely a process of [i]-epenthesis and a process of degemination. It was demonstrated that [i]-epenthesis applies after stem geminates, while degemination affects affix geminates.

CHAPTER SIX

ON THE INTEGRITY OF GEMINATES IN MOROCCAN ARABIC

1. Introduction

This chapter investigates the phonological behavior of geminate consonants in Moroccan Arabic. In particular, we focus on the issue of geminate integrity in the context of schwa epenthesis and word formation. We show that, despite the many apparent exceptions, the variable nature of geminate integrity in MA can be successfully accounted for along the lines of the Geminate Law (Benhallam, 1980) if the latter is reinterpreted in the Optimality Theory framework. In this regard, this chapter promises the following contributions: (i) It provides a unified analysis of geminate integrity in MA; (ii) it accounts for the variability of geminate integrity through constraint interactions á la Optimality Theory; and (iii) it reconciles the exceptional patterns of geminate integrity with the regular ones.

In MA, geminates are characterized by resisting rules of epenthesis (i.e. $*C_xəC_x$) which normally affect clusters of unlike consonants (i.e. $C_xəC_y$) (Kenstowicz and Pyle, 1973; Guerssel, 1977; Benhallam, 1980). This generalization is commonly referred to in the literature as geminate 'integrity' or 'inseparability'. However, as far as full vowels are concerned, the integrity of geminates in MA has been found to be regularly subject to breaking (i.e. C_xVC_x).

In this chapter, we will examine the issue of geminate integrity in MA in connection with schwa and full vowel placement, using the constraint-based framework of Optimality Theory (Prince and Smolensky, 1993/2004). More specifically, we will explore the inconsistent behavior displayed by geminates vis-á-vis normal consonant clusters. Geminates in MA have been shown to be typically immune to vocalic splitting by means of schwa epenthesis (e.g. ħ**ə**ll 'to open' as opposed to zr**ə**b 'to hurry'). However, it has been reported that, occasionally, geminates may succumb to schwa epenthesis (e.g. də**ff**a vs. d**f**ə**f** 'doors'). In this regard, we will attempt to discern the possible constraint interactions that make geminates behave the way they do with respect to schwa epenthesis. In connection to this, we also intend to address the behavior of geminates in relation to full vowels, which have been found to split geminates and clusters alike without any notable irregularities (e.g. sə**dd** vs. məs**d**u**d** 'closed').

The structure of this chapter is mapped out as follows. Section 2 provides a brief background on the issue of geminate integrity in the non-OT literature. Section 3 characterizes the phonology of schwa in MA and introduces our proposed analysis to capture the effect of the GL in the framework of OT. Section 4 examines some persisting issues to our account and offers some solutions. Section 7 concludes the chapter.

2. Schwa epenthesis in MA

Understanding the behavior of schwa is very central to our investigation of geminate behavior in Moroccan Arabic. Therefore, before we delve into the intricacies of our account of geminate integrity and its interaction with schwa epenthesis, we will first delineate the general role of schwa in the phonology of MA.

The nature and behavior of schwa are among the most studied phonological aspects of MA (see Benhallam, 1980, 1989/1990, 1991; Benkirane, 1982; Benkaddour, 1982; Al Ghadi, 1990/2014, 1994; Boudlal, 2001, 2006/2007; Bensoukas and Boudlal, 2012a-b to name but a few). In this body of research within the generative paradigm, it has been mostly claimed that schwa in MA is a purely phonetic vowel, epenthesized in order to break up impermissible sequences of consonant clusters.[15] The epenthetic status of schwa can be determined through comparing morphologically related words, whereby shifting between categories causes schwa to disappear or otherwise change position. To illustrate, we provide the following examples:

(1) The epenthetic nature of schwa in MA

a. Disappearing schwas

ktəb	‘to write’	ktuba	‘books’
ktəf	‘a shoulder’	ktaf	‘shoulders’
sləx	‘slay’	slix	‘slaying’
təṛʒəm	‘to translate’	ṭəṛʒmu	‘they translated’
ʃəṛʒəm	‘a window’	ʃṛaʒəm	‘windows’

[15] Benkirane (1982) and Benkaddour (1982) maintain that schwa is underlying in nouns and epenthetic in verbs only. These works adhere to a diachronic account of schwa in nouns whereby vowel reduction of Classical Arabic (CA) full vowels is what gave rise to underlying schwas in MA.

b. Moving schwas

ʃṛəb	'to drink'	ʃəṛbat	'she drunk'
ɦṛəb	'to escape'	ɦəṛbu	'they escaped'
wdən	'an ear'	wədnin	'ears'
ṣbəʕ	'a finger'	ṣəbʕu	'his finger'
srəħ	'to shepherd'	sərħa	'shepherding'

In an OT analysis, the process of schwa epenthesis outlined above is captured by the interaction of two basic constraints: PARSE-Segment and DEP-ə. The former requires all consonants to be parsed into syllables, whereas the latter demands that every schwa in the output have a correspondent in the input, prohibiting the insertion of schwa. Being the dominant constraint in this case, PARSE-Seg triggers the epenthesis of schwa at the expense of DEP-ə.[16]

(2) PARSE-Seg >> DEP-ə

/CCC/	PARSE-Seg	DEP-ə
☞a- CCəC		*
b- CCC	***!	

This tableau evaluates candidates for a tri-consonantal input, where the faithful candidate CCC loses for having no syllable structure. Candidate (2a) becomes the winner by satisfying PARSE-Seg through schwa epenthesis. Note that at this stage the justification for the position of schwa is not discussed yet. Any candidate with an additional schwa (e.g. CəCəC) not only would

[16] Schwa is the default epenthetic vowel in MA. This can be explained by the ranking: Dep-V >> Dep-ə (where V stands for any full vowel). Schwa epenthesis in MA has been described as an oddity of the language, given that other Arabic dialects lack schwa epenthesis; instead they rely on full vowels for epenthesis (see Bensoukas and Boudlal (2012a-b) and references therein for more on this). Bensoukas and Boudlal (2012a-b) describe this oddity as an Amazigh substratum in MA since schwa epenthesis in Amazigh is more fundamental and occurs across many Amazigh dialects.

incur multiple and gratuitous violations of DEP-ə but would also violate the undominated constraint against schwa in open syllables, i.e. *μ/ə.

(3) ***μ/ə:** moraic schwas are prohibited (Bensoukas and Boudlal, 2012a:23)

In MA, complex syllable margins are not permitted in MA, in accordance with the stipulation of *COMPLEX. As a result, in order for the latter to be satisfied, a syllabic consonant is created in violation of the constraint against moraic head consonants, i.e. $*\mu_{head}$/C.

(4) **$*\mu_{head}$/C**: moraic head consonants are prohibited.

The interaction between *COMPLEX and $*\mu_h$/C also helps create a bimoraic/disyllabic word that meets the word minimality condition of MA:

(5) *COMPLEX >> $*\mu_h$/C

/CCəC/	*COMPLEX	$*\mu_h$/C
☞a- C.CəC		*
b- CCəC	*!	

Thus, the candidate that avoids complex onset via creating a syllabic consonant emerges as the most harmonic form.

The epenthetic site of schwa in MA is marked by variability. In triliteral roots, a schwa could fall either between the last two consonants (i.e. CCəC) or the first two ones (i.e. CəCC). In nouns in particular, schwa insertion is subject to the sonority level of the relevant consonants. Al Ghadi (1990/2014:85) maintains that schwa is placed before the most sonorous consonant among the last two or otherwise between the last two consonants if they have the same sonority index:[17]

(6) Schwa epenthesis in triliteral nouns

a. CəCC		b. CCəC	
fərx	'bird'	qbəṛ	'grave'
qəṛd	'monkey'	kfən	'shroud'
wəld	'boy'	ʒbəl	'mountain'
qənt	'corner'	smən	'salty butter'

This sonority condition on schwa epenthesis in triliteral nouns is formulated as follows:

(7) **The sonority condition** (Al Ghadi, 1990/2014):[18]

a. $C_1əC_2C_3$: If C_2 is more sonorous than C_3

b. $C_1C_2əC_3$: If C_3 is more sonorous than C_2 or equal to it

Such a condition can be stated in the form of a violable constraint á la OT. Therefore, through its interaction with the previously established constraints, the constraint SON-Cond would define the epenthetic position of the schwa in triliteral nouns in the following way:

---

[17] The sonority of C_1 is kept out of the equation probably for the simple reason that placing a schwa before C_1 would be costlier than all the other possible options in terms of syllable well-formedness, resulting in an onsetless syllable, which the language strictly prohibits.

[18] The vast majority of nouns obey the sonority condition. However, there are a bunch of nouns that stray away, for example: ħnəʃ 'snake' and ħəbs 'jail'.

(8) SON-Cond, PARSE-Seg, *COMPLEX >> *μ_h/C, DEP-ə

/frx/$_N$	SON-Cond	PARSE-Seg	*COMPLEX	*μ_h/C	DEP-ə
☞a. fər.x				*	*
b. f.rəx	*!			*	*
c. .fərx.			*!		*
d. frx		***!			

This tableau illustrates the effect of the first part of the sonority condition. In the optimal candidate, schwa is inserted before the liquid /r/ since it is more sonorous than the fricative /x/. Candidate (8b) is ruled out for violating the sonority condition by placing the schwa before the less sonorous consonant of the C_2C_3 cluster. Candidate (8c) respects the sonority condition in the same way the optimal form does, but it is excluded for violating *COMPLEX. Candidate (8d) is rejected simply for having no syllable structure, going against the will of PARSE-Seg.

The following tableau shows the optimal position of schwa when C_3 is more sonorous than C_2:

(9) SON-Cond, PARSE-Seg, *COMPLEX >> *μ_h/C, DEP-ə

/ʒbl/$_N$	SON-Cond	PARSE-Seg	*COMPLEX	*μ_h/C	DEP-ə
☞a. ʒ.bəl				*	*
b. .ʒbəl.			*!		*
c. ʒəb.l	*!			*	*
d. ʒbl		***!			

In this tableau, however, it is the second part of the sonority condition that gets to define the place of the schwa. The optimal candidate has its schwa epenthesized right before the lateral /l/, for it is more sonorous than the bilabial /b/. *COMPLEX disqualifies candidate (9b) due to its complex onset. Candidate (9c) wrongly places the schwa before the less sonorous consonant, and hence loses in the competition. Candidate (9d) fails as it does not conform to the stipulation of the high-ranking PARSE-Seg.

Now, let us consider the case of schwa epenthesis in triliteral verbs and adjectives. Unlike in nouns, the sonority of the individual consonants in verbs and adjectives has no impact on the placement of schwa. The latter, however, is invariably epenthesized between the last two consonants. Consider the following examples:

(10) Schwa epenthesis in triliteral verbs and adjectives

a. Verbs		*b. Adjectives (comparative)*	
ɦṛəb	‘escape’	tqəl	‘heavy’
tləf	‘disappear’	ṭwəl	‘tall’
ʃṛəb	‘drink’	ɣləḍ	‘fat’
kbər	‘grow’	sɦəl	‘easy’

In the face of this situation, an alignment constraint has been devised by Boudlal (2001) to explain the invariability that characterizes schwa epenthesis in triliteral verbs and adjectives (also see Bensoukas and Boudlal (2012a-b)). The constraint proposed requires that the verb/adjective stem and the main syllable should be aligned at the right periphery, formally written as ALIGN-R-Maj-σ.

(11) **ALIGN-R-Maj-σ**:

The right edge of the verb/adjective stem should align with the right edge of the major syllable.

If we think of SON-Cond as a constraint that is applicable only to nouns, then ALIGN-R would not need to dominate it. This means that their effects would be in complementary distribution, with each one operating on different lexical categories. However, if SON-Cond is thought of as a general constraint that is blind to the lexical category of words, then ALIGN-R would have to dominate it to keep verbs and adjectives out of its reach. For us, SON-Cond is a general constraint that should be dominated by ALIGN-R as shown by the following tableau:

(12) ALIGN-R >> SON-Cond, PARSE-Seg, *COMPLEX >> *μ_h/C, DEP-ə

/tlf/$_V$	ALIGN-R	SON-Cond	PARSE-Seg	*COMPLEX	*μ_h/C	DEP-ə
☞a. t.ləf		*			*	*
b. .tləf.		*		*!		*
c. təl.f	*!				*	*
d. tlf			***!			

What is explained by this tableau is why the schwa of the optimal candidate is inserted before the fricative /f/ rather than the more sonorous lateral /l/. On the basis of the ranking ALIGN-R >> SON-Cond, verbs and adjectives break free from the stipulation of SON-Cond in order to satisfy ALIGN-R. On the other hand, the sonority-respecting candidate is ruled out for failing to satisfy the more important right alignment. Candidates (12b) and (12d) are suboptimal for violating *COMPLEX and PARSE-Seg, respectively.

For the sake of completeness, we now turn to the syllabification of quadrisegmental roots of the form /CCCC/. Roots of this type trigger the epenthesis of two schwas to form a sequence of two light syllables: [CəC.CəC]. This is illustrated by the following items:

(13) Schwa epenthesis in quadrisegmental roots

a. Quadrisegmental verbs[19]

/ṭṛʒm/	ṭəṛʒəm	'to translate'
/blbl/	bəlbəl	'to disturb'
/drdb/	dərdəb	'to roll down'
/ʃṛml/	ʃəṛməl	'to marinate'

b. Quadrisegmental nouns

/ʃṛʒm/	ʃəṛʒəm	'a window'
/mslm/	məsləm	'muslim'
/tnbr/	tənbər	'stamp'
/drɦm/	dərɦəm	'dirham'

Based on these items, we assume that schwa epenthesis in quadrisegmental roots is affected neither by the sonority of the individual consonants nor by the morphological category of the stems in question. Instead, their syllabification can be captured simply through the following constraint hierarchy: PARSE-Seg, *COMPLEX >> *μ_h/C, DEP-ə.

[19] Boudlal (2001) suggests a constraint of the type VERB ROOT= [μμ], which requires verbs to correspond to two moras. According to Boudlal, this constraint is responsible for the epenthesis of schwas in verbs, instead of full vowels, by outranking Dep-V. However, we argue that no such constraint is needed to motivate schwa epenthesis if the constraint interaction Parse-Seg >> Dep-V >> Dep-ə is posited. Under this ranking, any candidate with epenthetic full vowels would be ruled out by Dep-V and schwa epenthesis would be the least costly option.

(14) PARSE-Seg, *COMPLEX >> *μ_h/C, DEP-ə

/CCCC/	PARSE-C	*COMPLEX	*μ_h/C	DEP-ə
☞a. CəC.CəC				**
b. .CCəCC.		**!		*
c. C.CəC.C			**!	*
d. C.CəCC		*!	*	*
e. CCəC.C		*!	*	*
f. CC.CəC	**!			*

This tableau shows that any candidate with only one instance of schwa epenthesis would be less harmonic than the optimal form. Therefore, no reference needs to be made to the constraints ALIGN-R and SON-COND in this case. Candidate (14b) is excluded for its double violation of *COMPLEX. Candidate (14c) has no complex margins and parses all of its segments but at the expense of violating *μ_h/C on more than one occasion, twice to be specific. Candidates (14d) and (14e) violate *COMPLEX by having a complex coda and a complex onset, respectively. Candidate (14f) is ruled out by PARSE-Seg for failing to parse two of its segments. In addition, if the schwas occur anywhere apart from their optimal positions, other high-ranking markedness constraints would be violated, namely ONSET and *μ/ə. For example, possible candidates such as CəC.əC.C and Cə.CəC.C violate Onset and *μ/ə, respectively.

With this out of the way, we are now in a good position to introduce our analysis of the issue of geminate integrity in relation to schwa epenthesis and word formation in MA.

3. Geminate integrity: A review

3.1. Benhallam (1980): Rule types and geminate integrity

Cross-linguistically, the characterization of geminate behavior has been a recalcitrant issue for phonologists. In previous works (Kenstowicz and Pyle, 1973; Saib, 1977; Guerssel, 1977; Leben, 1980), it has been elaborately demonstrated that, within a single language, geminates may display a dual behavior by combining qualities of both cluster structures and unit structures. This has raised concerns about how to accurately represent geminates in phonological terms in order to account for the observed behavior. We weighed in on this issue by suggesting a hybrid representation of geminates that seems to justify their inconsistent patterning. In this chapter, we continue to unravel the mystery of geminate behavior. This time we focus on the issue of geminate integrity in the context of MA data.

Early works on the issue of geminate behavior have predominantly conceived of geminates as cluster structures that pattern with sequences of unlike consonants. However, upon further investigation, it has been observed that geminates prefer to keep the adjacency of their constituent members intact, exhibiting some degree of integrity that normal consonant sequences lack. One manifestation of this preference is their resistance to rules of epenthesis. To account for this anomaly, the Integrity Hypothesis (IH) (Kenstowicz and Pyle, 1973) and the Adjacency Identity Constraint (AIC) (Guerssel, 1977) have been proposed.

(15) The IH (Kenstowicz and Pyle, 1973)

> *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.*

(16) The AIC (Guerssel, 1977)

Given two segments A1A2 where A1=A2, a phonological rule can alter the adjacency of A1A2 if and only if it alters the identity of A1 or A2.

These took the form of general constraints on rules. Their function, then, consisted in excluding geminates from the effect of rules that apply to other clusters. The IH and AIC have been pronounced in an inflexible manner by predicting that the unbroken state of geminates is the most natural state of affairs, in contradiction to the facts shown by the data from a number of languages.

On this note, it was not long after Benhallam (1980) retorted that, contrary to the predictions made by the IH and the AIC, the adjacency of geminate clusters can sometimes be predictively altered. Benhallam (1980) suggested that not all rules are equal in relation to the adjacency of geminates. On this basis, a modification to the theory of geminate integrity was advanced, considering the nature of the rules in question. The author argued that geminates tend to have different reactions to different types of rules, such that they can be split by morphological rules but not by phonological ones. To capture this generalization, Benhallam (1980) formulated a constraint on rules dubbed the Geminate Law (GL):

(17) The Geminate Law (Benhallam, 1980:141)[20]

Geminates can be split up by morphological (or morpholexical) rules but not by phonological rules.

The GL is based on a close examination of data from Palestinian Arabic (PA), Standard Arabic (SA) as well as MA. Consider the following cases from PA:

(18) PA[21]

a. sad	'to close'	sad**i**d	'closed'
b. ħal	'to open'	ħal**i**l	'open'
c. xad	'cheek'	xd**u**d	'cheeks'

According to Benhallam, an analysis that does not recognize a distinction between phonological rules and morphological ones would have to posit ad hoc solutions to the problem of the inconsistency of geminate integrity. For example, in an attempt to make the IH reign supreme over all rules equally, Guerssel (1977) claimed that the geminates in the first two items from PA are underlyingly split by the vowel /i/. Benhallam criticized this by introducing the datum in (18c), also from PA, which is problematic for Guerssel's view since it entails that the plural [xdud] is the underlying form, from which the singular form is derived by means of vowel deletion, degemination and vowel insertion. For Benhallam (1980), this seems very unlikely, simply because such a claim would lead to a lot of confusion and inconsistency elsewhere. For elaboration, Benhallam presented the following data from SA to highlight the problem:

[20] A reservation about this terminology is in order. The term 'law' gives the impression that the relevant hypothesis is universal and unchallenged. The author himself admits that this is not meant to be a universal, but only a tendency of geminate behavior. A better name could be the Rule-type condition on geminate integrity.
[21] A rule of final degemination applies here (See Benhallam (1980)).

(19) SA

a. ʃubba:k	‘window’	ʃab**a:**bi:k	‘windows’
b. sadd	‘dam’	sud**u:**d	‘dams’
c. ħabi:b	‘lover’	ʔaħibba	‘lovers’

The point here is that if the integrity of geminates becomes a criterion for deciding on underlying forms, one would have to list the plural forms as underlying in the cases of (19a) and (19b), while listing the singular form as underlying in the case of (19c). Furthermore, in the case of the derivatives in (20), it would be unclear if one should posit the singular form of the adjective or its plural form as underlying since both occur with split geminates.

(20)

ʒadi:d	‘new (sg.)’
ʒudud	‘new (pl.)’
ʔaʒadd	‘newer’

Analogously, when inflected for person and tense, some geminated verbs give rise to paradigms whose members vary in terms of the integrity of their geminates. Some forms appear with unbroken geminates, while others have broken geminates. This is exemplified in (21) with the past tense paradigm for the verb ʃaqq ‘to crack’. Given Guerssel’s proposal to posit forms with split geminates as underlying, the question that Benhallam raises here is: what form in the paradigm below, among those with split geminates, should be the underlying form? Note that the paradigm includes more than one form where the geminate is split. Evidently, it would be very implausible to suggest positing more than one form as underlying.

(21) The paradigm problem:

ʃaqqa	‘he cracked’
ʃaqqu:	‘they cracked’
ʃaqaqtu	‘I cracked’
ʃaqaqta	‘you cracked’
ʃaqaqtum	‘you (pl.) cracked’

With these serious problems in the picture, Benhallam confidently suggested that the GL provides better insights into the issue of geminate integrity. Instead of meddling with underlying representations, the GL predicts that geminates get split under the effect of morphological processes. This is supported by the above-discussed data where splitting geminates correlates with the formation of new word forms. Phonological processes, on the other hand, fail to split geminates as has been evidenced by some of the data presented in Kenstowicz and Pyle (1973) and Guerssel (1977).

On a later note, Benhallam cautioned that there exit some notable contradictions to his GL, calling for a revised version that considers the new facts represented by the following set of data:

(22) Derived geminates split by phonology:

/m-mdd/	məmdud	‘stretch’
/m-mlk/	məmluk	‘own’
/m-mnʕ/	məmnuʕ	‘forbid’

The contradiction showcased by this data has to do with what appears to be the splitting of some geminates by means of a phonological rule. The relevant items above are all past participle forms whose geminates arise through affixing the past participle morpheme /m-u/ to base forms that happen to begin with the sound /m/. The result is a derived geminate whose integrity is later compromised by schwa epenthesis, in contradiction with the prediction of the GL. Therefore, to conform to the new facts, a revision proved necessary.

(23) The Geminate Law Revised (Benhallam, 1980:145)

> *Underlying geminate clusters can be split up by morphological or phonolexical rules but not by phonological rules. Derived geminate clusters can be split up by phonological rules.*

The revised version of the GL draws a further distinction between underlying geminates and derived geminates. Only the adjacency of underlying geminates is resistant to phonological rules under the new version. Derived geminates, however, are assumed to be prone to the effect of phonology.

Benhallam briefly pointed out to the fact that the new version of the GL does not say whether, like underlying geminates, derived geminates can be split by morphological rules or not. This is due to the absence of data that can support such a claim. In this regard, we suppose that the GL implicitly predicts that morphological rules should be able to split derived geminates, given their ability to split underlying ones. Coming up with data that could back up this prediction is difficult since it would require two morphological processes interfering with each other. For example, such a case would ideally involve a morphological material which splits a geminate derived through the concatenation of two morphological forms.

Later, it will be shown that the GL does not require any revision to account for this data. We will also show that the dictate of the GL can simply emerge from constraint interaction á la OT. Moreover, we will demonstrate that a distinction between phonological rules and morphological ones suffices to explain all cases of geminate breaking in MA, including those that were deemed as exceptions in Benhallam (1991).

In view of the GL, the alterable integrity of geminates in MA becomes justified. The items on the left in (24) show that schwa, which is phonologically motivated, is unable to break up the relevant geminates. In the items on the right, however, geminates are split by full vowels that are morphologically or lexically motivated.

(24) Geminate integrity and schwa epenthesis

a.	*Verb*		*PP*	
	s**ə**dd	*sd**ə**d	məs.d**u**d	'close'
	ħ**ə**ll	*ħl**ə**l	məħ.l**u**l	'open'
	ʃ**ə**dd	*ʃd**ə**d	məʃ.d**u**d	'catch'
	d**ə**qq	*dq**ə**q	məd.q**u**q	'knock'
b.	*Comparative*		*Adjective*	
	ṛ**ə**qq	*ṛq**ə**q	ṛq**i**q	'thin'
	x**ə**ff	*xf**ə**f	xf**i**f	'light'
	b**ə**nn	*bn**ə**n	bn**i**n	'tasty
	ṣ**ə**ħħ	*ṣħ**ə**ħ	ṣħ**i**ħ	'strong'

3.2. Benhallam (1991): Limitations of the GL

As a follow-up to his earlier work on geminates, in 1991, Benhallam published a new article wherein he revisited the issue of geminate integrity in light of new data. In this work, the GL was criticized for being unable to account for all cases of geminate breaking in MA. Specifically, the following items were cited as problematic counterexamples to the GL:

(25)

sləl	'baskets'
gfəf	'bags'
rẓəẓ	'turbans'
bərgəg	'to spy'

In these items, schwa epenthesis splits underlying geminates, in conflict with the prediction that all geminates should resist breaking by means of phonological rules. In accounting for this problematic data, Benhallam (1991) dispensed entirely with the idea of the GL and proposed a new account where geminates are treated as sequences of normal consonants with no integrity whatsoever. As a result, cases where geminates exhibit resistance to schwa epenthesis are now considered exceptions to the normal application of schwa insertion.

The new state of affairs looks as follows. Given every triliteral root, a schwa is inserted between C_2 and C_3. Therefore, as far as this general rule of epenthesis is concerned, the broken plural forms in (25) become the norm. Every form that deviates from this rule should then be treated as an exception. For example, words like sədd 'to close' and bənt 'girl' are considered as equally problematic; thus, the concept of geminate integrity is no longer necessary. Benhallam needed two rules to explain this pattern. The first one is a lexical rule that inserts a V position after C_1 of some lexical categories and subcategories. The rule is reproduced below, where X and Y stand for any number of segments and lexical categories, respectively:

(26) $\varnothing \longrightarrow V / \# C\_\ X]_{Y}$

He also proposed a second rule which applies postlexically to fill in the V position. The rule inserts a schwa in the empty nucleus, deriving the structure CəCC. This can be illustrated as follows:

(27) $\varnothing \longrightarrow$ ə / $\underline{V}$

This new way of looking at the issue comes at a cost, however. As has been mentioned before, the concept of geminate integrity becomes irrelevant in this analysis. This alienates the geminate patterns of MA from their crosslinguistic context since it has been shown that geminates do exhibit some degree of integrity across many languages. Thus, we end up having an analysis that works just for MA. Also, the treatment of schwa epenthesis, in forms like gfəf ‘bags’, as a merely phonological process leaves some unanswered questions: At worst, it seems like the morphology of such forms is left out; at best, it implies that the roots from which these plurals are derived are underlyingly marked for plurality without supporting such a claim.

4. An OT account of geminate integrity

In what follows, we will demonstrate that the OT framework allows for an elegant and unified analysis of geminate integrity in MA, which eschews the pitfalls of previous analyses. we will try to show that the inconsistent integrity of geminates in MA is best captured through constraint interactions. More specifically, we will argue for a fixed ranking of geminate integrity between morphology and phonology, such that morphology dominates geminate integrity, which in turn dominates phonology. By doing so, we will maintain that the broken plural exceptions mentioned in Benhallam (1991) are in fact morphologically relevant, hence put in the morphology side along with the case of the past participle derivation. As for schwa epenthesis, it will be shown that when schwa fails to break a geminate, it is because of its low-ranking status as a phonological process. However, when schwa splits geminates elsewhere, it is probably because the geminates in question are fake. Therefore, we will name our proposed account for geminate integrity in MA: The Fixed Integrity Ranking.

4.1. Schwa epenthesis in the context of geminate integrity

The GL, as argued for in Benhallam (1980), belongs to the pre-OT literature. It is formulated in the form of a general constraint on rewrite rules. It permits morphological rules to split up geminates while prohibiting phonological rules from doing so. In that framework, the GL can be criticized for being an extra cost on the grammar of the language, intuitively stated and allegedly inviolable. In OT, the effect of the GL can be achieved simply through the interaction of universal and violable constraints. To this end, a constraint named GEM-Integ (short for Geminate Integrity) is posited to represent the resistance to breaking exhibited by geminates:

(28) **GEM-Integrity**:

A geminate consonant cannot be split.

We have previously shown that schwa epenthesis in triliteral verbs and adjectives is subject to an alignment constraint that demands the main syllable to align with the word stem at the right periphery. However, lexically geminated verbs and adjectives appear to violate this requirement. Here comes the role of GEM-Integ. By dominating ALIGN-R, GEM-Integ prohibits the schwa from splitting the designated geminates. Rather, schwa epenthesis takes place between C_1 and G. The interaction between GEM-Integ and ALIGN-R is illustrated by the following tableau:

(29) GEM-Integ >> ALIGN-R

$/CC_xC_x/_{V/adj}$	GEM-Integ	ALIGN-R
☞a. $C\text{ə}C_xC_x$		*
c. $C.C_x\text{ə}C_x$	*!	

This tableau shows that candidate (29c) is penalized by the high-ranking GEM-Integ due to its broken geminate. Hence, it loses to the optimal candidate, which succeeds in maintaining the integrity of its geminate.

The tableau in (30) below presents other competing candidates and interacting constraints. Specifically, the markedness constraints *CON-GEM and *Empty-Headedσ are introduced to distinguish between the candidates *səd.d* and *sədd.* *CON-GEM disfavors candidates with tautosyllabic geminates whereas *Empty-Headedσ prohibits non-headed syllables:

(30) GEM-Integ, *CON-GEM, PARSE-Seg >> ALIGN-R, *Empty-Headedσ, DEP-ə

/sdd/	G-Integ	*CON-GEM	PARSE-Seg	ALIGN-R	*Empty-Headedσ	DEP-ə
☞a. səd.d				*	*	*
c. s.dəd	*!					*
d. .sədd.		*!		*		*
e. sdd			**!			

In this tableau, both the optimum and candidate (30d) respect geminate integrity. However, candidate (30d) is inferior to the optimum in that it contains a contained final geminate. In the optimal candidate, the first half of the geminate syllabifies into a coda whereas the second half functions as the onset of an empty-headed syllable in satisfaction of the high-ranking *CON-GEM, militating against tautosyllabic contained geminates. Candidate (30e) stands no chance in the competition since none of its segments is parsed.

4.2. Broken Plurals

In this section, we provide a fresh look at the issue of geminate integrity in MA. We show that, by rushing to undermine the generalization of the GL, Benhallam (1991) overlooks some important details about the issue. We argue that a distinction between phonological processes and morphological processes can be the basis of an adequate account of all the patterns of geminate integrity in MA. To achieve this goal, our analysis will be entertained under the rubric of OT.

It is true that the Fixed Integrity Ranking predicts that all underlying geminates should resist schwa epenthesis. Therefore, the items in (31) seemingly stand as counterexamples to this prediction. However, it could very well be that the status of geminates in these items follows from the provision of the Fixed Integrity Ranking itself if a careful look is taken. Recall that the Fixed Integrity Ranking also predicts that morphological processes can break geminates. Therefore, is not this a case of a morphologically motivated geminate breaking?

(31) Broken geminates

Sg.	*Pl.*	
səlla	sləl	‘basket’
gəffa	gfəf	‘bag’
rəẓẓa	rẓəẓ	‘turban’
səkka	skək	‘rail’
dərra	drər	‘veil’
dəffa	dfəf	‘door’
sədda	sdəd	‘mezzanine’
ʕəʃʃ	ʕʃəʃ	‘nest’
ħukk	ħkək	‘can’

In this set of data, the broken plurals on the right consist of a consonantal root, which they share with their singular counterparts, plus schwa that splits their geminates. Other broken plurals have full vowels, instead. Consider the following examples for illustration:

(32)

Sg.	*Pl.*	
muxx	mxax	‘brain’
kumm	kmam	‘sleeve’
muʃʃ	mʃaʃ	‘cat’
ʕəmm	ʕmam	‘uncle’
ʒədd	ʒdud	‘grandfather’
xədd	xdud	‘cheek’
qəṭṭ	qṭuṭ	‘cat’
kəff	kfuf	‘palm’
ṛəff	ṛfuf	‘shelf’
ʒənn	ʒnun	‘ghost’
xəṭṭ	xṭuṭ	‘line’

In particular, the geminates in the broken plurals in (32) are split either by the vowel /a/ or /u/. The nature of the vowel that each plural happens to have seems to be unpredictable.[22] We believe that an analysis which recognizes the morphological equivalence between the broken plurals in (31) and those in (32) is preferable. Also, such an analysis appears to require fewer assumptions about the phonology of schwa. Additionally, this will also help us maintain an analysis that could be extended to the items below:

[22] The unpredictability of the vowels in these plurals could be explained by a subcategorization mechanism, whereby different nouns can be lexically related to one vowel or the other.

(33)

a.	*Comparative*	*Adjective*	
	ṛəqq	ṛqiq	‘thin’
	xəff	xfif	‘light’
	bənn	bnin	‘tasty’
	ṣəħħ	ṣħiħ	‘strong’

b.	*Verb*	*Past participle*	
	sədd	məsdud	‘close’
	ħəll	məħlul	‘open’
	ʕəḍḍ	məʕḍuḍ	‘bite’
	ʃəqq	məʃquq	‘crack’
	ʃədd	məʃdud	‘catch’
	ħədd	məħdud	‘limit’

By doing this, we will be able to maintain that all cases of geminate breaking ensue from morphological demands.

To this end, two hypotheses will be considered. The first hypothesis that intuitively comes to mind would be to assume that the schwa of the broken plurals in (31) is a morphological marker, in the same way the full vowels in (32) through (33) are, arguably, morphologically induced. In this case, the schwa in (31) should be distinguished from the schwa used to break consonant clusters for purely phonological purposes. This way, we could account for our data in total agreement with the stipulation of the Fixed Integrity Ranking. However, this proposal seems unfavorable for two reasons. For one, granting a morphemic status to the schwa is not supported by other data in the language, unlike full vowels whose morphological use recurs on more than one occasion. For another, by distinguishing a morphological schwa and a phonological one, we sacrifice the phonological generalization of schwa epenthesis, meaning that not all schwas would have to be phonologically predictable.

For these reasons, an alternative treatment is suggested. We argue that the geminates in (31) are not broken by schwa epenthesis per se. Rather, they are split by means of morphological activity. To be more specific, we believe that the optimum is marked for plurality by means of neutralizing the geminate integrity of the input, meaning that the position between the parts of the geminate becomes available for schwa epenthesis. Schwa in this case has no morphological status by itself.[23] To make our account more concrete, we need the following constraints, some of which we have already presented:

(34)

a. **PARSE-Segment**: assign one violation to every unsyllabified segment.

b. **DEP-ə**: every element in the output must have a correspondent in the input (no insertion).

c. **RM-Plural**: the plural morpheme must have some phonological exponence in the output form, whereby the latter should be unfaithful to the root.

d. **GEM-Integrity**: a geminate consonant cannot be split.

[23] Our account for this data assumes a root-based approach to word derivation in MA. See Noamane (2014, 2018) for arguments in favor of the morphological status of the root in MA.

The interaction between the constraints PARSE-Seg and DEP-ə is conventionally responsible for schwa epenthesis for purely phonological reasons, namely breaking consonant clusters. The RM-Pl constraint in (34c) requires some degree of unfaithfulness between the input and the output for the sake of realizing the plural morpheme (Kurisu, 2001). The constraint in (34d) demands the maintenance of the integrity of geminates. Through the ranking of RM-Pl over GEM-Integrity, some degree of unfaithfulness can be achieved. In this case, the integrity of the underlying geminate in the input can be broken as a very subtle way to realize the plural morpheme. This is reflected by schwa epenthesis, which then interprets the relevant geminate as a sequence of two consonants.

(35) PARSE-Seg, RM-Pl >> GEM-Integrity, DEP-ə

√sll	PARSE-Seg	RM-Plural	GEM-Integrity	DEP-ə
☞a. sləl			*	*
b. səll		W*!	L	*
c. sll	W*!	W*	L	L

The inserted schwa in candidate (35b) cannot accomplish any morphological function. In fact, in this context, schwa epenthesis is a mere syllabification auxiliary. Thus, candidate (35b) fails to realize the plural morpheme by remaining faithful to the input. Arguably, such a candidate does not have any morphological exponence that could represent the plural morpheme. The optimal candidate (35a) satisfies RM-Pl at the expense of GEM-Integrity, thereby allowing the schwa to occur between the parts of the split geminate.

4.3. The past participle

This section will be devoted to investigating the morphological effect of the past participle (PP) derivation on geminate integrity. We look at those PPs derived from roots with final geminates. To illustrate, consider (36) below:

(36) məC.C_iuC_i:

Verb	*PP*	
sədd	məsdud	'close'
ħəll	məħlul	'open'
ʕəḍḍ	məʕḍuḍ	'bite'
ʃəqq	məʃquq	'crack'
ʃədd	məʃdud	'catch'
ħədd	məħdud	'limit'

Note that the same geminates that resist schwa epenthesis in the verb forms in (36) tend to be split by the vowel /u/ in the corresponding PP forms. In compliance with the line of reasoning that we have been pursuing in this paper, our analysis of this case of geminate breaking is also underlain by the basic assumption that geminates can only be split by morphological rules but not by phonological ones. Like in the case of the broken plural, we demonstrate that such assumption is supported by constraint interaction á la OT.

In this regard, the other assumption that follows is that the PP affix is represented by the discontinuous morpheme /m-u/. This means that the vowel /u/ is argued to be part of the PP morpheme. Such an assumption is substantiated by two main factors. First, we believe that assuming that the vowel /u/ is able to split geminates thanks to its morphological status resonates well with the postulations of the Fixed Integrity Ranking. Claiming otherwise would undermine the latter's generalization. Second, the vowel /u/ establishes morphological contrast between a large set of nouns and their PP counterparts, working as a meaning carrier morpheme. For the sake of illustration, some examples are listed below:

(37)

PP	*Noun*	
məktub	məktəb	'write'
məlʕub	məlʕəb	'play'
məḍrub	məḍrəb	'hit'
mədfun	mədfən	'bury'
mərbuṭ	mərbəṭ	'tie'
məxzun	məxzən	'store'
məʃṛub	məʃṛəb	'drink'
mədluk	mədlək	'flat'
məlbus	məlbəs	'wear'

Alternatively, one could attribute the morphological contrast between the nouns and PPs in (37) to their templates (see Boudlal, 2001). Under this view, the nature of the epenthetic vowel follows from the templatic properties of each form. However, the morphemic status of /u/ is further supported by the fact that no nouns of the shape [məC.C_iəC_i] are found in MA, where the last two consonants can be the result of geminate breaking. If both the vowel /u/ and the schwa were equally epenthetic vowels that are used to satisfy some templatic requirement, then nouns like the ones in (38) should also be possible. Said differently, for templatic reasons, one would also expect the schwa to break the geminates in the potential, yet nonexistent, nominals in (38). Nonetheless, the vowel /u/ is exclusively entitled to split geminates.

(38)

sədd	*məs.dəd
ħəll	*məħ.ləl
ʃədd	*məʃ.dəd
ʕəḍḍ	*məʕ.ḍəḍ
ʃəqq	*məʃ.qəq

Therefore, the fact that no such lexical items are attested could be interpreted as an indication of the inadequacy of templatic effects in predicting the nature of epenthetic vowels in MA. Accordingly, it is more convincing to treat the vowel /u/ as part of the PP morpheme (Joe Pater, personal communication).

In Alignment Theory, morphemes have no intrinsic affixal status. Their locations are defined by a category of constraints dubbed anchor or alignment constraints. Anchor constraints are one of Prince and Smolensky's (1993/2004) earliest contributions in OT. Building on that, McCarthy and Prince (1993) proposed a general family of constraints to capture the various constituent-edge effects in both morphology and phonology. Linguistic theory provides the grammar with a wide range of prosodic (PCat) and grammatical (GCat) categories. Thus, "a GA requirement demands that a designated edge of each prosodic or morphological constituent of type Cat1 coincide with a designated edge of some other prosodic or morphological constituent Cat2." (McCarthy and Prince, 1993:2) The general schema of this constraint family comes as follows:

(39)

Generalized Alignment: (McCarthy and Prince, 1993)

Align (Cat1, Edge1, Cat2, Edge2) =def

$\forall$ Cat1 $\exists$ Cat2 such that Edge1 of Cat1 and Edge2 of Cat2 coincide

Where

Cat1, Cat2 $\in$ PCat $\cup$ GCat

Edge1, Edge2 $\in$ {Right, Left}

Although alignment constraints are equally violated, their violation should be kept minimal. The designated affix should be as close as possible to the designated edge. Therefore, alignment constraints need to be gradiently assessed for violations, whereby the degree or multiplicity of violation is measured in terms of distance from the designated edge. The formal constraint which represents this general constraint family is ALIGN, which can be then specified for the targeted edges and the relevant categories.

One of the major functions of ALIGN constraints is the formation of new words by affixing morphemes to the left or the right of a stem. GA conceives of affixation as an edge-oriented phenomenon. Under this model, the prefixhood or suffixhood of morphemes is dictated by alignment constraints. In this context, prefixation and suffixation occur when ALIGN constraints refer to the left edge and the right edge, respectively.

Under this conception, the constraint that is responsible for the morphological distribution of the PP morpheme in our OT analysis is the following alignment constraint:

(40) **ALIGN (m, L, u, R)**

The right edge of /m-/ is aligned to the left edge of the stem, the left edge of /-u/ is aligned to the right edge to the stem.

This constraint characterizes the PP morpheme as a circumfix, in that the [m] of the affix is aligned to the left edge of the root while the [u] is aligned to the right edge of the root. However, while the [m] is consistently left aligned, the [u] of the morpheme always appears inside the derived forms, contrary to the stipulation of the posited alignment constraint. In order to account for the misalignment of the /u/, we postulate another alignment constraint, which we define as follows:

(41) **ALIGN- (Rt, R, PrWd, R)**

The right edge of the root should coincide with the right edge of the prosodic word.

The type of demand made by this alignment constraint is that the right edge of the root should match the right edge of the derived PP form. By being in a dominant position, this constraint pushes the [u] of the PP morpheme inside the prosodic word. The interaction between these two alignment constraints is illustrated by the following tableau:

(42) ALIGN- (Rt, R, PrWd, R) >> ALIGN (m, L, u, R)[24]

√sdd /m, u/	ALIGN- (Rt, R, PrWd, R)	ALIGN-affix-PP
☞a. məs.dud		*
b. məs.d.du	*!	

[24] This analysis can be extended to account for all the other PP classes. See section 6 in this chapter for a comprehensive analysis of the PP formation in MA.

Evidently, candidate (42b) loses in the competition for failing to bring the right edges of the root and the prosodic word together. In satisfaction of the dominating constraint, the winning candidate infixes the suffixal part of the PP morpheme, allowing the right edge of the root and that of the prosodic word to match.

This brings us back to the issue of geminate integrity. In this concern, when the /u/ of the PP morpheme moves inside the derived form, it splits the relevant geminates in the process. This means that the alignment constraint regulating the edges of the root and the prosodic word has to outrank the constraint GEM-Integrity as well.

(43) ALIGN- (Rt, R, Pwrd, R) >> GEM-Integrity

√sdd /m, u/	ALIGN- (Rt, R, Pwrd, R)	GEM-Integrity
☞a. məs.dud		*
b. məs.d.du	*!	

What disqualifies candidate (43b) is its persistence to observe the integrity of the geminate by keeping the /u/ of the PP morpheme at the right edge, hence violating the high-ranking alignment constraint. The winner, however, gives up the integrity of the geminate in question by allowing the right edges of the root and the prosodic word to match. A summary tableau is provided below:

(44) ALIGN- (Rt, R, Pwrd, R) >> ALIGN (m, L, u, R), GEM-Integrity

√sdd /m, u/	ALIGN- (Rt, R, Pwrd, R)	ALIGN (m, L, u, R)	GEM-Integrity
☞a. məs.dud		*	*
b. məs.d.du	*!W	L	L
c. m.sudd		**!	

Recall that, in the theory of GA, affixation is edge-oriented. Therefore, affixes should be as close to their designated edges as possible. This means that moving the /u/ of the PP morpheme further inside the prosodic word would be costlier, causing multiple violations of ALIGN-Affix. This situation is represented by candidate (44c), which is excluded exactly for this reason.

To recapitulate, we have been trying to prove that the variable nature of geminate integrity in MA follows from constraint interaction á la OT. Particularly, it has been shown that, on the one hand, the phonological process of schwa epenthesis fails to break geminates due to the ranking of Gem-Integrity above ALIGN- (Major-σ, R). On the other hand, it has been demonstrated that geminate integrity can be compromised only under morphological pressure like in the cases of the broken plural and the past participle, whereby the constraint Gem-integrity is outranked by RM-Pl and ALIGN (Rt, R, PWrd, R), respectively. In summary, the patterning of geminate integrity in MA can be captured by the general ranking of GEM-Integrity between morphology and phonology in the following way: Morphology >> GEM-Integrity >> Phonology.

5. Some persisting issues

For the sake of exhaustiveness, there are other interesting portions of the data that remain to be captured under our analysis. First, there is this set of items which does not comply with the stipulation of RM-Pl. In the plurals listed in (45) below, the plural morpheme is not realized by neutralizing the integrity of the relevant geminates as is the case in (31) above. Instead, the inserted schwa happens to obey geminate integrity. This could undermine the analysis we have been trying to construct. Consider the following examples:

(45)

Sg.	*Pl.*	
bəqqa	bəqq	'bug'
ħəbba	ħəbb	'grain'
bəẓẓa	bəẓẓ	'brat'
qəʃʃa	qəʃʃ	'junk'

The question we ask now is: why does the schwa not split the geminates in these plurals? To answer this question, we argue that the roots of these plurals are inherently marked for plurality. A simple diagnosis to confirm the plurality of the designated roots is to put them after the quantifying phrase 'bəzzaf djal', translated into 'plenty of'. Only plural nouns can come after such a phrase.

(46)

bəzzaf djal l- bəqq
bəẓẓ
ħəbb
qəʃʃ

The underlying assumption here is that the singular form of these nouns is derived from the mass plural forms stated above. However, it is worth noting that these nouns can also have count plurals, which are plurals that refer to discrete entities, derived though normal concatenation. This can be exemplified as follows:

(47)

Mass Pl.	*Sg.*	*Count Pl.*	
bəqq	bəqqa	bəqqat	'bug'
ħəbb	ħəbba	ħəbbat	'grain'
bəẓẓ	bəẓẓa	bəẓẓat	'brat'
qəʃʃ	qəʃʃa	qəʃʃat	'junk'

Therefore, since RM-Pl is inherently satisfied by the root of these nouns, there is no morphological motivation for schwa epenthesis to break the integrity of the geminates for the sake of deriving the plural form. Therefore, in the absence of the morphological incentive to split the geminates of the forms in (46), the latter would preserve the integrity of their geminates. Any gratuitous change of the integrity of the relevant geminates would be penalized by the constraint GEM-Integrity.

(48) PARSE-Seg >> RM-Plural >> GEM-Integrity, DEP-ə

√bqq (pl)	PARSE-Seg	RM-Plural	GEM-Integrity	DEP-ə
☞a. bəq.q				*
b. bqq	*!			
c. b.qəq			*!	*

Since the root is inherently marked for plurality, there is no need for candidate (48c) to give up the integrity of its geminate. Thus, candidate (48c) is ruled out due to its violation of GEM-Integrity. Candidate (48b) is excluded by PARSE-Seg for failing to parse its segments. Candidate (48a) emerges as the winner for keeping the integrity of its geminate intact.

There is another set of data that needs to be handled so that our analysis can stand firm. This time we are considering instances of underlying geminates that are spilt by schwa epenthesis in the absence of any morphological pressure. Some examples are shown below:

(49)

3p sg.	*2p sg.*	*3p pl.*	
bərgəg	bərgəgti	bərggu	'spy'
fərtət	fərtətti	fərttu	'crumb'
ħənẓəẓ	ħənẓəẓti	ħənẓẓu	'stare'

So far, we have been trying to substantiate the idea that geminates can break only when morphology compels them to do so. However, the items in (49) seemingly contradict this line of reasoning. There are two ways out. One way is to consider these cases as exceptions. The other way is to dig deeper and see if we are really dealing with geminates. By asking this question, we proceed to make a comparison between the verbs in (49) and those in (50). We notice that the 'geminates' in (49) do not exhibit any of the geminate qualities that we have come to know and are illustrated by the examples in (51). The latter show that geminates are characterized by resisting schwa epenthesis and triggering [i]-epenthesis in verb inflection. Rather, the 'geminates' in (49) seem to behave in the same way as any two unlike consonants, as shown by (50).

(50)

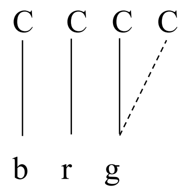
ṭərʒəm	ṭərʒəmti	ṭərʒmu	'translate'
kərkəb	kərkəbti	kərkbu	'roll'
dərdək	dərdəkti	dərdku	'roll over'

(51)

sədd	sədditi	səddu	'close'
ħəll	ħəlliti	ħəllu	'open'
ʕəḍḍ	ʕəḍḍiti	ʕəḍḍu	'bite'

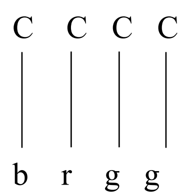
Therefore, the apparent geminates in (49) are but mere sequences of two identical singletons. Hence, this should explain why there is a schwa between them. Prior to the advent of OT, the OCP (McCarthy, 1986) could not allow the occurrence of two identical adjacent elements at the same melodic level. As a result, the items in (49) could only have geminates that are derived through rightward spreading of the last root consonant, creating a structure that would prevent vowel epenthesis in respect of the non-crossing association lines constraint.

(52) The pre-OT OCP prohibits identical melodic elements



In OT, however, there are no restrictions on underlying structures, in the sense that any type of input can be posited. This is referred to as 'richness of the base' (Prince and Smolensky, 1993/2004). Under this view, we can think of the items in (49) as being derived from inputs with sequences of two identical consonants.

(53) Identical melodic elements are possible in OT



Also, special to OT is the idea that all constraints are violable as long as they can be outranked by other constraints. On this basis, the OCP becomes violable under the pressure of higher-ranking constraints. For a structure of two identical consonants to surface, we suggest that the OCP should be dominated by a faithfulness constraint demanding the preservation of the featural identity of the input root nodes, hence IDENT-RN.

(54) IDENT-RN >> OCP

/brgg/	IDENT-RN	OCP
☞a. bərgəg		*
b. bərgəC_X	*!	

Candidate (54b) represents any possible candidate where the last consonant changes its featural identity to become different from the adjacent identical consonant. Any candidate of this form would be ruled out by IDENT-RN. The winning candidate under this ranking would always be the one where the identity of all root nodes is preserved.

Another explanation that we can give to this case of schwa epenthesis is to think of the alleged geminates, in the quadriliteral verbs above, as fake geminates, occurring across morpheme boundaries. For this to be possible, we will have to postulate that the second member of the fake geminate is a reduplicant morpheme whose role is to derive what we can call repetitive verbs (i.e. verbs that express repetition of the action). This way, schwa epenthesis in these verbs becomes justifiable, in the sense that it does not break a true geminate, but a fake one.

(55) The reduplicative approach

Root	*Root+RED*	*Derived verbs*
brg	brg+C	bərgəg
frt	frt+C	fərtət
ħnẓ	ħnẓ+C	ħənẓəẓ

Reduplication in MA is not unheard of since there is a number of verbs characterized by the repetition of their first syllable (see Imouzaz, 2002). Some examples are provided below:

(56)

Base verb	*Derived verbs*	
ʃəmm	ʃəm.ʃəm	'to smell'
ʒəṛṛ	ʒəṛ.ʒəṛ	'to pull'
dəqq	dəq.dəq	'to knock'

Next, we proceed to consider more cases of apparent exceptions to the generalization put forward by the Fixed Integrity Ranking. This time we deal with another case of schwa epenthesis which, supposedly, breaks geminate consonants. Note that the relevant 'geminates' in (57) are created through morpheme concatenation. They can only occur when the base form of the derived past participles arbitrarily begins with the consonant /m/, matching that of the past participle affix, believed here to be the discontinuous morpheme /m-u/.

(57) Fake geminates

məmdud	[hypercorrected as]	mmdud	'stretched'
məmnuʕ		mmnuʕ	'forbidden'
məmluk		mmluk	'owned'
məmʃuṭ		mmʃuṭ	'combed'

Previously, we mentioned that this data has led Benhallam (1980) to revise the initial version of the GL, thus distinguishing between underlying geminates and derived geminates. In the revised version, derived geminates were allowed to be split by schwa. We weigh in on this issue by claiming that the revised version of the GL is based on a false assumption and could have been avoided. Particularly, we believe that the 'geminates' in (57) ought to be classified as false (or fake) geminates.[25] These are structures that occur across morpheme or phrase boundaries and exhibit none of the properties of true geminates. The latter could be of two types: underlying or derived. Derived geminates can be phonological, created through assimilation, or morphological, created through mora affixation (e.g. √ktb >> kə**tt**əb. Be they underlying or derived, true geminates are expected to behave the same way vis-á-vis schwa epenthesis. This claim is substantiated by the following data:

(58) Derived geminates resist schwa epenthesis:

a.	/l-sma/	ssma	*səsma	'the sky'
	/l-ḍra/	ḍḍṛa	*ḍəḍṛa	'the corn'
b.	/ktra/	k**ə**tra		'plenty'
	/ħʃma/	ħ**ə**ʃma		'shyness'

As shown in (58), geminates derived through total assimilation between the definite article morpheme and the first root consonant display geminate integrity. Thus, it is not true that derived geminates can be split by schwa epenthesis as the revised version of the GL claims. In fact, the case of schwa epenthesis in (57) should not be of much concern to studies dealing with geminates. They are mere sequences of identical segments occurring across morpheme boundaries. Hence, schwa epenthesis treats them like any other consonant sequences. The fake geminates in (57) are sometimes misleadingly hyper-corrected as unbroken geminates, giving the impression that they are true geminates. This hypercorrection process could be seen as a reflection of native speakers' knowledge of the nature of true geminates in relation to schwa

[25] See chapter 2 for an elaborate discussion of the properties of true geminates vs. fake geminates.

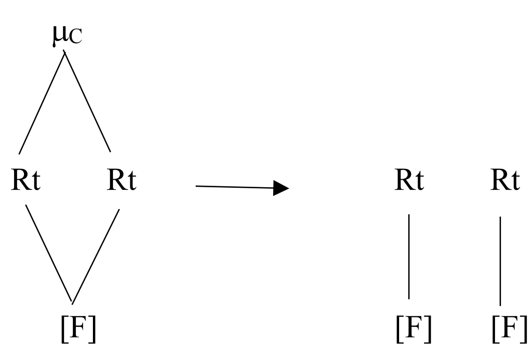
epenthesis. Yet, the forms with schwa epenthesis in (57) are in fact the default structures and not vice versa.

6. The fate of split geminates

We have previously argued for a hybrid phonological representation of geminates that combines both a moraic level and a two-root-node level. Particularly, in our model, a geminate is represented as two root nodes sharing a mora and a single place feature-set. In this chapter, we have been exploring the extent to which geminates preserve their integrity. In doing so, we have so far concluded that geminates are prone to breaking by means of morphologically induced vowels.

The question that arises from this fact is as follows: What form do geminates take when they get broken by vocalic material? Asked differently, how does vocalic breaking affect the phonological representation of geminates? In answering this question, we conjecture that when split up, the two root nodes constituting a geminate become independent of each other in two ways. First, each root node takes a copy of the feature set they have been sharing. Second, the mora dominating them gets erased after dissociating from both, by the stray erasure mechanism.

(59) The representation split geminates



In the pre-OT literature, the integrity of geminates is reflected by their representational properties. When a geminate is represented as a doubly-linked melodic element, its integrity is subject to the non-crossing association lines constraint, whereby a structure with two crossing lines is deemed ungrammatical. When a geminate is represented as a melodic element linked to a mora, its integrity arises from its built-in monolithic structure, in that nothing could infiltrate it as a unit segment. In OT, however, Geminate integrity is represented by a constraint that can be violated under the pressure of higher-ranked constraints.

7. Extending the past participle analysis

In section 4.3, we provided an account for the category of PPs derived from final geminated roots with the premise that geminate integrity in those forms is neutralized under morphological pressure. In what follows, the analysis proposed in that section would be extended to also account for the past participles derived from other base shapes. The next class to consider consists of PPs that are derived from tri-consonantal roots. Examples are provided below:

(60) məC.CuC:

Verb	*PP*	
ktəb	məktub	'write'
ḍrəb	məḍrub	'hit'
ʃṛəb	məʃṛub	'drink'
ʕrəḍ	məʕruḍ	'invite'
ħsəd	məħsud	'envy'
ʕṭəb	məʕṭub	'injure'

With no geminates in the picture, the interaction between our previously posited alignment constraints suffices to generate the right forms. This is illustrated by the following tableau:

(61)

√ktb /m, u/	ALIGN- (Rt, R, Pwrd, R)	ALIGN (m, L, u, R)
☞a. mək.tub		*
b. mək.t.bu	*!W	L

Considering the constraint hierarchy shown by this tableau, candidate (61b) is ruled out for failing to join the right edges of the root and the prosodic word. The optimal candidate in (61a) satisfies this requirement successfully. Seemingly, the possible candidate *[mək.təb] would seem to tie with the optimum, in that it too conjoins the right edges of the root and the prosodic word. However, this candidate can be discarded for the obvious reason that it does not realize the PP morpheme in its entirety by deleting the vowel [u], hence violating the faithfulness constraint MAX-Affix. Alhough it is not shown in our tableau, it should be noted that the independent ranking of PARSE-Seg over DEP-ə is responsible for the schwa epenthesis in the first syllable.

The class of PPs to consider next is derived from final weak triliteral roots (i.e. roots whose third segment is a vowel). Unlike all other classes, the PPs that belong to this class end with an open syllable. Also specific to this class is the fact that the [u] that appears with PPs from other classes is absent. We will see that these two characteristics are closely related. Consider the following illustrative items:

(62) məC.CV

Verb	PP	
kri	məkri	‘rent’
ʃri	məʃri	‘buy’
kwi	məkwi	‘weld’
ɣli	məɣli	‘boil’
kmi	məkmi	‘smoke’

In order to derive the appropriate forms constituting this class, two additional constraints are called for. These are as follows:

(63)

a. **MAX-Affix**: every element of the affix in the input must have a correspondent in the output.

b. ***High$^2_{word}$**: the repetition of the vocalic feature [+high] is banned in the domain of the prosodic word.

The role of the faithfulness constraint in (63a) is to ensure that the affix material in the input is fully preserved in the output. To account for the deleted [u] of the PP morpheme, this constraint would have to be dominated by some other constraint, which we believe to be the conjoined markedness constraint *High$^2_{word}$. Such a constraint penalizes the co-occurrence of two high vowels within the domain of the prosodic word. It is formalized using the mechanism of Constraint Conjunction (Smolensky, 1993, 1995, 1997), which enhances the effect of individual independently motivated well-formedness constraints by making their conjoined violation more serious that the violation of their single instantiations. This means that conjoined constraints are expected to always dominate their individual constituents (i.e. [C1&C2] >> C1, C2). The deletion of the vowel [u] is captured by the following constraint interaction:

(64)

√ʃri /m-u/	*High$^2_{word}$	MAX-Affix
☞a. məʃ.ri		*
b. məʃ.rui	*!	

Under the pressure of the high-ranking constraint *High$^2_{word}$, the vowel /u/ of the affix is deleted in the optimum, causing a less serious violation of the low-ranking Max-affix. Evidently, the vocalic material of the root is observed due to what we believe is the privileged status of root elements over affix elements. This means that *High$^2_{word}$ should be dominated by the faithfulness constraint MAX-Root, thus precluding the deletion of the root vowel. The following tableau summarizes the whole process of output selection:

(65)

√ʃri /m-u/	ALIGN- (Rt, R, Pwrd, R)	*High$^2_{word}$	MAX-Affix	ALIGN-Affix
☞ a. məʃ.ri			*	*
b. məʃ.riu	*!W	*	L	L
c. məʃ.rui		*!W	L	L
d. məʃ.ruj		*!W	L	L

In addition to violating *High$^2_{word}$, the sub-optimal candidate (65b) violates the alignment constraint on edges. Candidate (65c) satisfies the latter by infixing the [u], but still incurs a fatal violation of *High$^2_{word}$. Therefore, despite losing some of its morphological substance, candidate (65a) gets out of the competition victorious. The other candidate that should be considered is *məʃ.ruj. This candidate alters the consonantal specification of the root vowel, turning it into a glide. Glides are semi-vowels that are intrinsically [+high], too. Therefore, this candidate would also be ruled out by our conjoined constraint against the local repetition of high vocoids.

We now turn to the class of PPs whose bases are medial weak tri-segmental roots (i.e. roots whose second segment is a vowel). This class is characterized by turning the medial vowel of their base roots into a glide. The latter occupies the onset position of the second syllable.

(66) məC.juC:

Verb	*PP*	
biʕ	məbjuʕ	'buy'
dir	mədjur	'do'
ʃuf	məʃjuf	'see'
luħ	məljuħ	'throw'

This class of PPs shares with the previous one the fact that the high vowel of the root co-occurs with the high vowel of the affix in the same domain. However, the two classes differ in the way each one handles its double high vowel co-occurrence. Hence, the constraint set involved in deriving the PPs of the previous class (e.g. məʃri 'sold') yield the wrong forms in this particular case. This is illustrated by the following tableau:

(67)

√biʕ /m-u/	ALIGN- (Rt, Pwrd, R)	*High$^2_{word}$	MAX-Affix	ALIGN-Affix
a. məb.juʕ		*!W		*
b. məb.iuʕ		*!W		*
c. m.bi.ʕu	*!W	*		L
☹d. m.biʕ			*	L

In fact, in spite of comprising two high vowels, the class of PPs we are dealing with here does not resort to the deletion of the vowel [u]. Note that, unlike in the case of the previous class, deletion in this situation could give rise to forms with syllabic consonants (e.g.*m.biʕ). Therefore, deriving the right forms requires a constraint that could rule out this structure. The markedness constraint against syllabic consonants in MA is: *μ/C. Thus, the latter has to dominate *High$^2_{word}$ so that it can block deletion. This means that the grammar of MA prefers a structure with two high vocoids over one with a syllabic consonant. This can be summed up through the following ranking: *μ/C >> *High$^2_{word}$ >> MAX-Affix.

(68)

√biʕ /m-u/	ALIGN- (Rt, Pwrd, R)	*μ/C	*High$^2_{word}$	MAX-Affix	ALIGN-Affix
☞a. məb.iuʕ			*		*
c. m.bi.ʕu	*!W		*		L
d. m.biʕ		*!W		W*	L

Candidate (67a) is almost ready to become the optimal form. What needs to be accounted for now is the hiatus structure that we do not find in the output forms. Instead, the latter occur with a glide that corresponds with the high vowel of the root. To account for this alternation two more supplementary constraints are needed. These are *HIATUS and IDENT-Cons:

(69)

a. ***HIATUS**: assign one violation to every pair of adjacent vowels.

b. **IDENT-Cons**: correspondent segments in the input and output have identical values for consonantal.

The markedness constraint *HIATUS militates against sequences of vowels in adjacent syllables. This constraint is posited because a hiatus structure is formed when the vowel of the root and that of the affix are juxtaposed. Hiatus structures are cross-linguistically marked (Casali, 1996; 1997; 2011). Different languages employ different strategies to resolve hiatus structures. One of the most common hiatus resolving strategies is deletion, whereby one of the relevant vowels is deleted. Other strategies include consonant insertion, coalescence and glide formation. According to our data, the one strategy employed by the grammar of MA is glide formation. In particular, the high vowel of the root is turned to a glide to avoid a hiatus structure. Here comes the role of the faithfulness constraint IDENT-IO [Cons], which militates against the change of the consonantal specification of segments. Being outranked by *HIATUS, the demand made this constraint is overridden in favor of satisfying *HIATUS.

(70)

√biʕ /m-u/	*HIATUS	IDENT-Cons
☞ a. məb.juʕ		*
b. məb.iuʕ	*!	

As noted before, roots that constitute the base forms for this class of PPs contain a medial high vowel which can be either [+round] (i.e. /u/) or [-round] (i.e. /i/). Being the vowel at the edge of the second syllable, the root vowel turns into a glide, filling the place of an onset position. Typically, a [+round] high vowel corresponds to a [+round] glide (i.e. /w/). However, even in the case of roots whose medial vowels are [+round] (e.g. ʃuf 'see'), the formed glide is always [-round] (i.e. /j/). Following Boudlal (2001), we believe this to be yet another case of identity avoidance alternation, whereby a sequence of two adjacent [+round] vowels is not allowed. The constraints that we conjecture to be responsible for the observed alternation are *RdRd and IDENT-IO [round]. These interact as follows:

(71)

ʃuf /m-u/	*RdRd	IDENT-IO [round]
☞ a. məʃ.juf		*
b. məʃ.wuf	*!	

The last class of PPs to deal with is the one whose members are derived from quadrisegmental bases. These can be derived quadrisegmental bases like causatives or lexical quadrisegmental verbs. What uniquely characterizes the PPs belonging to this class is the fact that they are derived from an output form instead of a minimal root. Besides, this class is marked by the non-realization of the vocalic part of the PP morpheme (i.e. [u]). It will be shown that these two facts are connected. For illustration consider some examples below:

(72) m+CəCCəC

Output base	*PP*	
kəttəb	mkəttəb	'write'
dəwwəz	mdəwwəz	'pass'
wəlləf	mwəlləf	'accustomed'
ʕərri	mʕərri	'naked'
ṛəqqəq	mṛəqqəq	'slim'

As they get derived from an output form, it is suggested that the PPs cited above are subject to an output-output correspondence relation with their base forms. Output-output correspondence relations refer to a situation whereby morphologically related words are required to be phonologically identical (Benua, 1997). OO-correspondence relations may force a derived word to resist some rules to maintain its resemblance to its output base. On this note, the PPs under consideration are argued to deviate from the canonical form of the PP, which involves the affixation of the morpheme [m-u], by deleting the vocalic part of the morpheme to maintain a perfect match between the base forms and their corresponding derived forms. To accomplish this result, we posit the general faithfulness constraint FAITH-OO, which requires the output to be maximally faithful to the base form. The deletion of the vowel [u] of the PP morpheme takes place as a result of the following interaction between FAITH-OO and MAX-Affix.

(73)

Input: √ktb, m-u Base: [kət.təb]	FAITH-OO	MAX-Affix
☞a. m.kət.təb		*
b. m.kət.tub	*!	

This tableau demonstrates a competition between candidate (73b), which fully realizes the PP morpheme, and candidate (73a), which preserves the structure of the output base. Candidate (73b) is ruled out due its violation of FAITH-OO. Candidate (73a) emerges as the winner despite violating the dominated Max-affix. The subsequent tableau summarizes all the interactions involved in deriving the right forms:

(74)

Input: √ktb, m-u Base: [kət.təb]	FAITH-OO	ALIGN- (Rt- R, Pwrd, R)	ALIGN-Affix	MAX-Affix
☞ a. m.kət.təb				*
b. m.kət.tub	*W		*	L
c. m.kət.təb.u		*!W	L	L

This tableau evaluates an additional candidate, shown in (74c), which satisfies FAITH-OO by keeping the output base of the PP intact. However, this candidate incurs a fatal violation of the equally high-ranked ALIGN- (Rt- R, Pwd, R), which demands that the right edges of the root and the prosodic word should coincide. Therefore, the [u] deletes as it can neither stay at the right edge nor get infixed inside the PP form.

8. Conclusion

This paper has been devoted to the investigation of geminate integrity in MA. We have entertained an analysis that builds on Benhallam (1980), which claims that geminates in MA can be split by morphological rules but not by phonological ones. Cast in the constraint-based framework of OT, our account has been underlain by the basic idea that geminate integrity is a constraint that is not inviolable. As far as the effect of geminates on schwa epenthesis is concerned, we argued that the constraint on geminate integrity dominates the constraint responsible for word-to-syllable alignment: GEM-Integrity >> ALIGN-R. Under this ranking, schwa epenthesis gets blocked by geminates.

We have come to realize that schwa can split geminates on some occasions. We specifically treated the case of a class of broken plurals whose underlying geminates were shown to be prone to schwa epenthesis (e.g. sləl 'baskets'). In this regard, we have argued that there is a high-ranked plural Realize Morpheme (RM) constraint that dominates Gem-integrity: RM_{pl} >> GEM-Integrity. It has been shown that thanks to this ranking geminate integrity gets compromised for morphological purposes, namely the realization of the plural morpheme. Comparably, it was also demonstrated that when the full vowel [u] of the PP splits underlying geminates (e.g. ħəll vs. məħlul 'open'), they do so under morphological pressure, which consists in bringing together the right edges of the root and the prosodic word. This was captured by the ranking of the alignment constraint ALIGN-R (Rt, PWrd) over GEM-Integrity.

CHAPTER SEVEN

MOTIVATING THE ROOT IN MOROCCAN ARABIC: DERIVING THE CAUSATIVE AND THE COMPARATIVE

1. Introduction

The chapter examines the status of the root as a base of derivation in Moroccan Arabic morphology. This is mainly done for the purpose of paving the way to the following chapter, where we will be dealing with some aspects of word formation in MA. Another goal for this chapter is to define the status of geminates in the context of a root-based approach to morphology. To this end, in the course of this chapter, we will provide a set of empirical arguments that motivate the morphological function of the root (Cantineau, 1950; McCarthy, 1979, 1981). However, unlike the traditional view that sees roots as purely consonantal, we will be arguing for the existence of roots that can also involve vocalic material.

Evidence for the morphological status of roots is drawn from the formation of morphological causatives and comparatives in Moroccan Arabic. It is argued that a root-based approach to MCs and comparatives eschews problems of base inconsistency, base-to-output nonconformity and base unavailability. Such problems, however, are posed by a word-based approach (Ratcliffe, 1997, 1998, 2013; Benmamoun, 1999, 2003). More specifically, when a word-based approach is adopted, it is shown that (i) the causative could be inconsistently derived either from the perfective or imperfective verb form, (ii) some causatives and most comparatives randomly lose some of their bases' phonological material, and (iii) some causatives do not have any word-form corresponding base.

In what follows, a terse description of Arabic morphology under the auspices of Prosodic Morphology is provided (section 2). In Section 3, we review the arguments advanced against the adequacy of the root as a morphological entity. Section 4 highlights the inadequacies posed by a word-based approach to causatives (section 4.1) and argues for a root-based approach as an alternative (section 4.2). In section 5, the issue of the status of vowels in MA roots is addressed. Section 6 presents further evidence for the morphological status of the root from the derivation of the comparative form. Section 7 concludes.

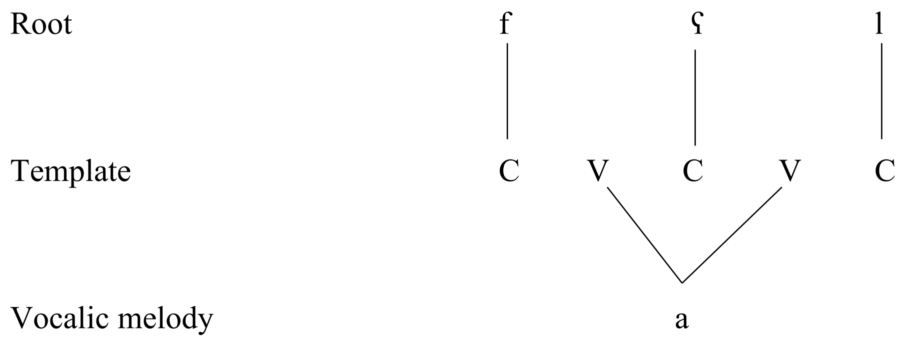
2. Setting the scene

Morphological models can be distinguished based on their units of analysis and the way they relate morphological forms to each other (Blevins, 2006). As a result, a major distinction is made between word-based systems and morpheme-based ones. In a word-based model, the grammatical word is viewed as the minimal unit of morphological analysis, while from the perspective of a morpheme-based model, a word is a combination of smaller meaningful sub-constituents (i.e. morphemes). Advocates of a word-based approach to morphology argue that morphological structure is much more diverse than simply putting morphemes together, in the sense that morphological meanings can be carried out by various other processes such as base modification, subtraction, metathesis, conversion and so on. Proponents of the morphemic approach, on the other hand, maintain that morpheme combination is cross-linguistically more common and allows for a restrictive architecture of description that unites morphology and syntax.

Arabic morphology has often been described as being nonconcatenative (McCarthy, 1979, 1981). That is, words are not constructed solely through a linear combination of discrete morphemes (e.g. English un-avoid-able), but also, and to a large extent, via interleaving discontinuous morphological forms (e.g. Arabic k-a-t-a-b 'write'). At the heart of this approach to Arabic morphology are the theoretical concepts of root and pattern (Cantineau, 1950). While the root, which is essentially tri-consonantal in nature, carries the basic lexical meaning of the word, the pattern expresses the grammatical meaning through a close-ended set of prosodic templates and vocalic melodies.[26] For this, every Arabic word would ideally comprise three types of morpheme: the root, the template and the vocalic melody. Such a characterization of Arabic morphology has become possible thanks to the representational mechanism of Autosegmental Phonology (Goldsmith, 1976), later developed by McCarthy (1979) for morphological systems, whereby features and morphemes can have independent lives on distinct tiers. The non-linear representation in (1) illustrates how Arabic words are represented under the purview of the Autosegmental framework:

[26] Arabic roots are mostly triliteral; however, there also exist biliteral and quadriliteral roots.

(1) Autosegmental representation of Arabic words



The above representation demonstrates the independence ascribed to each morpheme on its own tier. The root and the vocalic melody are then associated to the prosodic template. Later, the three morphemic levels are conflated into one linearly ordered string of segments to form the word ‘faʕal’. The citation root √fʕl ‘do’ stands for all the possible trilateral roots in Arabic (e.g. √ktb ‘write’ √ʃṛb ‘drink’ √ɦṛb ‘flee’…). The word form ‘faʕal’ is the simple perfective active verb form. Any change at the level of the template and/or the vocalic melody would yield a different word form. For instance, the vocalic melody /u-i/ would yield the perfective passive form ‘fuʕil’, and the prosodic template CVCCVC would produce the causative form ‘faʕʕal’. The verb forms ‘faʕal’, ‘fuʕil’ and ‘faʕʕal’ are semantically related as they share the same root, √fʕl.

As it stands, the root in Arabic morphology is attributed a pivotal morphological role as a morpheme upon which a multitude of grammatical constituents can be built. It is conceived of as the minimal meaningful lexical unit of word formation. Under this conception of the root, one could assume a lexicon that consists primarily of discontinuous roots, bearing a general meaning shaped and constrained by the patterns they associate with. (2) provides a specific example of the function of the root in Arabic, the case of the root √ktb in MA:

(2) Some derivatives of the root √ktb in MA

Form	*Gloss*
a. ktəb	'write'
b. kəttəb	'make write'
c. ktab	'book'
d. ktuba	'books'
e. katib	'writer'
f. məktub	'written'
g. məktəb	'office'
h. məktaba	'library'

Despite their different shapes and categories, the forms in (2) undeniably share the lexical meaning of 'writing'. The root √ktb remains unscathed in every form of the paradigm.

It will be shown that morphological causatives in MA can be accounted for simply by affixing a mora to a root. Hence, it is believed that any other approach with additional alternations would be less favorable.

3. The case against the root in Semitic

The morphological role of the root has been called to question, not only in Arabic (Ratcliffe, 1997, 1998, 2013; Benmamoun, 1999, 2003) but in other Semitic languages as well, such as Hebrew (Bat-El, 1994, 2003; Ussishkin, 1999). In these works, it has been shown that roots are inadequate to explain all cases of word formation, and that sometimes reference has to be made to other derived words as bases of derivation. Under this view, roots could exist only in abstraction whereby rules of word formation apply to them redundantly rather than generatively.

Ratcliffe (1997) suggests that a root and pattern approach to Arabic morphology proves insufficient to account for the derivation of word forms such as the plurals and the diminutives. Rather, these are better understood if Arabic morphology is viewed as a set of processes operating on words, which then lead to internal modifications in those words. Such a view stems from the fact that deriving the plural and the diminutive in Arabic involves the transfer of output material (e.g. affixes and vowels), which is not part of the consonantal root, between base and derived forms. As a result, Ratcliffe (1997:150) takes the following strong position: "But since some rules must operate on words, it is clearly preferable on grounds of simplicity to assume that they all do, and that the word, rather than the root is the listed form in the Arabic mental lexicon."

A word-based view of Arabic Morphology is also shared by Benmamoun (1999). Building on a number of syntactic arguments which indicate that the unmarked verb form in Arabic is the imperfective rather than the perfective, Benmamoun proceeds to demonstrate that there is also morphological evidence that the imperfective should be the default input for word formation in Arabic. Especially, the imperfective appears to be the ideal base form for the derivation of many nominal and verbal forms. For example, the vocalic melodies of many nominals can be traced back to their corresponding imperfective forms. Consider the following data for concreteness:

(3)

Imperfective		*Nominal*	
ʕallim	‘teach’	mu-ʕallim	‘teacher’
saʕid	‘assist’	mu-saʕid	‘assistant’
ʒlis	‘sit’	ma-ʒlis	‘council’
ʕriḍ	‘exhibit’	ma-ʕriḍ	‘exhibition’
sbaħ	‘swim’	ma-sbaħ	‘pool’

Likewise, the imperative form is believed to be derived from the imperfective given the strong phonological similarity they share. Some examples are shown below:

(4)

Imperfective	*Imperative*	
drus	ʔu-drus	‘learn’
qraʔ	ʔu-qraʔ	‘read’
ʒlis	ʔu-ʒlis	‘sit’

Evidence for the inadequacy of the root is also found in Hebrew. In light of issues related to cluster transfer in Modern Hebrew, which a theory of roots fails to account for, Bat-El (1994) decides to dispense with the notion of the consonantal root in the morphology of Modern Hebrew. To illustrate, the verbs derived from the base nouns praklit ‘lawyer’ and sandlar ‘showmaker’ are priklet ‘to practice law’ and sindler ‘to make shoes’, respectively. Note that it is not only the order of consonant that is observed, but also their adjacency. Therefore, the choice between the verb templates CCVCCVC (priklet) and CVCCCVC (sindler) appear to be dictated by the template of the base nouns. The theory of Root-to-Template Association clearly fails to predict which verb template is to associate with which root, since roots lack the information of consonant adjacency that characterizes each template.

Ussishkin (1999) presents more arguments that undermine the morphological status of the root in Semitic, in general, and Modern Hebrew, in specific. These arguments are based on observations about the derivation of denominal verbs, which shows that the nature of vowels in the nominal bases predicts the shape of the derived denominal verbs. For instance, biliteral denominal verbs, whose bases have two consonants, may take one of the following possible patterns (the data presented is from Ussishkin, 1999):

(5)

(a) $C_1iC_2eC_2$

cad	'side'	cided	'to side'
sam	'drug'	simem	'to drug, to poison'
dam	'blood'	dimem	'to bleed'

(b) C_1ijeC_2

tik	'file'	tijek	'to file'
bul	'stamp'	bijel	'to stamp'
ʔir	'city'	ʔijer	'to urbanise'

(c) C_1iveC_2

sug	'type'	siveg	'to classify'
ʃuk	'market'	ʃivek	'to market'
hon	'wealth'	hiven	'to capitalise'

(d) $C_1iC_2C_1eC_2$

hed	'echo'	hidhed	'to echo'
nam	'sleep'	nimnem	'to doze
daf	'page'	difdef	'to turn pages'

Ussishkin points to the fact that the shape of the majority of these denominal verbs can be predicted from the vowels of their nominal bases. To illustrate, the denominal verbs with consonant doubling are derived from nominals with the low vowel [a] (i.e. Pattern (a)). Also, the denominal verbs whose medial consonant is the glide [i] (i.e. Pattern (b)) are based on nominals whose medial vowel is high. Finally, denominal verbs whose medial consonant is the glide [v] (i.e. Pattern (c)) are derived from nominals with medial round vowels.

This chapter attempts to motivate a root-based approach to Moroccan Arabic (MA) causatives and comparatives. We argue that the derivation of the causative verb and the comparative adjective in Moroccan Arabic lends further support to the morphological status of the root. However, we by no means attempt to dismiss the fact that some derivational processes in Arabic have to be word-based as is the case with the broken plural in Classical Arabic (Hammond, 1988; McCarthy and Prince, 1990; Ratcliffe, 1990; Idrissi, 1997).

Unlike the sound plural, which is constructed through the suffixation of the masculine –uun or the feminine –aat to a singular form (e.g. musafir-uun 'travelers'), the broken plural in Arabic is formed mainly via the internal modification of the base. The derivation of the broken plural is characterized by the transfer of various qualities from the corresponding singular forms. McCarthy and Prince (1990) catalogues the following types of transfer:

(6)

(a) *Vowel length transfer*

Root	*sg.*	*pl.*	
/ʒndb/	ʒundub	ʒanaadib	'locust'
/slṭn/	sulṭ**aa**n	salaaṭ**ii**n	'sultan'

(b) *Prosodic size transfer*

Root	*sg.*	*pl.*	
/nfs/	nafs *bimoraic*	nufuus *disyllabic*	'soul'
/ʒndb/	ʒundub *trimoraic*	ʒanaadib *trisyllabic*	'locust'

(c) *Consonant spreading transfer*

Root	*sg.*	*pl.*	
/nwr/	nuwwaar	nawaawiir	'white flowers'
/ʒlb/	ʒilbaab	ʒalaabiib	'a type of garment'

(d) *Reduplication transfer*

Root	*sg.*	*pl.*	
/zl/	zalzal	zalaazil	'earthquake'
/ʒd/	ʒudʒud	ʒadaaʒid	'cricket'

(e) *Affix transfer*

Root	*sg.*	*pl.*	
/nbʔ/	janbuuʔ	janaabiiʔ	'spring'
/ftħ/	miftaaħ	mafaatiiħ	'key'

This data demonstrates that various phonological and morphological characteristics carry over from the singular to the plural. In view of these facts, it becomes obvious that the formation of the broken plural is word-based par excellence. The corresponding roots of the plurals mentioned above would be insufficient to produce the right forms in the absence of single unified broken plural template.

In fact, our main goal is to show the way in which a word-based approach to MA proves inadequate and a root-based approach emerges as a better alternative (see Noamane (2018) for a similar line of argumentation). The inadequacy of the word is justified by problems related to correspondence between the designated word forms and their corresponding derived causatives and comparatives, mainly the inconsistent choice of the aspectual form of base verbs, the loss of vocalic material of some base nouns and adjectives, and finally the lack of base forms for certain causatives. A theory of roots, however, has the following advantages: (i) roots are not specified for aspectuality (ii) all root material appears in the derived forms (iii) every causative can have some root.

4. Root-based vs. word-based approaches to Morphological Causatives

As far as deriving morphological causatives is concerned, the choice of the base form remains a recalcitrant issue. The main controversy revolves around whether morphological rules apply to words or roots. Two possible hypotheses suggest themselves. Under one hypothesis, causative verbs in MA are derived directly from roots whose syntactic categories and morpho-phonological information are not specified yet, hence the name the root-based hypothesis. Under the other hypothesis, the base forms of causatives are fully specified words for their syntactic categories and morpho-phonological information such as verbs, nouns and adjectives. Here, we argue in favor of the root-based approach.

4.1. Demotivating a word-based approach

In a purely word-based approach, causative verbs in MA would be derived from simple verbs, nouns, adjectives and sometimes comparative forms. As Aronoff (1976:21) suggests, "all regular word formation processes are word-based. A new word is formed by applying a regular rule to a single already existing word. Both the new and the existing one are members of major lexical categories". Building on this idea, the larger portion of causatives would derive from verbs. Each verb belongs to one of the following three major classes of segmental shape: strong tri-literal verb bases (e.g. ʃṛəb 'to drink', xdəm 'to work'), middle-weak tri-literal verb bases (e.g. fiq 'to wake up', gul 'to say') and final-weak tri-literal verb bases (e.g. ʒri 'to run', mʃi 'to walk').

Verbs of the first class are characterized by containing only consonants and no vowels. Some examples are provided in (7), where schwa is phonologically motivated to break up impermissible consonant clusters both in the base and the derived form (Benhallam 1989/1990; Boudlal, 2001; Bensoukas and Boudlal, 2012a-b; see chapter #6 for a thorough account of schwa epenthesis in MA):

(7) The verb base form: the strong triliteral bases

Base form		*Causative verb*
zʕəm	'to have courage'	zəʕʕəm
ṛħəl	'to leave'	ṛəħħəl
fzəg	'to get wet'	fəzzəg
nʒəħ	'to succeed'	nəʒʒəħ
wṣəl	'to arrive'	wəṣṣəl
rbəħ	'to win'	rəbbəħ
nʕəs	'to sleep'	nəʕʕəs
sɦəṛ	'to stay up late'	səɦɦəṛ
nʃəf	'to get dry'	nəʃʃəf
xdəm	'to work'	xəddəm

Verbs of the second class are marked by containing one of the full vowels /i/ or /u/ in the imperfective form, and the vowel /a/ in the perfective, in the middle of the verb, as shown in (8) below:

(8) Middle-weak triliteral bases

	Base form		*Causative verb*
a.	fiq	'to wake up'	fəjjəq
	ṭiħ	'to fall down'	ṭəjjəħ
	ṭir	'to fly'	ṭəjjər
	mil	'to lean'	məjjəl
	ʕif	'to get disgusted'	ʕəjjəf

b.	nuḍ	‘to get up’	nəwwəḍ
	gul	‘to say’	gəwwəl
	dux	‘to feel dizzy’	dəwwəx
	ʕum	‘to swim’	ʕəwwəm

What is unique about the group of items in (8) is the alternation between the high vowels /-i-/ and /-u-/ in the base and the geminate glides /-jj-/ and /-ww-/ in the derived causative. That is, deriving the causative here is not limited to lengthening the second segment of the base, but in addition, there is a change at the level of sonority, where a vowel becomes less sonorant by turning into a geminate glide. A fact like this can be straightforwardly explained since it is widely acknowledged that the high vowels /i/ and /u/ alternate with their glide counterparts /j/ and /w/, respectively, in many languages of the world (Rosenthall, 1994; Padgett, 2008). We account for this by assuming that the causative morpheme is a consonantal mora that turns any high vowel it attaches to into a corresponding geminate glide.

4.1.1. Base inconsistency

The third class, however, involves verbs that end with one of the vowels /a/ or /i/.[27] Causatives derived from these verbs have in common the pattern CəCCV. This is put in evidence by the following examples:

[27] Verbs ending in /u/ are unattested in MA. Hence, we assume that this is a lexical gap in the grammar of the language.

(9) Final-weak tri-literal verb base

	Base form		*Causative verb*
a.	qṛa	'to read'	qəṛṛi
	dfa	'to become warm'	dəffi
	nsa	'to forget'	nəssi
	ʕja	'to get tired'	ʕəjji
	wʕa	'to get conscious'	wəʕʕi
b.	ɣli	'to become boiled'	ɣəlli
	ʒri	'to run'	ʒərri
	mʃi	'to leave'	məʃʃi
	bki	'to cry'	bəkki

The base verbs in (9) are all in the imperfective aspectual form as those in (8). In addition to the singleton-geminate alternation between them and their corresponding causative forms, one can also notice that the vowel /a/ in the base verbs of (9a) alternates with the vowel /i/ in causatives. The base verbs in (9b), however, all end with the vowel /i/, which is preserved in the derived causatives. Such a problem arises as we try to be consistent in positing the imperfective aspectual form as the basic form of derivation across all the above-mentioned verb categories. One way to go around this in a word-based approach is to turn all the base verbs in (9) to the perfective form in which case they would all end with /a/ as would their corresponding causatives.

In this case, we can establish a perfect correspondence between vowels, but the produced causatives would appear in a different aspectual form from the causatives in (8) since those would be in the imperfective form to maintain the correspondence between the vowels /i/ and /u/ in their bases and the corresponding glides. This brings to the fore the question of what the right aspectual verb form is that would yield the appropriate correspondence between the base and the derived form, and which should be a form that would cause the least possible random alternations.

To address this issue, a holistic view of the data treated should be taken. First, all the base verbs listed in (7), (8) and (9) are consistently given in the imperfective, and as a result, the derived causatives also occur in the imperfective form. We could have posited the perfective form instead as basic; hence we would have no problem in accounting for the /a/-/i/ alternation that shows in (9), because then all the base verbs in (9) would have the vowel /a/ at the end, and the causatives would appear with the same vowel, as illustrated below:

(10) The perfective choice:

Base form		*Causative verb*
qṛa	'to read'	qəṛṛa
dfa	'to become warm'	dəffa
ʒra	'to run'	ʒərra
mʃa	'to leave'	məʃʃa

However, if we pursue this option, by assuming that base verbs and causative verbs alike should take the perfective form, another problem would emerge. In this case, it would be hard to account for the alternation between the vowels of middle-weak verbs and the glides of causative verbs in (8) since the base forms would all appear with the vowel /a/ and the causatives would have glides with unknown origin. For more clarity, consider the items in (11):

(11) The perfective choice

Base form		*Causative verb*
faq	‘to wake up’	fəjjəq
ṭaħ	‘to fall down’	ṭəjjəħ
naḍ	‘to get up’	nəwwəḍ
gal	‘to say’	gəwwəl

In effect, the reason the imperfective is chosen as basic in the first place is basically motivated by the items in (8), where an alternation between high vowels and glides takes place, for it is more natural and cross-linguistically common to have such an alternation.

4.1.2. Base-to-output nonconformity

In addition to verbs, causatives in MA can also get derived from nouns and adjectives. This is referred to as denomination, whereby nouns and adjectives become verb forms. Being the base form in this case, nouns and adjectives would function as the causee-event in the causative structure. In this respect, a given causative verb whose base form is a noun or an adjective would have the meaning “to cause someone/something to become Noun/Adjective”, or simply to cause that noun or adjective directly, i.e. “to cause N/Adj”. By way of illustration, a number of examples are listed in (12) for nouns and (13) for adjectives:

(12) The noun base form

Base form		*Causative verb*
ħməq	‘crazy person’	ħəmməq
təlʒ	‘ice’	təlləʒ
kəħt	‘misery’	kəħħət

məṛqa	'broth'	məṛṛəq
kuṛa	'ball'	kəwwəṛ
duṛa	'circle'	dəwwəṛ
rwina	'mess'	rəwwən
slaħ	'weapon'	səlləħ
ɣlaf	'cover'	ɣəlləf
ṭlaq	'divorce'	ṭəlləq
dlala	'auction'	dəlləl
ʕadab	'torture'	ʕəddəb
ɣawt	'screaming'	ɣəwwət
ħəlwa	'candy'	ħəlli
jtim	'orphan'	jəttəm
ṣabun	'soap'	ṣəbbən
fʃuʃ	'pampering'	fəʃʃəʃ

(13) The adjective base form

Base form		*Causative verb*
ṣʕib	'difficult'	ṣəʕʕəb
ṣɣir	'small'	ṣəɣɣər
wasəʕ	'wide'	wəssəʕ
ʕwəʒ	'curved'	ʕəwwəʒ
rṭəb	'smooth'	rəṭṭəb
kħəl	'black'	kəħħəl
ħməṛ	'red'	ħəmməṛ
qwi	'strong'	qəwwi
nqi	'clean'	nəqqi
ṣafi	'clear'	ṣəffi
ħafi	'blunt'	ħəffi
saħi	'sober'	səħħi

Once again, the problem of vowel alternation between the base forms and the derived forms comes to the surface. In this context, not only do some vowels alternate with others, but there are base forms whose vowels do not appear in the causative form. For example, the causative verbs [kəwwəṛ] 'to ball', [rəwwən] 'to cause a mess', [jəttəm] 'to orphan', [ṣəʕʕəb] 'to make difficult' and [ħəffi] 'to make blunt' from the nouns and adjectives [kuṛ**a**] 'a ball', [rw**i**n**a**] 'a mess', [jt**i**m] 'orphan', [ṣʕ**i**b] 'difficult' and [ħ**a**fi] 'blunt', respectively, do not preserve the vowels of their bases. In this case, we would need additional ad-hoc (i.e. unnatural) rules to account for these deletions. For adjectives, however, a word-based approach can resort to the comparative form to minimize the scale of alternation between some base forms and causative verbs, as shown in (14):

(14) The comparative base form

	Base form		*Causative verb*
a.	ṣʕəb	'difficult'	ṣəʕʕəb
	ṣɣər	'small'	ṣəɣɣər
	wsəʕ	'wide'	wəssəʕ
b.	ṣfa	'clear'	ṣəffi
	ħfa	'blunt'	ħəffi
	sħa	'sober'	səħħi

Nevertheless, the issue of random vowel alternations persists, as illustrated in (14b), where it is the vowel /a/ of the comparative forms that corresponds to the vowel /i/ of the causatives. The comparatives in (14b) are posited to substitute the adjectives [safi], [ħafi] and [saħi] as base forms whose internal vowel does not show up in their corresponding causatives. Yet, both forms seem to be inadequate.

4.1.3. Base unavailability

To cap it all, there are causative verbs that do not seem to have any corresponding base word forms. If a word-based approach were enforced, these causatives would remain baseless and treated as underived items. The mere fact of being baseless means that causatives are derived from some forms other than complete words. If a root-based approach is advanced, a unified and consistent view will be shared by all the morphologically derived causatives. Some examples are provided in (15):

(15) Baseless causative verbs:

xəbbi	‘to hide’
ṛəbbi	‘to bring up’
ʕəlləq	‘to hang up’
wərri	‘to show’
bəddəl	‘to change’
ṣəwwəṛ	‘to take a photo’

On this basis, it is assumed that causatives are derived from a variety of forms, including verbs, nouns and adjectives. However, in the subsequent section, we will be capitalizing on the above-mentioned irregularities to defend a root-based approach to deriving the causative.

4.2. Motivating a root-based approach

The root has been defined as a semantic minimal core element «which remains invariant when all identifiable morphological formatives have been abstracted away» (Acquaviva, 2009). It is characterized by being lexically non-decomposable and phonologically abstract. When combined with phonological and morpho-syntactic features, roots build up larger constituents like words (Arad, 2005). As noted by Ussishkin (2006), linguistic theory would favor an economic view in which the lexicon stores only roots. A root-based approach means that storage would be limited to abstract, idiosyncratic and non-redundant information. As for the burden of explaining the alternations between roots and surface forms, it remains the task of the grammar. Put differently, the human capacity for acquiring a Semitic language like MA would be much simpler if word derivation were based on roots (see Prunet (2006) and Idrissi et al. (2008) for psycholinguistic evidence in favor of the morphological status of roots).

As noted earlier, it has long been observed that Semitic languages display different linguistic traits in comparison with the Indo-European languages. In Semitic languages, words are not formed on the basis of isolable strings of segments that are linearly concatenated. Rather, they consist of consonantal roots indicating their lexical meaning and vocalic patterns, which signify grammatical categories. This entails that consonantal roots constitute the basic ground upon which words are derived. Change in form or category is due to the insertion of vocalic patterns or the affixation of some consonantal material which either modifies or expands the basic meaning contributed by the root.

Adding to this general property of Semitic languages, the motivation for a root-based approach to causatives in MA is further corroborated by the irregularities that are shown by a word-based approach and that a root-based approach eschews. As we have already noted, a word-based approach towards causatives is laden with random variation and inconsistency between the relevant word forms and the derived causatives, which weakens the prospect of such an approach.

To start with, it is worth stating that the root in MA is not a pure abstraction. In fact, the dominant pattern of word forms, verbs in particular, is one where the only material present is the consonantal root as is the case with most triliteral verbs. The interesting point is that it is only in this category of base verbs and their causative counterparts where none of the problems discussed earlier are posed. This shows that the causative is derived simply by lengthening the second segment of the base without any further vocalic alternations or deletions. For convenience, this is schematized as follows:

(16) Deriving causatives from strong triliteral roots

Base	*Mora affixation*	*Misalignment*	*Output form*
√ktb	μ+ktb	kμ+tb	kəttəb

In an attempt to limit the irregular alternations between certain word forms and their corresponding causatives in a word-based approach, Bennis (1992) suggests "la condition de généralisation variée" -*the condition of varied generalization*-, whereby all the segments of the base form should have corresponding elements in the derived form. According to Bennis (1992), this condition allows a word-based approach to choose the appropriate base form, regardless of its grammatical category, to control any possible unnatural alternations.

For example, in (17) below, the base form chosen is one whose segments completely correspond with the segments of the derived causatives. That is, if an adjective contains a vowel that does not show up in the output, it should be discarded from being the base form and instead the comparative form is posited as basic. The same applies to verbs. The base verb can be in the imperfective for medial weak verbs or the perfective for final weak verbs.

(17) The condition of varied generalization (Bennis, 1992)

a. wsaʕ (V); wasəʕ (Adj); wsəʕ (Adj.Comp) → wəssəʕ

b. rṭab (V); rṭəb (Adj); rṭəb (Adj.Comp) → rəṭṭəb

c. naḍ (perf.); nuḍ (imperf.)/ → nəwwəḍ

d. mʃi (imperf.); mʃa (perf.) → məʃʃi

e. nsa (imperf.)/ nsa (perf.) → nəssa

However, this means that the data will be presented in a random and inconsistent way, in the sense that the causative verbs will be derived from the perfective form, only when there is a need for maintaining the vowel /a/ as a common segment between the base and the causative. Otherwise, causatives are derived from the imperfective form so as to establish the high vowels' alternation with glides. For example, the word-based treatment of Bennis (1992) posits perfective verbs like [nsa] 'to forget' as the base form to get causatives of the form [nəssa] 'to cause to forget', as opposed to [nəssi]. In other instances, the imperfective forms like [nuḍ] 'to get up' are postulated instead as the base form to get causatives of the form [nəwwəḍ] ‘to cause to get up’; what we do not know is whether it is going to re-express its base and be imperfective as well, or to express the aspectual form of the other causative forms and thus be perfective.

For the sake of consistency, a word-based approach should either posit all the base forms in the perfective and face the /a/-glide alternation in cases like [naḍ]→ [nəwwəḍ], or maintain that the base forms are all imperfective and give us imperfective causatives, and hence face the /a/- /i/ alternation in cases like [nsa]→ [nəssi]. Instead, such an approach keeps moving from one grammatical form to another in order to shun those problematic alternations in every way possible. As it stands, this condition seems to be far from being psychologically and cognitively economical since it means that the derivational mechanism should simultaneously consider a couple of possible base forms and then look for the most appropriate one.

4.2.1. Underspecified roots

In a theory of roots, however, we propose that all causatives are derived from roots whose aspectuality is not specified yet. This way, every medial weak root will appear with one of the underlying vowels /i/ or /u/. The causatives derived from this category of roots will have corresponding geminate glides /jj/ or /ww/. As for final weak roots, we purport that they all end with the vowel /i/. The causatives derived from these will have a corresponding vowel /i/ as well. For illustration, consider the following examples:

(18)

	Root	Imperfective	Perfective	Causative	
a.	√ṭir	ṭir	ṭar	ṭəjjər	‘fall’
	√gul	gul	gal	gəwwəl	‘say’
b.	√ʃri	ʃri	ʃra	ʃərri	‘buy’
	√mʃi	mʃi	mʃa	məʃʃi	‘go’
c.	√nsi	nsa	nsa	nəssi	‘forget’
	√qṛi	qṛa	qṛa	qəṛṛi	‘teach’

In a word-based approach, the starting point of causative derivation is either the imperfective form or the perfective one, depending on the correspondence to be maintained. A theory of roots allows us to go back further and posit underspecified roots, which can guarantee us a systematic correspondence without causing any inconsistency. The verbs in (18c) appear with the low vowel /a/ both in the perfective and imperfective. This vowel is believed to be strictly related to perfective forms in (18a) and (18b) verbs. Nonetheless, a closer look at the derivatives of the problematic items in (18c) shows us that the underlying vowel is in fact the high vowel /i/:

(19)

Verb	*Causative*	*Agent Noun*	*Action Noun*	*1sPST*		*Root*
nsa	nəss**i**	nəssa**j**	nəs**j**an	ns**it**	→	√nsi
qṛa	qəṛṛ**i**	qəṛṛa**j**	qṛa**j**a	qṛ**it**	→	√qṛi

On the basis of this, it is safe to say that the verbs [nsa] and [qṛa] have roots that end with /i/ since the latter appears in other derived forms, though sometimes as the glide /j/. The /a/ in the agent and action nouns is specific to their structure and has nothing to do with the /a/ that appears in [nsa] and [qṛa].

Accordingly, roots would be underspecified for aspectuality. When verbs are derived, they then get inflected for the imperfective or the perfective. We assume that the imperfective marker in MA is a zero morpheme. Thus, verbs in the imperfective happen to be identical to the roots they are derived from. When it comes to deriving the perfective form, however, we believe that it is morphologically marked by the vocalic feature [-high], which changes the high vowels of both medial weak roots and final weak roots into the low vowel /a/. As for strong triliteral roots, they have no vowels that can carry this feature, hence the perfective form of their corresponding verbs is interpreted as a zero morpheme as well.

4.2.2. Minimal roots

In response to the problem of vowel loss that characterizes causatives derived from certain nouns in a word-based approach, we propose to consider minimal roots that contain only segments that are consistently carried over by the derived causatives. By doing so, the vowels that mark the feminine or nominal forms in the relevant nouns are left out and only the segments that appear in the causative form (and other derived forms) are sustained in the shape of a root.

(20) *Root* *Causative*

√mṛq	'broth'	məṛṛəq
√kuṛ	'ball'	kəwwəṛ
√duṛ	'circle'	dəwwəṛ
√rwn	'mess'	rəwwən
√slħ	'weapon'	səlləħ
√ɣlf	'cover'	ɣəlləf
√ṭlq	'divorce'	ṭəlləq
√dll	'auction'	dəlləl
√ʕdb	'torture'	ʕəddəb
√ɣwt	'screaming'	ɣəwwət
√ħli	'candy'	ħəlli
√jtm	'orphan'	jəttəm
√ṣbn	'soap'	ṣəbbən
√fʃʃ	'pampering'	fəʃʃəʃ

The result is that the causative derivation becomes simpler. It all boils down to affixing a consonantal mora to an underspecified minimal root without having to account for the deleted vowels.

As for the irregularity between some adjectives and their causative counterparts, a word-based approach resorts desperately to the comparative form, as a new base of derivation, to overcome those types of alternations illustrated in (13). This means that the lexicon has to store more word forms and mark them as [+causative]. Obviously, it does not seem to be an economical strategy to postulate more word forms as bases to avoid alternations that we are not obliged to avoid if roots are regarded as base forms instead.

To extend this argument, sometimes, it appears that causative verbs can be derived from two different possible base forms, say, a noun and an adjective. For example, the causative [ħəmməṛ] 'to make red' can be said to be derived from the noun [ħməṛ] 'red' or the comparative form [ħməṛ] 'redder', which happen to be homophonous. This means that the human computational system would get confused upon which form the relevant causative is derived from. In the case of roots, no such anomaly is expected to occur since there is only one base form that supplies all the possible surface forms. Therefore, it seems to be more feasible to claim that the causative [ħəmməṛ], along with the comparative [ħməṛ] and the noun [ħməṛ], is derived from the root √ħmṛ as a unique and non-redundant form.

4.2.3. Abstract roots

The third advantage of a theory of roots is that every word form can be said to have a minimal abstract core. For those causatives that lack base word forms, the root comes in handy to function as an ultimate base form. In a word-based approach, it could be argued that the structure of such causatives is base-generated. However, when you look at derivatives related to them, you notice that the geminate they contain appears as a singleton elsewhere, which suggests that those geminates must have been derived as well via morphological lengthening.

(21)

Causative		*Action Noun*	*Root*
xəbbi	'to hide'	txəbja	√xbi
ṛəbbi	'to bring up'	tṛəbja	√ṛbi
ʕəlləq	'to hang up'	tʕlaq	√ʕlq
wərri	'to show'	twərja	√wri

bəddəl	'to change'	tbdal	√bdl
ṣəwwəṛ	'to take a photo'	tṣwaṛ	√ṣwṛ

As a result, our main claim is that no derivational relation holds between word forms and the causatives that are semantically related to them. In fact, all word forms in MA are most likely to be derived directly from invariant roots that are abstract and underspecified, unless proven otherwise. Hence, any material in surface forms is not the result of an alternation of some other material; however, it is the outcome of the derivational system which generates forms directly from roots.

After having argued for a root-based approach to causative derivation, we now comment on some of the arguments that were advanced against such an approach. It has been shown elsewhere (Imouzaz, 1991), cited in Bennis (1992), that the root hypothesis displays an excessive generative power, in the sense that non-attested forms could be produced from roots, like the case of /qtl/ 'kill'→*[qəttəl]. Still, the question is whether a word-based hypothesis gives a better account since nothing prevents the form *[qəttəl] to be derived from the verb [qtəl]. In both cases, there is a need for a [+causative] feature to mark off those base forms that give rise to causatives from those that do not.

In addition, a variety of other arguments were meant to refute the root-based approach to causative derivation, without posing any serious problems to this view. This includes the following (see Bennis, 1992:84):

(22)

(i) The root cannot predict whether it can be geminated for causatives or not.

(ii) Predictable properties such as [+verb], [+noun], [+adjective] and [+causative] cannot be predicted by the root.

(iii) The root does not allow a distinction between causatives derived from verbs and causatives derived from nouns and adjectives.

Again, all these statements pose no embarrassment for the root-based hypothesis since it is taken for granted that the root should not predict any of the above-stated facts. It has already been assumed that the lexicon stores only abstract and idiosyncratic information, whereas the derivational system accounts for whatever appears in the output forms. For these reasons, the hypothesis proposed in this study is that morphological causatives in MA are derived directly from the root via geminating the second radical element.

5. The status of geminates in the root

Harrell (1962) identifies three basic root types in MA, namely triliteral roots, quadriliteral roots, and atypical roots. This typology is based on the number of constituent segments that each root incorporates. For triliteral roots, the number is limited to three constituent elements, for example √ḍrb 'hit', √ɦṛb 'escape' and so on. Second, quadriliteral roots consist of four constituent elements such as in √ṭrʒm 'translate' and √krkb 'roll'. Finally, atypical roots have more than four or less than three constituent elements, as in the word √ma 'water' or √ʒa 'come'. Each of the aforementioned root types can be either strong or weak. Strong roots comprise only consonants, whereas weak ones may include a vowel or a glide.

Also, in the context of geminates, a distinction is made between sound roots and doubled roots (Harrell, 1962). Sound roots consist of consonants that are different from each other. Doubled roots, on the other hand, are casually described as involving a pair of identical consonants. The existence of doubled roots follows from the fact that geminates in MA are underlying. This means that these roots would be associated to words whose geminates phonemically contrast with the singletons of other words. Some examples of geminated roots are provided below:

(23) Geminated roots in MA

ʒṛṛ	'drag'
ħll	'open'
sdd	'close'
ʃdd	'catch'
ʕḍḍ	'bite'
ṛqq	'thin'
dqq	'knock'
dmm	'blood'
ʕʃʃ	'nest'

In the course of this dissertation, it has been shown that geminates are much more complex than a sequence of two identical consonants. Instead, they are conceived of as moraic two-root node structures (see chapter 3 for details about the representation of geminates in MA).

6. The status of vowels in the root

Traditionally, roots in Semitic languages are assumed to be purely consonantal (Cantineau, 1950). That is, roots are viewed as sequences of discontinuous consonants, loosely associated with a specific lexical meaning. Vowels, on the other hand, come in to mark grammatical functions. Contrary to this classical view of roots in Semitic, we claim that roots in MA could be made up of consonants and vowels alike. Such a view has been argued for in relation to other Afroasiatic languages like Amazigh (Bensoukas, 2001, 2018). This can be based on the observation that certain vowels are consistently carried over by derived word forms sharing the same root. Generally, it will be argued that roots in MA contain no more than one single vowel. These vowels may occur either root-medially or root-finally, as shown in (24) below, where the following abbreviations apply: Imp= imperative; Impf= imperfective; Pf= perfective; Pv= passive; Caus= causative; AcN= action noun, AN= Agent noun; IN= instrument noun.

(24)[28]

I. Final vowels paradigm

a. Verb morphology

Imp	*Impf*	*Pf*	*Pv*	*Caus*
bk**i**	bk**i**	bk**a**	təbk**a**	bəkk**i**
ʃr**i**	ʃr**i**	ʃr**a**	təʃr**a**	ʃərr**i**
mʃ**i**	mʃ**i**	mʃ**a**	təmʃ**a**	məʃʃ**i**

[28] For lack of space, glosses will be given here: bki 'cry'; ʃri 'buy'; mʃi 'walk'; fiq 'wake up'; ʕum 'swim'; dir 'do'

b. Noun morphology

AcN	*AN*	*IN*
bək**j**an	bəkka**j**	_
ʃər**j**an	ʃərra**j**	_
məʃ**j**an	məʃʃa**j**	məʃʃa**j**a

II. Middle vowels paradigm

a. Verb morphology

Imp	*Impf*	*Pf*	*Pv*	*Caus*
f**i**q	f**i**q	f**a**q	_	fə**jj**əq
ʕ**u**m	ʕ**u**m	ʕ**a**m	tʕ**a**m	ʕə**ww**əm
d**i**r	d**i**r	d**a**r	td**a**r	_

b. Noun morphology

AcN	*AN*	*IN*
f**i**qan	fə**jj**aq	_
ʕ**u**man	ʕə**ww**am	_
d**j**ar	_	_

In these two paradigms, it is observed that the relevant verbal and nominal morphological classes share and retain certain vocalic material. The latter is believed to be part of a minimal root from which all these forms are derived. We maintain that some roots have medial vowels, while others have final ones. More specifically, both the high vowels /i/ and /u/ can occur root-medially whereas only the front high vowel /i/ can appear root-finally. The low vowel /a/, however, is never a part of roots. The quality of root vowels changes due to various morphological or phonological processes. For instance, they may turn into a vowel /a/ to mark the perfective form for verbs or to a corresponding geminate glide to derive causatives or agent nouns. They may also alternate with glides to improve syllable structure (e.g. resolving a hiatus).

With this in mind, one could claim that causative verbs in MA are derived directly from the root. In particular, causatives are derived from tri-segmental roots, which can be strong or weak. Accordingly, the data here is assumed to fall into the following categories:

(25)

a. Strong roots

√fzg	‘wet’	fəzzəg
√nʕs	‘sleep’	nəʕʕəs
√mṛq	‘broth’	məṛṛəq
√tlʒ	‘snow’	tølləʒ
√wsʕ	‘wide’	wəssəʕ
√ħmṛ	‘red’	ħəmməṛ

b. Final weak roots

√kmi	‘smoke’	kəmmi
√bki	‘cry’	bəkki
√xwi	‘emptiness’	xəwwi
√ɣli	‘boil’	ɣəlli

c. Middle weak roots

√fiq	'wake up'	fəjjəq
√ṭir	'fly'	ṭəjjər
√dux	'feel dizzy'	dəwwəx
√ʕum	'swim'	ʕəwwəm

To sum up, it has been shown that a word-based approach to causatives is inadequate, in that it faces many challenges related to correspondence between word forms and their corresponding causatives. Alternatively, a root-based approach has been defended. A theory of roots has allowed us to rid the data of any random and unnatural alternations by positing underspecified, minimal and abstract roots as bases of derivation. These roots have been shown to be of two main categories: strong or weak.

6. The comparative: Further evidence for the morphological status of the root

This section provides supplementary evidence in support of the active morphological status of the root in Moroccan Arabic morphology. Davis (2016) shows that the formation of comparative adjectives in Egyptian Arabic (EA) is better explained by a root-based derivation. In EA, the derivation of comparatives depends solely on the nature of the underlying root, which happens to be the only morphological material that gets mapped onto the comparative template aCCaC / aCaCC (e.g. akbaṛ 'bigger'; araqq 'thinner'). This means that, unlike in the case of the broken plural, there are no transfer relations between the comparative form and its corresponding base adjective form.

With this in mind, we would like to demonstrate that, despite the glaring phonological and morphological differences between EA and MA, the facts about the formation of the comparative form in MA corroborate Davis (2016), such that only the root material gets carried over from base adjectives to the comparatives. This is put into evidence by the following data:

(26) Deriving the comparative in MA

Root	*Adj.*	*Comp.*	
√kbr	kbir	kbəṛ	'big'
√tql	tqil	tqəl	'heavy'
√ɣlḍ	ɣliḍ	ɣləḍ	'fat'
√ṭwl	ṭwil	ṭwəl	'tall'
√ʕṛḍ	ʕṛiḍ	ʕṛəḍ	'wide'
√qsṛ	qsiṛ	qsəṛ	'short'
√smn	smin	smən	'obese'
√ṛxs	ṛxis	ṛxəs	'cheap'

The data in (26) shows that the derivation of the comparative forms in question is based on their corresponding minimal roots rather than their equivalent base adjectives. This is corroborated by the non-transfer of the pre-final vowel [i] of the base adjectives. Instead, the comparatives in (26) consist exclusively of the tri-segmental content of their corresponding roots. To further emphasize this point, let's consider additional data:

(27)

Root	*Adj.*	*Comp.*	
√sɦl	saɦəl	sɦəl	'easy'
√ʕmṛ	ʕaməṛ	ʕməṛ	'full'
√xnz	xanəz	xnəz	'smelly'
√qsħ	qasəħ	qsəħ	'hard'
√bsl	basəl	bsəl	'tasteless'

Likewise, the derivation of the comparatives in (27) ignores the vocalic material of the corresponding base adjectives. In fact, these comparatives are comprised solely of the root material of their corresponding adjectives, proving once again that the comparative form is root-based. A root-based approach to comparatives is further substantiated by the following class of comparative forms:

(28)

Root	*Adj.*	*Comp.*	
√ṛqq	ṛqiq	ṛəqq	'thin'
√ṣħħ	ṣħiħ	ṣəħħ	'strong'
√xff	xfif	xəff	'light'
√ħnn	ħnin	ħənn	'affectionate'
√qll	qlil	qəll	'scarce'
√bnn	bnin	bənn	'tasty'

The comparatives in (28) are characterized by containing final geminates whose integrity is preserved, matching the integrity of the geminates of their corresponding roots. In the base adjectives, however, the integrity of the relevant geminates is compromised, hence the information about the integrity of the geminates is lost. This clearly demonstrates that the comparatives of this class are also derived directly from their corresponding roots. An additional argument in favor of the active status of the root in the formation of the comparative is shown by the following items:

(29)

Root	*Adj.*	*Comp.*	
√wsx	musəx	wsəx	'dirty'
√fid	mufid	fjəd	'interesting'
√smm	məsmum	səmm	'poisonous'
√zħm	mzaħəm	zħəm	'crowded'
√mll	mumill	məll	'boring'

This data illustrates that any external affixal material of the base adjectives fails to be carried over in the corresponding comparatives. Instead, the only morphological material that appears in these comparatives is that of the underlying roots. To further expand the range of our data, let's also consider the derivation of comparatives from final weak bases:

(30)

Root	*Adj.*	*Comp.*	
√ʕli	ʕali	ʕla	'high'
√ɣli	ɣali	ɣla	'pricy'
√mḍi	maḍi	mḍa	'sharp'
√ħfi	ħafi	ħfa	'blent'

These items further prove that the derivation of comparatives is root-based. Here, the comparatives under scrutiny occur without the internal vowel characterizing their corresponding basic adjectives. However, notice that the final vowels are preserved in the comparative forms. This means that they are taken to be part of the root material. The alternation from a high vowel to a low one can be explained using our claim that high vowels represent the default nature of vowels that are part of the root and that there is a feature [+low] that marks the perfective aspect, turning high vowels into [a]. This is exactly why the comparatives above, which happen to be imperfective forms, occur with a final [a].

7. Conclusion

In this chapter, we have argued that the derivation of causative verbs and comparative adjectives is better accounted for if roots are taken as base forms. A theory of roots holds that only idiosyncratic information should be listed in the lexicon, while alternations in derivatives are taken care of by the computational system. Under this conception, causatives and comparatives are derived directly from their corresponding roots in the lexicon, instead of other surface forms. We have shown that a word-based approach poses issues related to (i) the inconsistent choice of the aspectual form of verb base forms, (ii) the random alternation of vowels, and (iii) the lack of base forms for certain causatives. In response to these, we have demonstrated that a root-based approach allows for a simpler analysis of causatives and comparatives that is limited to affixing a mora to the root to derive MCs and simply maintaining the root material in the case of the comparative.

CHAPTER EIGHT

A MORAIC APPROACH TO MORPHOLOGICAL GEMINATION IN MOROCCAN ARABIC

1. Introduction

Morphological gemination (MG) "consists of the systematic gemination of a segment associated with the systematic change in meaning of the affected base" (Samek-Lodovici, 1993). In Moroccan Arabic (MA), MG characterizes the derivation of causative verbs, agent nouns and instrument nouns. It involves the lengthening of the second segment of some base root to express the intended morphological function (e.g. √ktb 'to write' >> kə**tt**əb 'to make write'). In the case of the agent and the instrument nouns, lengthening the second segment is espoused with the presence of some vocalic material, namely the vowel /a/ (e.g. fə**lla**ħ 'farmer' and sə**mma**ʕ**a** 'headset'). This chapter will try to answer the following questions: (i) What is the morphological process responsible for morphological gemination in MA? (ii) What is the morphological exponence of the causative, agent and instrument morphemes? (iii) How does the templatic shape of each form come to be?

Often, the process of lengthening is described as a case of stem modification, whereby an operation applies to a base form and changes it without adding any segmental material. Morphologists see stem modification as difficult to accommodate in a concatenation-based model (Anderson, 1992). This chapter, however, seeks to argue in favor of a morpheme-based approach to deriving MG in MA. The ultimate goal is to extend the range of concatenative patterns to include instances of stem modification (lengthening as a case in point), and hence emphasize the claim that concatenation is an inherent property of the morphological system despite the apparent exceptions (Trommer and Zimmermann, 2014; Zimmermann, 2017).

To this end, we are proposing a constraint-based analysis of morphological gemination in MA cast within the framework of Optimality Theory (OT) (Prince and Smolensky, 1993/2004; McCarthy and Prince, 1993a, 1995, 1999). We show that OT offers an appropriate analytical device which successfully eschews problems that can be encountered elsewhere. In particular, this chapter addresses the substantial relevance of three OT offshoots, notably the theories of Generalized Alignment (McCarthy and Prince, 1993b), Correspondence Theory (McCarthy and Prince, 1995, 1999) and Generalized Template Theory (McCarthy and Prince, 1993a). Precisely, it will be demonstrated that an alignment constraint affixes a consonantal mora to the right edge of the root. The realization of the morphological mora is ensured by the interaction between the faithfulness constraints Max-Affix and IDENT-Root-Weight. The mora is then infixed due to some phonological requirement on output well-formedness, namely syllabic well-formedness. The final result is the occurrence of an infixal geminate in the relevant forms, namely causative verbs, agent nouns and instrument nouns (for more on infixation and infixal patterns, see Ultan (1975); Moravcsik (1977); Urbanczyk (1993) and Yu (2007)).

The organization of this chapter is mapped out as follows. Section 2 describes and compares the data of MCs, ANs and INs in MA. Section 3 outlines our proposal to answer how morphological gemination is derived in MA through the interaction of constraints à la Optimality Theory. In section 3.1. the theoretical background of our analysis will be delineated. Section 3.2. specifies and identifies the constraints involved in deriving the patterns in question. In Section 3.3., we show the constraint interactions responsible for yielding the derived forms. The subsequent section (i.e. Section 4) summarizes previous accounts of morphological gemination in MA. Section 5 concludes.

2. Data description and basic assumptions

2.1. Morphological causatives

In MA, morphological causatives are characterized by being morphologically marked, hence the name.[29] They are built on some base form and interpreted as verbs with an extended meaning. The first meaning is that of the base and the second meaning is contributed by the causative affix attached to it. Here, the causative affix is realized by means of lengthening the second segment of the base. The examples in (1) are illustrative of morphological causatives in MA:

(1) Causatives in MA

	Base form		Causative Form
a.	ktb	'to write'	kəttəb
	ɦṛb	'to run away'	ɦəṛṛəb
	ʃṛb	'to drink'	ʃəṛṛəb
	xrʒ	'to go out'	xərrəʒ
	dxl	'to get in'	dəxxəl

[29] Two other types of causative constructions have been discerned in the literature, namely lexical causatives and analytical causatives (Comrie, 1981). Lexical causatives are not morphologically marked, and hence inherently express the meaning of causativity. Analytical causatives, on the other hand, express causativity periphrastically.

b.	kmi	'to smoke'	kəmmi
	xwi	'to empty'	xəwwi
	mʃi	'to leave'	məʃʃi
	bki	'cry'	bəkki
c.	fiq	'to wake up'	fəjjəq
	ṭiħ	'fall down'	ṭəjjəħ
	gul	'to say'	gəwwəl
	ʕum	'swim'	ʕəwwəm

For example, the causative verb [kəttəb] consists of two different morphemes. First, the base morpheme [ktb] has the form of a trisegmental root and indicates the meaning of writing. Second, the causative morpheme contributes the meaning of causativity and is encoded in the form of a geminate [-tt-]. The geminate of every causative must match the featural specifications of the second segment of its corresponding base form.

In MA, each and every morphological causative verb falls into one of the following patterns: CəCCəC or CəCCV -where C and V indicate a consonant and a full vowel, respectively, as shown in (2):

(2) Causative patterns:

a. $C\partial C_iC_i\partial C$		b. $C\partial C_iC_iV$	
bəjjəḍ	'to whiten'	ləwwi	'to twist'
məṛṛəḍ	'to sicken'	qəṛṛi	'to teach'
zərrəb	'to speed up'	dəffi	'to warm'
fəjjəq	'to waken up'	nəqqi	'to clean'
məlləs	'to smooth'	wərri	'to show'

The difference between the items in (2a) and those in (2b) stems from the fact that they are derived from distinct base shapes. Pattern (b) includes those causative verbs that are based on final weak tri-literal verbs, where a full vowel appears at the end of the verb. Pattern (a), however, incorporates causative verbs that are derived from the other types of base forms. Hence, the pattern CəCCəC is more productive and subsumes the majority of causative verbs in MA. The existence of two morphological templates for causative verbs poses a challenge for traditional templatic theory, whereby templates have a morphological status. Later in this chapter, we will show that templates are not base-generated but follow from independently motivated demands on output well-formedness.

2.2. Agent nouns

An agent noun is usually described as a deverbal noun (i.e. derived from a verb) that refers to the doer of the action expressed by the base verb. In English, for example, agent nouns are derived through the suffixation of -er to verbs (e.g. driver, fighter, etc.). Deverbal nouns tend to lose all the verbal characteristics of their corresponding verbs. Thus, they should be distinguished from verbal nouns (e.g. gerunds in English), which maintain some of the verbal characteristics of their corresponding verbs. The grammatical process of deriving nouns from verbs is called nominalization. Based on a sample of 42 languages, Bauer (2002) determined that agent nouns are one of the most-frequent derivational nominal categories, second only to action nouns.

Just like MCs, agent nouns in MA are marked off by medial gemination; but, in addition, they incorporate a pre-final vowel [a]. Agent nouns in MA can be derived from tri-consonantal verbs, middle weak verbs as well as final weak verbs. Examples from every class are provided below:

(3) Agent nouns in MA

a.	fləħ	‘to farm’	fəllaħ	‘farmer’
	ʃfəṛ	‘to steal’	ʃəffaṛ	‘thief’
	nʒəṛ	‘to sharpen’	nəʒʒaṛ	‘carpenter’
	gzər	‘to butcher’	gəzzar	‘butcher’
	ṣbəɣ	‘to paint’	ṣəbbaɣ	‘painter’
	sṛəṭ	‘swallow’	səṛṛaṭ	‘a glutton’
	nʕəs	‘to sleep’	nəʕʕas	‘someone who sleeps a lot’
	mrəḍ	‘get sick’	mərraḍ	‘someone who gets sick a lot’
	gəbṣ	‘plaster’	gəbbaṣ	‘plaster craftsman’
	qtəl	‘to kill’	qəttal	‘serial killer’
	nqəʃ	‘to engrave’	nəqqaʃ	‘engraving artist’
b.	ħḍi	‘to watch over’	ħəḍḍaj	‘nosey person’
	ʃri	‘to buy’	ʃərraj	‘a buyer’
	kri	‘to rent’	kərraj	‘a renter’
	bni	‘to build’	bənnaj	‘mason’
	bki	‘to cry’	bəkkaj	‘someone who cries a lot’
	ʒri	‘to run’	ʒərraj	‘someone who runs fast’
	kmi	‘to smoke’	kəmmaj	‘heavy smoker’
c.	ʃuf	‘to foretell’	ʃəwwaf	‘fortune-teller’
	biʕ	‘to sell’	bəjjaʕ	‘a seller’
	bul	‘to urinate’	bəwwal	‘someone who urinates a lot’
	gul	‘to say’	gəwwal	‘someone who gossips a lot’
	fiq	‘to wake up’	fəjjaq	‘someone who gets up a lot’
	xuf	‘to get scared’	xəwwaf	‘someone who has phobia’

All the agent nouns in MA share the pattern $C\partial C_iC_iaC$. The vowels of tri-literal weak verbs turn into glides (see (3b) and (3c)). The ANs derived from final weak verbs consistently end with the glide [j], which corresponds to the unround high vowel [i]. Those ANs derived from medial weak verbs also occur with a medial glide [ww] or [jj], depending on the corresponding high vowel of the base.

2.3. Instrument nouns

Analogously, instrument nouns are also derived from verbs to indicate an inanimate doer of the action expressed by the corresponding verb. For instance, English instrument nouns are formed from verbs via the affixation of the suffix '-er' (e.g. cooker, dryer etc.), which means that the agent and instrument suffixes in English are homonymous, or otherwise the suffix -er is polysemous.[30] Instrument nouns rank number six in Bauer (2002)'s list of the most frequent derivational nominal categories.

Again, as is the case with MCs and ANs, INs in MA are also characterized by internal gemination. On a different note, INs can only come in the feminine form. They share the pattern $C\partial C_iC_iaCa$, which contains an internal vowel [a] and a final one. Some examples are shown below:

[30] See Luschützky & Rainer (2011) for more on the relation between the agent and instrument morpheme in Indo-European languages.

(4) Instrument nouns in MA

a.	ɣsəl	‘to wash’	ɣəssala	‘washer’
	nʃəf	‘to dry’	nəʃʃafa	‘dryer’
	sməʕ	‘to listen’	səmmaʕa	‘phone handset’
	skət	‘to be quite’	səkkata	‘pacifier’
	rḍəʕ	‘to suck milk’	rəḍḍaʕa	‘nurser’
	təlʒ	‘ice’	təllaʒa	‘fridge’
	frək	‘to scrub’	fərraka	‘washing board’
	ʕləf	‘to feed on’	ʕəllafa	‘horse feeder’
b.	kwi	‘to weld’	kəwwaja	‘welding machine’
	ʃwi	‘to grill’	ʃəwwaja	‘grill’
	ṭfi	‘to put out’	ṭəffaja	‘ashtray’
	mʃi	‘to walk’	məʃʃaja	‘baby walker’
	sqi	‘to water’	səqqaja	‘water supplier’

Just like in the case of ANs, the INs derived from final weak verbs end up containing a glide that corresponds to the final vowel of the base. The only notable difference is that the glide in INs appears intervocalically.

2.4. Basic assumptions

2.4.1. Extending the root-based approach

There are several ANs and INs that would have to be derived from morphologically complex base verbs if a word-based approach to derivation is maintained. This would create a conflict between the morphological make-up of the base and that of the agent and instrument forms. In particular, given the fact that the designated base verbs of this class of agent and instrument nouns are already geminated, lengthening the second segment of the base becomes redundant and unnecessary. For illustration, consider the following examples:

(5) Agent nouns from morphologically complex bases

kəssəl	'to stretch'	kəssal	'masseur'
xərrəz	'to cobble'	xərraz	'cobbler'
bərrəħ	'to publicize'	bərraħ	'publicizer'
dəlləl	'to auction'	dəllal	'auctioneer'
fəṛṛəʃ	'to furnish'	fəṛṛaʃ	'street vendor'
xəmməs	'to slave away'	xəmmas	'overtime worker'

(6) Instrument nouns from morphologically complex bases

ṣəlli	'to pray'	ṣəllaja	'praying rug'
ṣəffi	'to strain'	ṣəffaja	'strainer'
ɣəṭṭi	'to cover'	ɣəṭṭaja	'cap'
ʃəṭṭəb	'to sweep'	ʃəṭṭaba	'broom'
kərrəṭ	'to wipe'	kərraṭa	'wiper'
ḍəwwi	'to light'	ḍəwwaja	'glass roof'
ʕəlləq	'to hang'	ʕəllaqa	'clothes hanger'

In the previous chapter, we have motivated a root-based approach to deriving MCs and comparatives, given the correspondence problems it resolves. Generalizing a root-based approach to deriving ANs and INs would mean that every AN or IN should be derived from a minimal root underspecified for its grammatical category. This way, all ANs and INs would be equally derived from bases with the same amount of morphological complexity (i.e. roots). To prove that the geminates in the base forms above are not basic, we can simply refer to some other semantically related word-forms with singletons instead of geminates, like in the case of *bə**rr**əħ* 'to announce' vs. *tbriħa* 'announcement' and *ɣəṭṭi* 'to cover' vs. *ɣṭa* 'cover'. This could well be understood as another argument in favor of a root-based approach to MA morphology. Consequently, the ANs and INs shown in (5) and (6), respectively, would have to be derived from the following corresponding roots:

(7) A root-based derivation of agent nouns

√ksl	'stretch'	kəssal	'masseur'
√xrz	'cobble'	xərraz	'cobbler'
√brħ	'announce'	bərraħ	'announcer'
√dll	'auction'	dəllal	'auctioneer'
√fṛʃ	'furnish'	fəṛṛaʃ	'street vendor'
√xms	'slave away'	xəmmas	'overtime worker'

(8) A root-based derivation of instrument nouns

√ṣli	'to pray'	ṣəllaja	'praying rug'
√ṣfi	'to strain'	ṣəffaja	'strainer'
√ɣṭi	'to cover'	ɣəṭṭaja	'cap'
√ʃṭb	'to sweep'	ʃəṭṭaba	'broom'
√kṛṭ	'to wipe'	kəṛṛaṭa	'wiper'
√ḍu	'to light'	ḍəwwaja	'glass roof'
√ʕlq	'to hang'	ʕəllaqa	'clothes hanger'

2.4.2. The templatic morphology of the causative, the agent and the instrument

By comparing and contrasting MCs, ANs and INs, one easily notices that they all undergo morphological gemination. In every case, consonantal lengthening affects the second segment of the base, be it a consonant or a vowel. When the second segment is a vowel, it changes to a corresponding geminate glide. In the causative case, gemination by itself is the sole morphological marker of the derivation in question. In the agent and instrument cases, gemination is supplemented by other morphological markers, namely an internal vowel [a] that is reminiscent of the [a] in non-derived nominals. Some examples are shown below:

(9) Non-derived nominals in MA

dəft**a**ṛ	'notebook'
sərw**a**l	'trousers'
qal**a**m	'pencil'
ṣəbb**a**ṭ	'shoe'
sṭ**a**ħ	'roof'
ʕəwd**a**n	'horse'

In addition, INs are also characterized by a final vowel [a] which is indicative of their feminine form. The final vowel [a] is a typical feminine marker in many non-derived nouns as well. Consider the examples below:

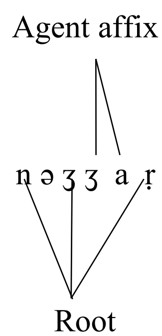
(10) Non-derived feminine nouns:

ħadiq**a**	'zoo'
luʕb**a**	'game'
wəṛq**a**	'paper'

ʒuml**a**	‘sentence’
luħ**a**	‘board’
misṭar**a**	‘ruler’

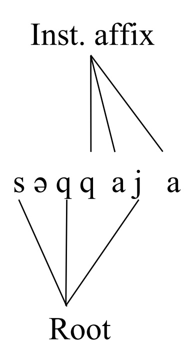
In this chapter, it will be assumed that the agent morpheme in MA is represented by the gemination of the second segment as well as the infixation of an internal [a]. This represents a case of a hybrid morpheme, whereby one part is vocalic and the other is prosodic (i.e. a mora).

(11) The morphological content of agent nouns in MA



Likewise, the instrument morpheme is assumed to be expressed by the gemination of the second segment of the base in addition to the affixation of an internal [a] and a final one.

(12) The morphological content of instrument nouns in MA



Such an assumption resonates well with the premises of Generalized Template Theory (McCarthy and Prince, 1993a), which will work as our background approach to describing the templatic morphology of MA. At the center of this approach is the idea that templates (e.g. the agent and instrument templates) are emergent structures that have no basic morphological status. This means that the internal [a], that characterizes the derivation of both ANs and INs, and the final [a], which is specific to INs, should have a basic morphological status, instead of being prosodically motivated by some template.

Despite the deceptive templatic morphological similarities between the causative, the agent and instrument templates, we will be assuming that there is no direct morphological connection between them, except that they can be derived from the same root type. That is, ANs are not nominalized causatives, and instrument nouns are not feminized ANs. Rather, each is derived directly from a root base by referring to an independent morpheme. This also means that the gemination in these forms performs a different morphological function in every case:

(i) It represents the causative morpheme in MCs

(ii) It forms the agent morpheme together with the internal [a] in ANs

(iii) It realizes the instrument morpheme jointly with the internal [a] and the final [a] in INs.

On a different matter, every agent noun can be inflected for the feminine form, by suffixing the feminine marker vowel [a], without turning into an instrument noun. Feminine ANs have the same exact phonological shape of INs, except that the latter are feminine by default.

(13)

Masc. ANs		*Fem. ANs*
fəllaħ	‘farmer’	fəllaħa
ʃəffaṛ	‘thief’	ʃəffaṛa
nəʒʒaṛ	‘carpenter’	nəʒʒaṛa
gəzzar	‘butcher’	gəzzara
ʃəwwaf	‘fortune-teller’	ʃəwwafa

This fact proves two points. First, it shows that INs are not derived from ANs since the latter already have their independent feminine forms. Second, it demonstrates that the feminine quality of INs is not optional, meaning that they can only be feminine. That is, in the event of removing the final [a] of the instrument, the meaning will be incomplete (e.g. *təllaʒ, *səmmaʕ, *səkkat etc.). This seems as an interesting case of overlap between inflection and derivation, in that the [a] in question serves as an exponent of both. Put simply, despite the apparent inflectional feminine quality of the final [a] of INs, it actually does more than being an inflectional marker (John McCarthy, personal communication). In fact, it is also derivationally relevant.

2.4.3. The moraic two-root node representation in the context of MG

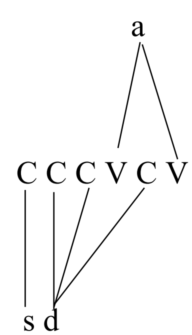
Now, let’s wind up this section by commenting on the class of INs whose base roots include an underlying final geminate. We will proceed to call this the ‘overgemination’ problem as it seems to involve the gemination of what is already an underlying geminate. Some illustrative examples are shown below:

(14)

√sdd	‘shut’	səddada	‘cover’
√ħkk	‘scratch’	ħəkkaka	‘mince’
√ṛʃʃ	‘spray’	ṛəʃʃaʃa	‘sprayer’

The question here is: how does this structure get derived? One answer could be that a geminate is simply a cluster of two identical consonants. Therefore, this can be seen as a case of morphological doubling of one member of the cluster. Such an answer echoes the SPE conception of geminates as sequences of identical segments. Another possibility might be to explain this in terms of melodic spreading to a skeletal tier á la autosegmental phonology. This would mean that a single melodic element would have to be multiply associated to three slots on the skeletal tier as a result of mapping a bi-consonantal root to the instrument template (i.e. CCCVCV).

(15) The autosegmental account



[səddada] ‘cover’

Despite the fact that both the non-linear and the autosegmental approaches to consonantal length appear to provide a simplified answer to this case of ‘overgemination’, we believe that the shortcomings and inadequacies of these approaches elsewhere have been well noted (see chapter #3 for a detailed discussion of these shortcomings).

However, considering a moraic approach to gemination, such as the one we entertain in this dissertation, it seems more challenging to account for this data. In particular, this is said because the geminates of the base roots in question behave like two-part segments, whereby only one part is reproduced. Thus, even if we could assume that an underlying geminate can be further geminated for morphological purposes by means of moraic affixation, it would be hard to explain why only 'half' of the geminate is replicated and not all of it (e.g. *ṛəʃʃaʃʃa). This would be an embarrassment especially to a moraic one-root node theory.

In the face of this situation, the two-root node aspect of our model comes into good service, in the sense that it allows morphological gemination to target only one root node. What happens is that when the mora of the instrument morpheme attaches to the first root node of the underlying geminate, the latter's structure collapses into a sequence of two independent segments. This is akin to the structure created via geminate breaking by means of morphological processes. Now that the first root node is an independent segment, it can geminate to encode the moraic part of the instrument morpheme, creating a geminate structure similar to that of underlying geminates. The second root node of the underlying geminate maintains its independence, behaving like any normal short consonant.

3. Analysis

3.1. The Unified Theory of Morphological Gemination

Our proposal to account for morphological gemination by dint of mora affixation finds further support in previous analyses, namely Lombardi and McCarthy (1991), Samek-Lodovici (1992) and Bensoukas (2001).[31]

Lombardi and McCarthy (1991) argues that the formation of completive verbs in Choctaw, a Muskogean language, draws on, among other things, the prefixation of a mora to a circumscribed base à la prosodic circumscription. The affixed mora copies the featural content of the adjacent consonant by a rule of spreading, hence geminating the affected consonant. Next, the derived geminate occurs word-medially after concatenating the extraprosodic constituent. A portion of the data dealt with in Lombardi and McCarthy (1991) is reproduced below for easy reference:

(16) Morphological gemination in Choctaw

Base	*completive verb*	
talakci	ta**ll**akci	'to be tied'
falama	fa**ll**aama	'to return'
takci	ta**yy**akci	'to tie'
pisa	pi**yy**iisa	'to see'
oktabli	okta**yy**abli	'to dam up'
toksali	toksa**yy**aali	'to work'

[31] For more recent moraic accounts of morphological gemination, see Trommer and Zimmermann (2014) and Zimmermann (2017).

To take an example, the completive verb *tallakci* is derived from the base *talakci,* whose first mora (i.e. *ta*) is rendered extraprosodic by a rule of prosodic circumscription. A mora is then prefixed to that base, causing the lengthening of the adjacent segment. Later, when the circumscribed mora is reinstated, the gemination appears internally.

Samek-Lodovici (1993) capitalizes on Lombardi and McCarthy (1991) and extends the analysis to provide a unified account of cross-linguistic morphological gemination, using the constraint-based framework of OT. The essence of the analysis established in this work is that MG in Keley-i and Alabama arises from edge-oriented mora affixation. The optimal output is then decided on by the syllable well-formedness constraints of each language. The data set of MG in Keley-i and Alabama looks fascinating. For convenience, some portions of the data are repeated below:

(17) Morphological gemination in Keley-i

Base	pi.li 'to choose'	du.yag 'to pour'	ʔag.tu 'to carry on'	dun.tuk 'to punch'
Subject focus				
Input	um.pi.li	um.du.yag	man.ʔag.tu	um.dun.tuk
Output	um.pi**l.l**i	um.du**y.y**ag	man.ʔag. tu	um.dun.tuk
Object focus				
Input	pi.li	du.yag	ʔag.tu	dun.tuk
Output	pi**l.l**i	du**y.y**ag	ʔag.tu	dun.tuk
Access. focus				
Input	ʔi.pi.li	ʔi.du.yag	ʔi.ʔag. tu	ʔi.dun.tuk
Output	ʔi**p.p**i.li	ʔi**d.d**u.yag	ʔi**ʔ.ʔ**ag. tu	ʔi**d.d**un.tuk

(18) Morphological gemination in Alabama

a. Words with open antepenultimates always geminate the following onset

a.ta.kaa-li	a.tá**k**.**k**aa.li	'hang one object'
a.caa.-pa	á**c**.**c**aa.pa	'object to vocally'
a.fi.nap-li	a.fí**n**.**n**ap.li	'lock up'
a.tak.-li	á**t**.**t**ak.li	'hang more than one object'
ho-co.ba	hó**c**.**c**o.ba	'big (pl)'
a.taa.nap.-li	a.tá**n**.**n**ap.li	'rancid'

b. Bisyllables and light penultimate geminate vocalic nucleus

co.ba	c**óo**.ba	'big (sing)'
i.s-i	**íi**.si	'take, catch'
cam.po.-li	cam.p**óo**.li	'taste good'
i.bak.pi.la	i.bak.p**íi**.la	'turn upside down'
ho-f.na	h**óo**f.na	'smell'
is.-ko	**íi**s.ko	'drink'

In both Keley-i and Alabama, MG marks the imperfective aspect of verbs. The data above illustrates how the process of MG works in tandem with the syllabification process to produce structures that best comply with the well-formedness constraints of each language. In the case of Keley-i, MG cannot get realized if it would create superheavy syllables. (e.g. see 'ʔag.tu' and 'dun.tuk' above). The keley-i data also shows that tautosyllabic geminates are not allowed, which explains why vowels are never geminated. As for Alabama, final syllables are skipped by MG due to their extrametrical nature. In Alabama, however, tautosyllabic gemination is possible (e.g. c**óo**.ba)

Along the same lines, Bensoukas (2001) also accounts for Tashlhit imperfective morphological gemination in terms of the affixation of a mora to a base root. Specifically, despite the complex nature of the imperfective derivation in Tashlhit, the moraic conception of morphological gemination proved to be uselful in developing a unified analysis for the different classes of the imperfective. For easy illustration, only the portion of the data where MG straightforwardly affects one of the root segments is cited below:[32]

(19) Morphological gemination in Tashlhit

a. Geminating the first radical element

krz	**kk**rz	'plow'
frd	**ff**rd	'graze'
ħrt	**ħħ**rt	'go ashore'
krf	**kk**rf	'tie'

b. Geminating the second radical element

kla	k**ll**a	'spend the day'
ftu	f**tt**u	'go'
mgr	m**gg**r	'harvest'
rgl	r**gg**l	'lock'

In analyzing the imperfective gemination in Tashlhit, Bensoukas (2001) posits an abstract prosodic affix that consists of a consonantal mora. The latter is freely attached to the base root to derive the imperfective form. The landing position of the affix, which decides what segment to geminate, is then determined through the interaction of well-formedness constraints on syllable and prosodic structure.

[32] There is a third class of imperfectives where the mora is realized outside of the root material in the form of a prefixed geminated [tt] e.g. ttkka 'pass', ttumu 'contain' and ttfrtal 'escape'.

According to Samek-Lodovici (1992), any account of MG should be able to address the following two questions: (a) how does morphological gemination occur? And (b) where does gemination happen (i.e. which segment geminates)? To answer these questions, it was suggested that MG should be broken down into two main modules:

(20)

a. *Affixation module*: freely add the moraic morpheme to the base by altering the base's underlying prosodic configuration

b. *Selection module*: select the derivation whose phonological structure optimally satisfies the constraints of the language

The labor of MG is divided between these two components in the following way. The first module represents the part of MG that is common between the various languages where MG is attested. It simply consists of affixing a moraic morpheme that performs the intended morphological function. If successfully realized, the designated affix changes the prosodic as well as the segmental configuration of the base. The second module, however, relies on the framework of OT to account for the cross-linguistic variations concerning the landing position of the moraic affix. This module is composed of universal constraints on syllabic well-formedness, which are independently motivated by the phonology of the relevant languages. Since the ranking of these constraints varies from one language to another, distinct patterns of morphological gemination are produced.

3.2. Generalized Template Theory: A Root-and-Prosody Approach

As has been pointed out earlier, MCs, ANs and INs are all characterized by an invariant templatic shape. The template of MCs consists of two light syllables (i.e. LL), which could have the form [CəC.CəC] or [CəC.CV], depending on the nature of the base root. ANs also come in the form of a disyllabic template, except that one syllable is light and the other is heavy (i.e. LH): [CəC.CaC]. Finally, the template of INs is composed of three light syllables (i.e. LLL): [CəC.Ca.Ca].

To account for the templatic morphology of the forms under study, we adhere to the premises of Generalized Template Theory (GTT; McCarthy and Prince, 1994)), whereby templates are believed to derive from the interaction of independently motivated constraints on the well-formedness of output prosody. GTT represents a departure from earlier templatic theories, namely CV-theory (McCarthy, 1979, 1981) and Prosodic theory (McCarthy and Prince, 1993a), in which templates are assumed to be listed in the lexicon as morphological entities. Within OT, the existence of constraints which explicitly dictate the prosodic shape of templates was found to result in predicting grammars that are, in fact, nonexistent.[33]

In the framework of GTT, the Fixed Prosody approach (FP; Ussishkin, 1999, 2000, 2005) has been proposed to handle the templatic effects in Semitic without the need for templatic specific constraints. In this approach, roots have no specific morphological status. Instead, word formation is believed to be word-based, in that new word forms are derived from other output forms through melodic overwriting. According to the FP approach, templates are emergent

[33] This is dubbed the Kager-Hamilton problem. See McCarthy and Prince (1999) and Ussishkin (2000) for more on arguments against templatic constraints.

structures that follow from general constraints of prosody, namely constraints of prosodic minimality and maximality.

In this chapter, a different approach towards templatic effects will be adopted. The approach is referred to as the Root-and-Prosody approach (RP; Kramer, 2007; Tucker, 2010, 2011). Like the FP approach, the RP approach assumes that templates arise from the interaction between independently needed constraints. However, the RP approach differs from the FP one in terms of being root-based. This makes the RP approach more aligned with our root-based approach to word formation in MA. The main underpinnings of this approach are summarized as follows:

(21) Central claims of the Root-and-Prosody approach (Tucker, 2010)

a. *Roots and vowels are morphemes*: the input to nonconcatenative templatic morphology (NTM) forms consists of the consonantal root and a vowel affix

b. *Templates are given by prosody*: Templates are emergent properties of words in NTM languages which surface from the necessary satisfaction of high-ranking prosodic markedness constraints

According to the RP approach the nonconcatenative templatic morphology of Semitic stems from the low-ranking status of contiguity faithfulness, hence the constraint CONTIGUITY, in relation to prosodic well-formedness: Prosodic Markedness >> CONTIGUITY. This means that the discontinuous linearization of the root and the affixal material is obtained simply through constraint interaction. The RP approach requires only the following types of constraints to derive all template shapes and sizes:

(22) Constraints in an RP approach

a. *Prosodic/Syllabic Constraints*: Constraints on prosody/syllable structure independently needed in the language (FtBin, *Complex, Onset, etc.).

b. *Morphological Constraints*: Constraints which align morphemes in linear prosodic structure (Align-L (n, ω), Align-L (-t-, ω), etc.).

c. *Faithfulness Constraints*: Faithfulness constraints of the usual family (Ident[F], Dep-Root, Max, Contiguity, Linearity etc.).

So far, the RP approach has been used to account for the templatic morphology of verbal forms in Coptic (Kramer, 2007) and Iraqi Arabic (Tucker, 2010). In this chapter, we look toward extending the premises of the RP approach to other lexical categories and languages, namely agent and instrument nouns in MA.

.

3.3. Identifying the constraints and their interactions

3.3.1. Affixation module

To account for the various patterns of MG in MA, we follow Lombardi and McCarthy (1991), Samek-Lodovici (1992) and Bensoukas (2001) in theorizing that a prosodic mora can be used by morphology to play the role of an affix. Therefore, we believe that MCs, ANs, and INs in MA are derived, fully or partially, by the affixation of a consonantal mora which is not associated with any melodic material underlyingly. Besides its morphological role as an affix, the posited mora changes the phonological make-up of the base form by lengthening the medial segment of the root. Accordingly, the postulated mora in our analysis has the following properties:

a. Morphologically, it is an affix material that plays the role of a morphological marker.
b. Phonologically, it is a prosodic constituent that changes the prosodic weight and segmental length of the root.

With this in hand, we proceed to show that the affixation module in our analysis calls for the service of two constraints. First, the faithfulness constraint MAX-Affix is responsible for the realization of the moraic affix in the output form. Particularly, this constraint is posited to guarantee that the mora is both prosodically parsed and segmentally filled. To obtain this result, the MAX-Affix constraint has to dominate the faithfulness constraint IDENT-Root-Weight, which resists any change of the weight configuration of the input root. In this sense, IDENT-Root-W should be violated so that the moraic affix can be realized in the output. In the case of ANs and INs, the scope of MAX-Affix is extended to ensure the realization of the vowels characterizing their derivation. In other words, the deletion of the internal [a] of the AN and the IN or even the final [a] of the IN would too incur a violation of MAX-Affix.

Second, the analysis also involves an alignment constraint that specifies the edge targeted by the moraic affix. Recall that affixes in OT are either left-aligned or right-aligned. The infixation of a morpheme is normally the result of phonological pressure. That is, before a morpheme gets infixed, it starts at some edge. Defining the canonical edge of the moraic affix in our analysis is no easy task since all forms realize the mora word medially. There are no forms where the mora is realized on its canonical edge. Also, since the infixed mora is equally close to both edges, there is no way to tell where the mora has originated based on its closeness to one of the edges. As a consequence, we would just assume that the affix is right-aligned since suffixation represents the least marked form of affixation. In the case of MCs, where the mora fully represents the causative morpheme, the alignment constraint is defined as follows:

(23) **ALIGN-R (μ)**

The moraic causative affix should be right-aligned

This constraint requires that the right edge of the moraic affix coincide with the right edge of the base root, resulting in the suffixation of the causative morpheme. Regardless of where the mora is realized, the preliminary outcome of this process is the derivation of causative verbs. For instance, given a root form like /tlʒ/ 'snow', ALIGN-R (μ) places a mora at its right edge for the sake of deriving the causative verb [təlləʒ] 'to make snow', as depicted in (24), where brackets refer to the edges of each constituent.

(24)

Root	Affix	Alignment (Root +Affix)
[tlʒ]	[μ]	[[tlʒ]μ]

Otherwise, if the purported affix is aligned to an edge other than the right edge, say the left one, this constraint would be violated, as displayed in the incomplete tableau below:

(25)

Input: √tlʒ	ALIGN-R (μ, Root)
a. [μ_c [tlʒ]]	Violated
b. [tlʒ[μ_c]]	Satisfied

One last fact is that infixation itself is an edge-oriented phenomenon, whereby the affix occurs as close as possible to the designated edge. The more misaligned the affix, the more punished by the alignment constraint in a gradient fashion, as is shown in the following incomplete tableau:

(26)

Input: √nʕs	ALIGN-R (μ, Root)
a. **nn**ʕəs	**!
b. nə**ʕʕ**əs	*
c. nʕə**ss**	Satisfied

As for ANs and INs, the alignment constraint involved in their derivation should be specified for the whole affixal material. This means that when considering the agent and instrument morphemes, which are partly moraic and partly vocalic, their versions of the alignment constraint will have the following form and specifications:

(27) **ALIGN-R (μ-a)**

The agent affix must be right-aligned.

(28) **ALIGN-R (μ-a-a)**

The instrument affix must be right-aligned.

Next, we will see how the ultimate position of the affix material of each morpheme is selected. That is, we will be looking at the factors that lead the designated affixes to miss the right alignment and occur inside the base.

3.3.2. Selection module

Since MCs, ANs and INs all consistently come with medial geminates, it is obvious that the moraic affix is invariably realized on the second segment. This means that the moraic affix gets misaligned (i.e. infixed). The question now is what causes the moraic affix to be realized in a position other than the right edge?

As has been previously mentioned, the canonical position of the causative morpheme is determined by the constraint ALIGN-R (μ), demanding that the right edge of the affix coincide with the right edge of the root. However, it is argued that the infixation of the mora is enforced by some phonological restrictions on output forms. We believe that the insertion site and filling segment of the mora are decided on by the need to conform to the prosodic and syllabic well-formedness of MA. In particular, it is suggested that what forces infixation in the relevant forms is the quest for the least marked possible syllabification of the input material, in accordance with the syllabic well-formedness constraints of MA. This is captured by the constraint σWF (i.e. syllabic well-formedness). As a result, being high-ranked, σWF bans the moraic affix from being realized on the edge as it is going to lead to a marked structure. This can be expressed as follows:

(29) Syllable well-formedness forces violation of alignment:

σWF >> ALIGN-R (μ, Rt)

The σWF (i.e. syllable well-formedness) constraint is an umbrella constraint that subsumes a variety of specific constraints on syllable structure. These include:

a. **ONSET**: syllables must have onsets
b. ***μ/C_h**: a consonant should not be the head of a syllable
c. ***EMPTY-headedσ**: a syllable must have a nucleus
d. ***μ/ə**: schwas are nonmoraic

For example, the right alignment of the moraic affix in the case of MCs would lead to the creation of a word-final geminate, which, in turn, would project an empty-headed syllable to avoid being fully contained in the coda. Thus, in order to avoid this scenario, the mora should be realized in a position other than the right edge. Here comes the role of infixation as a way out. Another example concerns the vowels of the instrument noun, where the first /a/ is infixed alongside the mora. If both vowels are kept side by side on the right edge, the ONSET constraint will be violated. Hence, under the pressure of the ONSET constraint, the first /a/ moves inside the root to look for an onset, leaving its canonical onset to the final /a/.

After having presented the relevant constraints in an incremental and detailed way, we intend to show, in the subsequent section, how these constraints together with their ranking lead to the formation of MCs, ANs and INs.

3.3.3. Constraint interaction

The central idea that underlies our OT analysis is that there are four major constraints responsible for the derivation of morphologically geminated forms in MA. First, MAX-Affix is a faithfulness constraint which stipulates that the input affixal material must be fully manifested in the output form for the sake of realizing the designated morpheme. This constraint militates against the nonrealization of the affix material in the output. Hence, for this constraint to be satisfied, gemination should take place.

The Second constraint, IDENT-Root-Weight, is a faithfulness constraint which requires corresponding output and input root segments to be identical in terms of weight. As a result, this constraint disallows input segments from gaining or losing phonological weight. Third, ALIGN-R is an alignment constraint that demands coincidence between the right edge of the affix in question (i.e. the consonantal mora) and the right edge of the root. Given this demand, ALIGN-R (μ, Rt) is violated whenever the edges of the specified constituents fail to coincide. Fourth, σWF is a constraint on prosodic well-formedness which subsumes the main constraints on syllable structure. Fifth, CONTIGUITY is a faithfulness constraint that demands the contiguity of input segments to be preserved in the output. These are summarized below:

(30) Constraints responsible for deriving morphological causatives in MA:

a. **ALIGN-R (μ, Root)**: The right edge of the moraic affix must coincide with the right edge of the root.
b. **MAX-Affix**: The input affixal material should be preserved in the output form.
c. **IDENT-Root (Weight)**: The weight specification of the root must be preserved in the output.
d. **σWF**: The output form should satisfy the following markedness constraints on syllable well-formedness:

- **ONSET**: syllables must have onsets
- ***μ/C$_h$**: a consonant should not be the head of a syllable
- ***Empty-headedσ**: a syllable must have a nucleus
- ***μ/ə**: schwas are nonmoraic

e. **CONTIGUITY**: The contiguity of input material should be preserved in the output.

We proceed by putting forward the appropriate rankings that would yield the expected output forms. In particular, we highlight the existing ranking arguments between the constraints at play. Ranking arguments will provide us with evidence for ranking certain constraints over others or leaving them unranked.

First, tableau (31) compares the winning candidate [kəttəb] 'to make write' with the losing candidate *[ktəb], whereby the former satisfies Max-affix and violates IDENT-Root-W, and the latter violates MAX-Affix and satisfies IDENT-Root-W. This shows that these constraints do really have conflicting demands. Consequently, MAX-Affix is ranked high because it favors the winning candidate while IDENT-Root-W is ranked low as it favors the loser.

(31) Ranking argument (1): MAX-Affix >> IDENT-IO (Weight)

Input: μ √ktb	MAX-Affix	IDENT-IO (Weight)
☞a. kəttəb		*
b.ktəb	W*!	L

Second, tableau (32) displays a conflict between σWF and ALIGN-R (μ) as they disagree on the assessment of both the winner and the loser. Since σWF appears to favor the winning candidate, it ranks high in the hierarchy above ALIGN-R (μ).

(32) Ranking argument (2): σWF >> ALIGN-R (μ)

Input: μ √ktb	σWF	ALIGN-R (μ)
☞a. kəttəb		*
b. k.təb.b	W*!	L

Third, tableaux (33) and (34) show us that the relevant constraints agree on the assessment of at least one candidate, which means that they do not conflict or dominate each other. As a result, these constraints will be left unranked with respect to each other.

(33) No ranking argument: MAX-Affix and ALIGN-R (μ) have no effect on each other

Input: μ √ktb	MAX-Affix	ALIGN-R (μ)
a. kəttəb		*
b. k.k.təb		

(34) No ranking argument: σWF and IDENT-Rt-W have no effect on each other

Input: μ √ktb	σWF	IDENT-IO (Weight)
a. kəttəb		*
b. k.təb.b	*	*

What these rankings tell us is that there is no one absolute and total ordering between every pair of our constraints. Rather, the constraint set consists of two independent ranked sets in the same hierarchy whereby output forms are evaluated by each in a parallel fashion. This can be vividly illustrated by the following Hasse diagram:

(35) Ranking summary for causatives in MA

MAX-Affix	σWF
|	|
IDENT-Rt-W	ALIGN-R (μ)

An important fact about OT is that no ranking should be enforced between two constraints where there is no evidence to that. Also, it is not necessary to provide a total ordering of the relevant constraints. A constraint may stay unranked in relation to some other constraint, or it may have no ordering relation at all with another constraint, as far as they agree on the assessment of the winning candidate. As noted by McCarthy (2008:43), "it's perfectly OK if the process of analysis leads to a partial ordering: CONST1 and CONST3 both dominate CONST2, but the ranking between CONST1 and CONST2 is unknown." In this case, only the crucial constraint rankings that are important for the analysis should be determined and highlighted.

In support of our constraint ranking, we demonstrate the possible repercussions that can emerge if the ranking suggested is altered. The re-ranking of IDENT-Root-Weight over MAX-Affix in (36) yields forms with no gemination, while the re-ranking of alignment over syllabic well-formedness in (37) gives rise to forms with edge geminates at the right periphery. In each case, the winning candidates are not attested in MA. Conventionally, the symbol "💣" indicates this fact.

(36) IDENT-Root (Weight) >> Max-affix

Input: µ √ktb	IDENT-IO (Weight)	Max-affix
a. kəttəb		*
💣b.ktəb	*!	

(37) ALIGN-R (µ) >> σWF

Input: µ √ktb	ALIGN-R (µ)	σWF
a. kəttəb		*
💣b. k.təb.b	*!	

3.3.4. Output selection: the causative

Bearing in mind the constraints and the ranking arguments presented above, we proceed to illustrate the overall interactions of those constraints and their role in the selection of the optimal structures. To begin with, let's look at how the causative *kəttəb* 'to make write' emerges as the winner in its competition with other candidates.

(38) MAX-Affix >> IDENT-Root (Weight); σWF >> ALIGN-R (µ, Rt)

Input: µ √ktb	MAX-Affix	σWF	IDENT-Root-W	ALIGN-R (µ)
☞ a. kəttəb			*	*
b. ktəb	*!			
c. k.kət.b		*!	*	**
d. k.təb.b		*!	*	

This tableau demonstrates how the optimal template of MCs emerges. For example, given the input /µ/ √ktb, the optimal form is [kəttəb]. This form realizes the moraic causative affix by geminating the second segment of the root for the sake of achieving the least marked syllable structure possible. The other competing candidates shown in the tableau represent cases of candidates that either fail to parse/fill the moraic affix (i.e. 38b) or geminate the peripheral consonants of the root, creating marked syllable structures (i.e. 38c and 38d), hence violating the high-ranking constraints MAX-Affix and σWF, respectively. For example, candidate (38d) is ruled out because of the empty-headed syllable projected by its final geminate. Note that candidate (38c) is harmonically bounded to the optimal form since it is suboptimal for other reasons, namely the multiple violation of ALIGN-R (μ). This clearly shows that prosody is responsible for the linearization of the input material and the construction of invariant templates.

3.3.5. Output selection: the agent noun

The idea of how prosody could affect input linearization and template derivation would be clearer as we consider the cases of the AN and IN. Particularly, in the context of ANs and INs, the low-ranking status of CONTIGUITY gets more highlighted since the root material gets interrupted by some of the affixal material. We have mentioned before that the contiguity of the input elements gets neutralized under the pressure of constraints on prosodic well-formedness. The following tableau illustrates how the interaction between prosodic well-formedness and contiguity faithfulness leads to the formation of the AN template:

(39)

√nʒṛ /μ-a/	σWF	CONTIGUITY
a. nəʒ.ʒ.ṛa	*!	
☞b. nəʒ.ʒaṛ		*

This tableau reveals that neutralizing the contiguity of the root, by means of infixing the vowel /a/ of the agent affix, is necessary to avoid a syllable structure like the one derived in candidate (39a), whereby the medial geminate projects an empty-headed syllable. Therefore, infixing the vowel /a/ is driven by the need to provide a nucleus for that syllable. The other constraint interactions involved in the derivation of ANs are shown in the tableau below:

(40) Deriving the agent from tri-consonantal roots

Input: √nʒṛ /μ-a/	MAX-Affix	σWF	IDENT-Root (Weight)	ALIGN-R (μ-a)	CONTIGUITY
☞a. nəʒ.ʒaṛ			*	*	*
b. n.ʒaṛ	*!	*		*	*
c. n.nəʒ.ṛa		*!	*	**	
c. n.nə.ʒaṛ		*!	*	**	*
d. n.n.ʒaṛ		*!	*	**	*
e. nəʒ.ʒ.ṛa		*!	*	*	
f. nəʒ.ʒəṛ.a		*!	*	*	
g. nəʒ.ṛ.ṛa		*!	*		

This tableau evaluates the most plausible candidates that could be generated from the preliminary input: √nʒṛ-/μ-a/. Candidate (40b) is ruled out by MAX-Affix for failing to realize the moraic part of the agent morpheme. Candidates (40c) through (40g) represent different permutations of input linearization. All these candidates violate some specific aspect of syllabic well-formedness. In particular, (40c) violates *μ/C_h and *μ/ə, (40d) violates both *μ/C_h and *EMPTY-headedσ, (40e) violates *EMPTY-headedσ, (40f) violates ONSET and (40g) also violates *μ/C_h. The only linearization that succeeds in avoiding the violation of the high-ranked constraints of prosodic markedness is the one that gives us the optimal form in (40a).

3.3.6. Output selection: the instrument noun

In the same vein, the templatic shape of INs also follows from the interaction between σWF and CONTIGUITY, such that σWF dominates CONTIGUITY. This way, the vocalic material of the instrument morpheme is licensed by prosodic well-formedness to disrupt the contiguity of the root material.

(41)

√tlʒ /μ-a-a/	σWF	CONTIGUITY
a. təl.l.ʒ.a.a	*!	
☞b. təl.la.ʒa		*

The candidate in (41a) above incurs numerous violations of different well-formedness constraints, preferring to keep the contiguity of its root material intact. The winning candidate in (41b), however, neutralizes the contiguity of its constituents in order to conform to the syllabic well-formedness of the language.

In the tableau below, it will be shown that the same constraint hierarchy involved in the derivation of MCs and ANs can account for the derivation of INs as well. In particular, the following tableau shows how the IN *təllaʒa* ‘refrigerator’ is derived from the input: √tlʒ /μ, a, a/.

(42) Deriving the instrument noun

Input: √tlʒ /μ, a, a/	MAX-affix	σWF	IDENT-Root (Weight)	ALIGN-R (μ-a-a)	CONTIGUITY
☞ a. təl.la.ʒa			*	**	*
b. t.ləʒ.ʒa.a		*!			
c. t.t.la.ʒa		*!	*	**	*
d. t.təl.ʒa.a		*!	*	**	
e. t.la.ʒa	*!	*		*	*
f. təl.l.ʒa	*!	*		*	
g. təl.laʒ	*!		*	*	*

This tableau demonstrates that the way the optimal form in (42a) linearizes the input material does not violate any of the syllabic well-formedness constraints. On the contrary, candidates (42b) through (42f) are ruled out for violating some of the well-formedness constraints. For example, the candidate in (42b) specifically violates ONSET and $*\mu/C_h$. Candidates (42e) and (42f) are sub-optimal for other reasons as well, namely the violation of MAX-Affix. In both candidates, the affix material is not realized in its entirety. Particularly, candidate (42e) does not realize the prosodic part of the instrument morpheme, i.e. mora, while candidate (42f) fails to realize the nominal /a/ of the affix. Candidate (42g) is excluded for similar reasons, specifically the non-realization of the feminine /a/ of the instrument morpheme.

Another way of looking at INs could be to argue that they form one class with ANs. This means that they would have to be viewed as sharing the semantic affinity expressed by the morpheme /μ-a/. The difference, however, lies in the grammatical specificity brought by the final /a/ of the instrument form. Such an approach to deriving INs suggests a case of output-based derivation. We previously pointed out to the fact that the final /a/ of the instrument form serves a derivational function as well as an inflectional one. Concerning this point, an output approach to instrument formation further asserts this fact since the /a/ occurring at the end of the instrument form becomes the major characteristic of its derivation. Therefore, feminizing ANs becomes the main way to express the instrument meaning. Nevertheless, it should be noted that such an account does not comply with our view of maximizing a root-based perspective to word derivation in MA. Besides, this account would create many forms which morphologically qualify as potential ANs, but they are, in fact, semantically unattested (e.g. *təllaʒ, *səmmaʕ, *səkkat etc.).

4. Earlier accounts of morphological gemination in MA

The analysis that was developed in this chapter built on previous works that have dealt specifically with morphological gemination in causatives. To the best of our knowledge, no accounts have been devoted to investigating the morphology of ANs and INs in MA. The treatments to be reviewed and discussed here primarily include the pre-OT circumscription analysis of Bennis (1992), the pre-OT autosegmental analysis of Bennis and Iazzi (1995) and the OT partial reduplication analysis of Boudlal (2001).

4.1. The prosodic circumscription analysis

To start with, Bennis (1992) maintains that the causative in MA is derived from word forms (i.e. nouns, verbs and adjectives) through the affixation of a heavy syllable (i.e. bimoraic) to a circumscribed prosodic constituent that can be of the shape -CəC or CV. For instance, the causative [ləʕʕəb] ‘to make play’ is derived from the verb [lʕəb] by prefixing the heavy syllable [σμμ] to the prosodic domain [-ʕəb]. The idea is that one mora generates a geminate while the other mora triggers schwa epenthesis.

(43) The prosodic circumscription analysis

Aff Caus: lʕəb = σμμ- ʕəb * l

= -əʕʕəb * l

= ləʕʕəb

There are two aspects for which this analysis can be criticized. First, it rests heavily on the notion of extraprosodicity, in that it assumes that a segment remains inert and stays out of the derivational procedure, while the morphological process of affixation targets a well-defined prosodic constituent. Extraprosodicity is needed to explain the occurring infixation phenomenon in causatives. In our analysis, however, the infixation of the causative morpheme could be attained simply via the interaction of universal constraints.

Second, the causative affix is believed to be a heavy syllable (i.e. σμμ) which consists of two moras, one to motivate schwa epenthesis and the other to trigger consonant gemination. Nonetheless, it has been shown elsewhere (Bensoukas and Boudlal, 2012a-b) that schwa in MA is nonmoraic. Instead, the nucleus schwa and the following coda consonant share one single branching mora. Grounded on this fact, our analysis limits the prosodic shape of the causative morpheme to one mora responsible for consonant lengethening. Schwa, on the other hand, is independently motivated and is more general, in that it is epenthesised to break up any impermissible consoant clusters.

4.2. The skeletal analysis

Bennis and Iazzi (1995) accounts for the derivation of causatives in MA using the representational premises of autosegmental phonology. The basic claim of this work is that the causative template has the shape XXXX, where X refers to an underspecified position in the segmental tier. This template is then associated to a lexical entry as follows. First, the second radical element is pre-associated to the third position in the template. Second, the other segments spread to associate with the remaining corresponding positions in the template. Finally, the rules of syllabification apply. Below is an example of deriving the causative [dəxxəl] 'to make enter':

(44) The autosegmental analysis

Lexical entry	*Pre-association*	*Spreading*	*Syllabification*
	X X X X	X X X X	
√dxl	d x l	d x l	[dəxxəl]

A major shortcoming of this analysis is that it resorts to an ad-hoc solution to explain the internal gemination that has come to characterize the causative verb in MA. Neither right-to-left nor left-to-right spreading could yield the right forms. It is only through the mechanism of pre-association that the correct forms can be produced. Again, our analysis can explain internal gemination by means of interaction between universally attested constraints.

4.3. The reduplication analysis

Boudlal (2001) analyzes the causative pattern by means of constraint interaction à la OT. At the heart of this analysis is the claim that the causative affix is a reduplicative one. To be more specific, it was maintained that causatives are derived through partially reduplicating the base form (e.g. RED, ktb). The reduplicant takes the form of an underspecified segment that copies the features of the segment it gets to duplicate. After being infixed, the RED affix copies the features of the second consonant of every root it is affixed to. For a detailed criticism of this analysis, see section 4 from chapter 4.

According to Boudlal (2001), the causative form is derived through the interaction of four main constraints. First, there is an alignment constraint, ANCHOR (Base, L, RED, L), which requires the reduplicant to be left aligned to the base. To account for the medial position of the geminate in the optimal form, this constraint must be dominated by *GEM-σ_1, a markedness constraint banning the occurrence of geminates in initial syllables.

Also, there is another alignment constraint, ALIGN-E (Root, PWd), which demands a complete match between the edges of the root and the prosodic word. The role of this constraint consists in preventing the total reduplication of the base by dominating the faithfulness constraint MAX-Rt-BR. The violation of ALIGN-E (Root, PWd) and ANCHOR (Base, L, RED, L) is assessed gradiently, in the sense that the degree or multiplicity of violation in terms of distance from the designated edge is what makes a candidate pass or fail. The following tableau provides a visual demonstration of these constraint interactions:

(45)

/RED, ktb/	*GEM-σ1	ALIGN-E (Root, PWd)	ANCHOR (Base, L, RED, L)	MAX-Rt-BR
☞ a. kə**t**.təb			*	**
b. **k**ək.təb	*!	**		**
c. kə**k**.təb	*!			**
d. kət.**b**əb			**!	**
e. kə**b**.təb			**!	**
f. **k.təb**. k.təb		****!		

This tableau shows that candidate (45a) is the optimal output as it satisfies the structural demands of the higher-ranked constraints: *GEM-σ1 and ALIGN-E (Root, PWd). Candidates (45b) and (45c) are immediately ruled out by the undominated constraint *GEM-σ_1. Furthermore, candidate (45b) incurs multiple violations of ALIGN-E (Root, PWd) since the left edge of the prosodic word no longer corresponds to the left edge of the root. The latter constraint also excludes candidate (45f), which represents a case of total reduplication. As for candidates (45d) and (45e), they seem to tie with the optimal candidate at the level of *GEM-σ1 and ALIGN-E (Root, PWd). The constraint ANCHOR (Base, L, RED, L) becomes the deciding factor in this case. Here, it unties what seems to be a deadlock through gradient assessment, where candidates (45f) and (45e) fare worse on this constraint; hence, [kət.təb] wins out at the end.

7. Conclusion

In this chapter, we have argued that morphological gemination in MA is achieved via the affixation of a moraic affix to a root, which later translates into a word-medial geminate. More specifically, we have made the claim that the designated mora is initially suffixed to the root. The infixation of the mora, however, happens under the pressure of certain phonological requirements, namely syllabic well-formedness. We have postulated that there is an alignment constraint which stipulates that a mora should be right-aligned to the edge of the root, hence ALIGN-R (μ). The realization of the mora in the output is ensured by the faithfulness constraint Max-affix, militating against the non-parsing/non-filling of the moraic affix in the input. The alignment constraint is believed to be outranked by a markedness constraint on prosodic well-formedness, σWF. This very specific ranking forces the aligned mora not to be prohibited from attaching to the right edge or any edge for that matter because it will always cause some violation of prosodic markedness. The latter was also shown to be responsible for the emergence of the shape invariant templates of the forms under study. This effect is obtained through the ranking of σWF over CONTIGUITY.

GENERAL CONCLUSION

This dissertation has been concerned with the phonological and morphological patterning of geminates in MA. Precisely, we have addressed the issues of geminate distribution, geminate representation, geminate behavior, geminate syllabification, word minimality, prosodic lengthening, schwa epenthesis and morphological gemination. Besides, this dissertation has also tackled issues related to the status of the root in word derivation and the interaction between morphology and phonology. Concerning the latter issue, this work has specifically explored the kinds of effects that morphology and phonology can exert on the distribution and patterning of geminates as well as the role they could play in deriving invariant templatic forms. Several chapters have been devoted to the investigation of these topics. The content of each chapter is summarized below.

Chapter 2 has provided a taxonomy of geminate consonants in MA based on how they can be created, where in a word they can occur and what type of sounds they can be. It has been shown that while geminates in MA are basically lexical, as they contrast phonemically with their singleton counterparts, they can also be phonologically and morphologically derived. Phonologically, it has been revealed that geminates can arise through the processes of total assimilation and prosodic lengthening. Morphologically, it has been disclosed that geminates may also function as morphological markers in MCs, ANs and INs. In addition, this chapter has demonstrated that geminates in MA may come from all the major class features and appear in all word positions. The other major contribution that chapter two has achieved is distinguishing between fake and true geminates in MA. In this concern, we have shown that true geminates, which can be underlying or derived, are mainly characterized by resisting schwa epenthesis. Fake geminates, however, do not resist schwa epenthesis, behaving more like typical sequences of unlike consonants.

Chapter 3 has dealt with the behavior and representation of geminates in MA. In this chapter, geminates have been shown to display a dual behavior that combines properties of singletons and clusters alike. In patterning with singletons, it has been demonstrated that geminates tend to resist schwa epenthesis. In patterning with clusters, it has been illustrated that geminates act as consonant sequences in connection with schwa triggering, syllabic heterosyllabicity and full vowel placement. In the face of this duality, various models of geminate representation have been evaluated and checked against the data of MA. In doing so, we have shown that, in nonlinear generative phonology, there are two major ways to represent geminates phonologically: they can be either segmentally represented as two skeletal slots linked to a single melodic element or prosodically represented as a mora underlyingly associated to a single root node. In this respect, we have discussed the predications and implications associated with each model as far as MA geminates are concerned. It has been revealed that both the skeletal model and the moraic model have limitations related to other aspects of MA phonology and morphology. On the one hand, the skeletal model has been proved unsuccessful in making the right predictions about syllable weight and prosodic size. On the other hand, the moraic model has been considered inadequate in explaining the bi-positional tendencies of geminates in MA. For this reason, this chapter has culminated in developing a hybrid model of geminate representation that combines a segmental level consisting of two root nodes and a prosodic level consisting of a mora.

The fourth chapter has provided evidence from MA for the moraic nature of geminates. Three major morpho-phonological phenomena have been relied on for this purpose: word minimality, syllabification, and morphological gemination. By referring to word minimality, it has been shown that monovocalic words with initial geminates (i.e. GV) meet the bimoraic threshold imposed on MA words. In this case, the geminate appears to contribute a second mora besides the mora of the vowel. In addition, since prosodic lengthening has been described to give rise to a moraic geminate structure, this was believed to constitute an argument in favor of the moraic conception of geminates in MA. Finally, it has been demonstrated that word formation by means of morphological gemination favors a moraic analysis over a segmental one thanks to the cross-linguistic generalizations it allows.

In chapter 5, the syllabification of geminate consonants is addressed. It has been argued that, despite being represented with two-root nodes at the segmental level, geminates in MA do not constitute complex syllable margins. The motivation behind this reasoning is to curb the overpredictive power of our moraic two-root node model. Instead, it has been suggested that the syllabification of geminates in MA follows exclusively from their moraic structure. On this basis, the heterosyllabicity that characterizes medial geminates has been shown to ensue from a demand on geminates to be parsed as onsets while their underlying moras are to be parsed as codas. Comparably, initial geminates in MA have also been claimed to get heterosyllabic under similar circumstances. Yet, in the absence of a coda position, the underlying mora of an initial geminate has been argued to project a syllable of its own. Final geminates, however, have been shown to get heterosyllabic for different purposes. While the underlying mora of a final geminate is parsed into a coda position, an empty-headed syllable is projected, for which the geminate works as a potential onset, in order to avoid contained geminates. On this note, it has been demonstrated that, whenever possible, the phonology of MA employs different strategies to avoid empty-headed syllables. Such strategies include: [i]-epenthesis in a non-final context for stem geminates and degemination before an onset for geminated affixes.

The sixth chapter of this dissertation has been devoted to the issue of geminate integrity. The main objective has been to provide a constraint-based explanation of the inconsistency that characterizes the integrity of geminates in MA. To this end, we have built on the stipulation of the GL (Benhallam, 1980), which stipulates that geminates may get split by morphological rules but not by phonological ones. Besides formalizing the generalization made by the GL in terms of constraint interaction, we solved many problems that faced it. In particular, we have argued that the GL applies equally to derived and underlying geminates, rejecting the revised version of Benhallam (1991), which claimed otherwise. The argument was that splitting a sequence of identical heteromorphemic consonants by means of schwa epenthesis does not qualify as a case of geminate breaking. In fact, derived geminates, like the ones created through total assimilation or morphological gemination, have been shown to exhibit the same properties of underlying ones. In addition, this chapter has accounted for a problematic class of broken plurals whose underlying geminates were found to be split by a schwa. In reaction to this anomaly, we have maintained that the geminates of the broken plurals in question are not split by schwa itself. Rather, it was claimed that they are split by means of morphological activity, namely the realization of the plural morpheme, in total compliance with the GL. In a similar vein, the systematic splitting of underlying geminates in PP derivation was also demonstrated to follow from morphological pressure, especially the regulation of the right edges of the root and the prosodic word.

Chapter 7 has motivated a root-based approach to word formation in MA. Particularly, we have done so by demonstrating that a word-based approach to deriving MCs and comparatives raises problems related to base inconsistency, base-to-output nonconformity and base unavailability. First, it has been shown that, in a word-based approach, medial weak and final weak verb bases could be stated either in the imperfective form or the perfective form, depending on the type of correspondence to be maintained between the base vowels and the vowels or glides of the derived causatives. Second, it has been demonstrated that the derivation of MCs and comparatives alike does not involve the transfer of some vocalic and affixal material of their corresponding output forms. Third, some causatives have been shown to lack bases that have a word-form status. In response to these problems, it has been suggested that MCs and comparatives should be derived from roots that are underspecified, minimal, and abstract.

The last chapter (chapter 8) has provided an Optimality-theoretic account for the derivation of MCs, ANs and INs. What these forms happen to have in common is that they are all characterized by morphological gemination. The central claim of our analysis has been that morphological gemination in MCs, ANs and INs is the result of the affixation of a consonantal mora. It has been argued that the realization of the moraic affix is ensured by the faithfulness constraint MAX-Affix while its distribution is regulated by the alignment constraint ALIGN-Right. The final position of the affix, however, is then decided by the well-formedness constraints on syllable structure. On this note, it has been shown that realizing the moraic affix on the edge tends to give rise to marked syllable structures, like syllabic consonants and empty-headed syllables. To avoid these structures, the mora is forced inside the root and realized on the second radical segment. This chapter has also accounted for the templatic shapes of MCs, ANs and INs by adhering to the Generalized Template Theory, whereby templates are emergent structures that follow from the interaction of independently motivated constraints. Interestingly, it has been demonstrated that the templatic shapes of these forms ensue from the same constraints postulated to account for morphological gemination.

The findings of this dissertation have far-reaching theoretical as well as empirical implications. To begin with, the theoretical importance of this work is many-fold. First, we have been able to propose and argue in favor of a new model for representing geminates with the potential to account for the cross-linguistic complex behavior that geminates display with regards to various phonological and morphological processes. In particular, the model proposed in this dissertation conceives of geminates as underlyingly moraic two-root node structures. This way of representing geminates has proved to be successful in explaining the complex patterning of geminates, in that it allows reference to a moraic prosodic level and a two-root node segmental level. This way we could reconcile the bi-positional properties exhibited by geminates in MA with the well attested moraic patterning of geminates across many languages of the world.

Second, as far as moraic theory is concerned, we have introduced some pieces of evidence that support the moraic conception of geminates in MA. The utility of this achievement consists in promoting moraic theory in the context of geminate representation, which has been a challenging aspect for the theory, especially when it comes to the moraic representation of edge geminates. As a result, thanks to the evidence provided, we have been able to incorporate MA geminates under the scope of a moraic approach to geminate representation, hence overcoming the hardships posed by the language's edge geminates and the bi-positional patterning of these geminates.

The third theoretical contribution of this dissertation has to do with the morphological theory of roots. In this concern, and in light of data pertaining to the formation of causative verbs and comparative adjectives, we have argued that some instances of word formation in MA have to be root-based. Assuming a root-based approach to word formation has been useful in obviating many problems of correspondence between the base forms and the derived forms. Especially, it has been shown that, in many cases of deriving causative verbs and comparative adjectives, only root material gets carried over to the derived forms.

Furthermore, this dissertation has weighed in on the issue of templatic morphology in the context of word formation in MA. Our contribution to this theoretical point was in the form of an extension of the root-and-prosody approach, set within Generalized Template Theory, to account for the derivation of new word forms in a different Semitic language, namely causative verbs, agent nouns and instrument nouns in MA. With respect to this, we have argued that templates are emergent structures that follow from the pressure exerted by well-formedness constraints on contiguity faithfulness.

Empirically speaking, this dissertation has featured a number of typologically generalizable and theoretically well-corroborated analyses of complicated phenomena that include: the syllabification of edge geminates in moraic terms, the behavior of geminates in relation to schwa epenthesis, the formation of the past participle and word derivation by means of morphological gemination.

In dealing with the syllabification of geminates in MA, we have developed an analysis that treats medial, initial and final geminates equally. We have argued that geminates can be moraic regardless of their position in a word as long as the mora is parsed in a mora-bearing position. On this basis, we proceeded to show that the heterosyllabicity of medial and initial geminates follows from two basic demands: (i) supply an onset and (ii) preserve the mora somewhere other than the onset (e.g. a coda or a nucleus). In the case of a final geminate, however, it has been argued that heterosyllabicity is motivated by the need to avoid contained geminates.

As far as geminate integrity is concerned, we have presented an analysis that proved successful in accounting for the inconsistent behavior of geminates in relation to schwa epenthesis and full vowel placement. Our analysis has capitalized on the idea that geminates can be split up by means of morphological processes but not phonological ones. With this in mind, we have shown how every case of geminate breaking is morphologically relevant, including those cases where geminates are counterintuitively split by a schwa.

We also considered the formation of past participles in the context of geminate integrity and beyond. We have argued for an analysis which assumes that the PP morpheme consists of the discontinuous affix /m-u/. In the context of geminate integrity, we have shown that the /u/ of the PP morpheme splits geminates due to an alignment demand to bring together the right edges of the root and the prosodic word. Then, we extended the analysis to the rest of the PP categories, whose challenging aspects were also accounted for. Especially, we have shown that the absence of the /u/ in some PP categories is the outcome of some phonological demands, including hiatus avoidance and output-output faithfulness.

Finally, this dissertation has developed a unifying account for the patterns of morphological gemination attested in MA, especially MCs, ANs and INs. In dealing with this morpho-phonological phenomenon, we posited that the derivation of those forms simply involves the affixation of a mora to a root base. Then, we proceeded to treat the more challenging aspect of this issue, which is the linearization of the affix material in relation to the root material. In this concern, we have argued that the position of the moraic affix is determined by the syllabic well-formedness of MA.

REFERENCES

Acquaviva, P. (2009). Roots and Lexicality in Distributed Morphology. *York-Essex Morphology Meeting, 5*, 1-21.

Al Ghadi, A. (1990/2014). *Moroccan Arabic Plurals and the Organization of the Lexicon*. D.E.S. Thesis. Rabat: Publications of Mohammed V University.

Al Ghadi, A. (1994). An OT Account of Moroccan Arabic Prosody. Ms. Delaware University, and Mohammed V University, Rabat.

Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.

Arad, M. (2003). Locality Constraints on the Interpretation of Roots: The Case of Hebrew Denominal Verbs. *Natural Language & Linguistic Theory, 21*, 737–778.

Aronoff, M. (1976). *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Bat-El, O. (1994). Stem Modification and Cluster Transfer in Modern Hebrew. *Natural Language and Linguistics Theory, 12*, 571-596.

Bat-El, O. (2003). The Fate of the Consonantal Root and the Binyan in Optimality Theory. *Recherches linguistiques de Vincennes, 32*, 31-60.

Bauer, L. (2002). What you can do with Derivational Morphology. *Morphology 2000*, 37-48.

Beckman, J. (1998). *Positional Faithfulness*. Doctoral dissertation. University of Massachusetts Amherst.

Benhallam, A. (1980). *Syllable Structure and Rule Types in Arabic*. Ph.D. Dissertation. University of Florida.

Benhallam, A. (1989/1990). Moroccan Arabic Syllable Structure. *Langues et Littératures, 8*, 177-191.

Benhallam, A. (1991). On Geminates in Moroccan Arabic. *Linguistica Communicatio*, 28-40.

Benkirane, T. (1982). *Etude Phonétique et Fonction de la Syllabe en Arabe Marocain*. Thése de 3éme cycle. Université de Provence.

Benmamoun, E. (1999). Arabic Morphology: The Central Role of the Imperfective. *Lingua 108*, 175-201.

Benmamoun, E. (2003). Reciprocals as Plurals in Arabic. In J. Lecarme (Ed.), *Research in Arabic grammar II* (pp. 53–62). Amsterdam: John Benjamins.

Bennis, S. (1992). *La Formation du Causatif en Arabe Marocain*. Mémoire de D.E.S. Faculté des Lettres, Rabat.

Bennis, S. and E. Iazzi. (1995). Morphologie du Verbe Causatif en Arabe Marocain. *Langues et Littératures 13*, 79-96.

Bensoukas, K. (2001). *Stem Forms in the Nontemplatic Morphology of Berber*. Doctorat d'Etat thesis. Mohammed V University, Faculty of Letters, Rabat.

Bensoukas, K. (2006). Variable Syllable Weight in Amazigh. *Languages and Linguistics*, 18&19, 31-58.

Bensoukas, K. (2018). Against the Consonantal Root in Tashlhit. In A. Boumalk & K. Bensoukas (Eds.), *ASINAG 13, La Racine dans les Langues Chamito-sémitiques: Nature et Function* (pp. 53-90). Rabat: IRCAM.

Bensoukas, K. & A. Boudlal. (2012a). The Prosody of Moroccan Amazigh and Moroccan Arabic: Similarities in the Phonology of Schwa. In T. Borowsky (et al.). (Eds.), *Prosody Matters: Essays in Honor of Lisa Selkirk* (pp. 3-42). London: Equinox.

Bensoukas, K. & A. Boudlal. (2012b). An Amazigh Substratum in Moroccan Arabic: The Prosody of Schwa. *Langues et Littératures 22*, 179-221.

Benua, L. (1997). *Transderivational Identity: Phonological Relations between Words*. Ph.D. dissertation. University of Massachusetts Amherst.

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.

Blevins, J. (2005). The Typology of Geminate Inventories: Historical Explanations for Recurrent Sound Patterns. In *Seoul Linguistics Forum 2005* (pp. 121-137). Language Education Institute, Seoul National University.

Blevins, J. (2006). Word-based Morphology. *Journal of Linguistics,42*, 531-571.

Boudlal, A. (1993). *Moroccan Arabic Glides: A Lexical Approach*. D.E.S. Thesis. Faculté des Lettres, Rabat.

Boudlal, A. (2001). *Constraint Interaction in the Phonology and Morphology of Casablanca Moroccan Arabic*. Doctorat d'Etat thesis. Mohammed V University, Rabat.

Boukous, A. (1982). Les Contraintes de Structure Segmentale en Berbère (dialecte tachelhit). *Langues et littératures, 2*, 9-27.

Cantineau, J. (1950). Racines et Schèmes. In *Mélanges offerts à William Marçais* (pp. 119-124). Paris: A. Maisonneuve.

Casali, R. F. (1996). *Resolving Hiatus*. Ph.D. dissertation. University of California Los Angeles.

Casali, R. F. (1997). Vowel Elision in Hiatus Contexts: Which Vowel goes? *Language*, 493-533.

Casali, R. F. (2011). Hiatus Resolution. In M. van Oostendorp, C. J. Ewen, K. Rice, & E. V. Hume (Eds.), *The Blackwell Companion to Phonology, 5 (2)* (pp. 1-27). John Wiley & Sons.

Chomsky, N. (1957). *Syntactic Structure*. Mouton.

Chomsky, N. & M. Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Cohn, A. C., Ham, W. H., & Podesva, R. J. (1999). The Phonetic Realization of Singleton-geminate Contrasts in Three Languages of Indonesia. In *Proceedings of the 14th International Congress of Phonetic Sciences, 1*, 587-590.

Comrie, B. (1981). Causative Construction. In *Language Universals and Linguistic Typology* (pp. 158-177). Oxford: Blackwell.

Clements, G. N. & Keyser, S. J. (1983). *CV Phonology: a Generative Theory of the Syllable*. Cambridge, MA: MIT Press.

Clements, G. N. (1985). The Geometry of Phonological Features. *Phonology, 2(1)*, 225-252.

Curtis, E. (2003). *Geminate Weight: Case Studies and Formal Models*. Doctoral dissertation. University of Washington.

Davis, S. (1994). Geminate Consonants in Moraic Phonology. *West Coast Conference on Formal Linguistics 13,* 32-45.

Davis, S. (1999a). On the Moraic Representation of Underlying Geminates: Evidence from Prosodic Morphology. In R. Kager et al. (Eds.), *The Prosody-Morphology Interface* (pp.39–61). Cambridge: Cambridge University Press.

Davis, S. (1999b). On the Representation of Initial Geminates. *Phonology 16*, 93–104.

Davis, S. (2003). The Controversy over Geminates and Syllable Weight. In C. Féry & R. Van de Vijver (Eds.), *The Syllable in Optimality Theory* (pp. 77–98). Cambridge: Cambridge University Press.

Davis, S. (2011). Geminates. In M. van Oostendorp, C. J. Ewen, E. Hume & K. Rice (Eds.), *The Blackwell Companion to Phonology, 2* (pp. 837-859). Malden, MA & Oxford: Wiley-Blackwell.

Davis, S. (2016). The Arabic Comparative and the Nature of Templatic Mapping in Arabic. In L. Körtvélyessy, P. Štekauer & S. Valera (Eds.), *Word-formation across Languages* (pp. 73-90). Cambridge Scholars Publishing.

Davis, S. & Ragheb, M. (2014). Geminate Representation in Arabic. *Perspectives on Arabic Linguistics XXIV-XXV* (pp. 3-19). Philadelphia/Amsterdam: John Benjamins Co.

Davis, S. & Topintzi, N. (2017). On the Weight of Edge Geminates. In H. Kubozono (Ed.), *The Phonetics and Phonology of Geminate Consonants* (pp. 260-282). Oxford University Press.

Dmitrieva, O. (2012). *Geminate Typology and the Perception of Consonant Duration*. Doctoral dissertation. Stanford University.

Dmitrieva, O. (2017). Production of Geminate Consonants in Russian: Implication for Typology. In H. Kubozono (Ed.), *The Phonetics and Phonology of Geminate Consonants* (pp. 34-65). Oxford University Press.

Goldsmith, J. (1976). *Autosegmental Phonology*. Ph.D. dissertation. Cambridge, MA: MIT Press.

Guerssel, M. (1977). Constraints on Phonological Rules. *Linguistic Analysis 3(3)*, 267-305.

Guerssel, M. (1978). A Condition on Assimilation Rules. *Linguistic Analysis 4(3)*, 225-253.

Hajek, J., & Goedemans, R. (2003). Word-initial Geminates and Stress in Pattani Malay. *Linguistic review, 20(1)*, 79-94.

Hall, N. (2006). Cross-linguistic Patterns of Vowel Intrusion. *Phonology, 23(3)*, 387-429.

Hammari, A. (2000). *Aspects of Language Change in Rhafsai Moroccan Arabic*. Doctorat d'Etat thesis. Mohammed V University, Faculty of Letters, Rabat.

Hammond, M. (1988). Templatic Transfer in Arabic Broken Plurals. *Natural Language & Linguistic Theory, 6(2)*, 247-270.

Hankamer, J. & Lahiri, A. (1988). The Timing of Geminate Consonants. *Journal of Phonetics, 16*, 327-338.

Harrell, R. (1962). *A Short Reference Grammar of Moroccan Arabic*. Washington, DC: Georgetown University Press.

Hart, M. E. (1991). The Moraic Status of Initial Geminates in Trukese. In *Annual Meeting of the Berkeley Linguistics Society, 17(1)*, 107-120.

Hayes, B. (1986). Inalterability in CV Phonology. *Language*, 321-351.

Hayes, B. (1989). Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry 20(2)*, 253-306.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. University of Chicago Press.

Haselwood, B., & Watson, J. (2013). The Arabic Definite Article does not Assimilate. *Leeds Working Papers in Linguistics and Phonetics, 18*, 34-53.

Hume, E., Muller, J. & van Engelenhoven, A. (1998). Non-moraic Geminates in Leti. *Phonology, 14(3)*, 371-402.

Hyman, L. (1985). *A Theory of Phonological Weight*. Dordrecht: Foris.

Idrissi, A., Prunet, J. F. & Béland, R. (2008). On the Mental Representation of Arabic Roots. *Linguistic Inquiry, 39(2)*, 221-259.

Imouzaz, S. (1991). *La Dérivation du Nom d'Action en Arabe Marocain de Casablanca: Approche Prosodique*. Mémoire de D.E.S. Faculté des Lettres, Rabat.

Imouzaz, S. (2002). *Interaction des Contraintes dans la Morphologie Nongabaritique de l'Arabe Marocain de Casablanca.* Doctoral dissertation. Université Hassan II, Casablanca.

Inkelas, S. (2006). Reduplication. In *Encyclopedia of Language and linguistics* (pp. 417-419).

Jones, A. (1999). Vowel Copy Epenthesis in Kolami. Ms., UCLA.

Kager, R. (1999). *Optimality Theory*. Cambridge, MA: Cambridge University Press.

Kawahara, S. (2005). Voicing and Geminacy in Japanese: An Acoustic and Perceptual Study. *University of Massachusetts occasional papers in linguistics, 31*, 87-120.

Kawahara, S. (2007). *The Emergence of Phonetic Naturalness*. Ph.D. dissertation, University of Massachusetts Amherst.

Kawahara, S. (2015). The Phonetics of Sokuon, or Geminate Obstruents. In H. Kubozono (Ed.), *Handbook of Japanese Phonetics and Phonology* (pp. 185-265). Berlin/Boston/Munich: Walter de Gruyter, Inc.

Kawahara, S & P. Melanie. (2017). Spectral Continuity, Amplitude Changes, and Perception of Length Contrasts. In H. Kubozono (Ed.), *The Phonetics and Phonology of Geminate Consonants* (pp. 13-33). Oxford University Press.

Kenstowicz, M. & C. Pyle. (1973). On the Phonological Integrity of Geminate Clusters. In M. Kenstowicz & C. Kisseberth (Eds.), *Issues in Linguistic Theory* (pp. 27-43). The Hague: Mouton.

Khattab, G. & J. Al-Tamimi (2014). Geminate Timing in Lebanese Arabic: the Relationship between Phonetic Timing and Phonological Structure. *Laboratory Phonology, 5(2),* 231-269.

Kiparsky, P. (2008). Fenno-Swedish Quantity: Contrast in Stratal OT. *Rules, Constraints, and Phonological Phenomena*. Oxford: OUP.

Kramer, R. (2007). Nonconcatenative Morphology in Coptic. In A. Kaplan & D. Teeple (Eds.), *Phonology at Santa Cruz, 7*. Santa Cruz, CA: Linguistics Research Center.

Kubozono, H., Ito, J., & Mester, A. (2008). Consonant Gemination in Japanese Loanword Phonology. In *18th International Congress of Linguistics, Seoul*, 953-973.

Kubozono, H. (Ed.). (2017). *The Phonetics and Phonology of Geminate Consonants*. Oxford University Press.

Kurisu, K. (2001). *The Phonology of Morpheme Realization*. Ph.D. dissertation, University of California Santa Cruz.

Laaboudi, D. (2003). *Issues in the Phonology of Bni Zeroual Moroccan Arabic: An Optimality-theoretic Account*. Doctorat National thesis. Mohammed V University, Faculty of Letters, Rabat.

Ladefoged, P. & I. Maddieson. (1996). *Sounds of the World's Languages*. Oxford: Blackwell.

Leben, W. R. (1973). *Suprasegmental Phonology*. Ph.D. dissertation. Cambridge, MA: MIT Press.

Leben, W. R. (1980). A Metrical Analysis of Length. *Linguistic Inquiry, 11(3)*, 497-509.

Lehmann, C. (2005). Latin Causativization in Typological Perspective. In *Actes du 13ème Colloque International de Linguistique Latine*, 1-18.

Levin, J. (1985). *A metrical Theory of Syllabicity*. Ph.D. dissertation. Cambridge, MA: MIT Press.

Lombardi, L. and J. McCarthy. (1991). Prosodic Circumscription in Choctaw Morphology. *Phonology, 8*, 37-71.

Luschützky, H. C. & F. Rainer (2011). Agent-noun Polysemy in a Cross-linguistic Perspective. *STUF-Language Typology and Universals Sprachtypologie und Universalienforschung, 64(4)*, 287-338.

Maddieson, I. (2008). Glides and Gemination. *Lingua, 118(12)*, 1926-1936.

McCarthy, J. (1979). *Formal Problems in Semitic Phonology and Morphology.* Ph.D. dissertation. Cambridge, MA: MIT Press.

McCarthy, J. (1981). A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry 12*, 373-418.

McCarthy, J. J. (1986). OCP effects: Gemination and Antigemination. *Linguistic inquiry, 17(2)*, 207-263.

McCarthy, J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, J. (2005). Optimal Paradigms. In L. Downing, T. A. Hall & R. Raffelsiefen (Eds.), *Paradigms in Phonological Theory* (pp.170– 210). Oxford: Oxford University Press.

McCarthy, J. (2008). *Doing Optimality Theory: Applying Theory to Data*. Oxford: Blackwell.

McCarthy, J. & A. Prince. (1986). *Prosodic Morphology*. Ms. University of Massachusetts, Amherst and Brandeis University.

McCarthy, J. and A. Prince. (1990a). Foot and Word in Prosodic Morphology: The Arabic Broken Plurals. *Natural Language and Linguistic Theory, 8*, 209-83.

McCarthy, J. and A. Prince. (1993a). *Prosodic Morphology I: Constraint Interaction and Satisfaction*. Ms. University of Massachusetts at Amherst and Rutgers University.

McCarthy, J. & A. Prince. (1993b). Generalized Alignment. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology* (pp. 79-153). Dordrecht: Kluwer.

Moravcsik, E. (1977). *On Rules of Infixing*. Bloomington: Indiana University Linguistics Club.

Morén, B. (1999). *Distinctiveness, Coercion, and Sonority: A Unified Theory of Weight*. Ph.D. dissertation. University of Maryland.

Morén, B. (2003). Weight Typology: An Optimality Theoretic Approach. *The Linguistic Review, 20*, 281–304.

Muller, J. S. (2001). *The Phonology and Phonetics of Word-initial Geminates*. Doctoral dissertation. The Ohio State University.

Noamane, A. (2014). *Aspects of the Morpho-phonology of Causative Verbs in Moroccan Arabic: An Optimality-theoretic Approach*. MA thesis. Lambert Academic Publishing.

Noamane, A. (2018). Morphological Causatives in Moroccan Arabic: Root-based or word-based? In A. Boumalk & K. Bensoukas (Eds.), *ASINAG 13, La racine dans les langues chamito-sémitique: nature et function* (pp. 217-240). Rabat: IRCAM.

Padgett, J. (2008). Glides, Vowels and Features. *Lingua, 118*, 1937-1955.

Pajak, B. (2009a). Context-dependent Perception of Geminates. In Poster presented at the 83rd LSA annual meeting, San Francisco.

Pajak, B. (2009b). Contextual Constraints on Geminates: The Case of Polish. In *Annual Meeting of the Berkeley Linguistics Society, 35(1)*, 269-280.

Pajak, B. (2013). Non-intervocalic Geminates: Typology, Acoustics, Perceptibility. *San Diego linguistics papers, 4*, 2-27.

Pater, J. & A. Werle. (2001). Typology and Variation in Child Consonant Harmony. In C. Féry, A., D. Green & R. van de Vijver (Eds.), *Proceedings of HILP 5* (pp.119-139). Potsdam: University of Potsdam.

Pater, J. & A. Werle. (2003). Direction of Assimilation in Child Consonant Harmony. *Canadian Journal of Linguistics, 48(3-4)*, 385-408.

Perlmutter, D. (1995). Phonological Quantity and Multiple Association. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 307-317). Oxford: Blackwell.

Podesva, R. (2000). Constraints on Geminates in Burmese and Selayarese. In *Proceedings of West Coast Conference on Formal Linguistics, 19*, 343-356. Cascadilla Press Somerville.

Podesva, R. (2002). Segmental Constraints on Geminates and their Implications for Typology. Paper presented at the LSA Annual Meeting, January.

Polgárdi, K. (2005). Geminates and Degemination in Hungarian. In International Conference on the Structure of Hungarian.

Prince, A. & P. Smolensky. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Malden, MA: Blackwell.

Prunet, J. F. (2006). External Evidence and the Semitic root. *Morphology, 16(1)*, 41-67.

Ratcliffe, R. (1990). Arabic Broken Plurals: Arguments for a two-fold Classification of Morphology. In M. Eid & J. McCarthy (Eds.), *Perspectives on Arabic linguistics II: Papers from the second annual Symposium on Arabic Linguistics* (pp. 94-119). John Benjamins Publishing.

Ratcliffe, R. (1997). Prosodic Templates in a Word-based Morphological Analysis of Arabic. In M. Eid & R. Ratcliffe (Eds.), *Perspectives on Arabic Linguistics X: Papers from the Tenth Annual Symposium on Arabic Linguistics* (pp. 147-171). John Benjamins Publishing.

Ratcliffe, R. (1998). *The "Broken" Plural Problem in Arabic and Comparative Semitic: Allomorphy and Analogy in Non-Concatenative morphology*. Amsterdam/Philadelphia: John Benjamins.

Ratcliffe, R. (2013). Morphology. In J. Owens (Ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 71-91). Oxford University Press.

Rguibi, S. (2001). *Phonological Variation and Change in Moroccan Arabic: The Case of the Tazi Dialect*. Doctorat d' État diss., M5 University, FLHS, Rabat.

Ridouane, R. (2010). Geminates at the Junction of Phonetics and Phonology. *Laboratory phonology, 10*, 61-90.

Rosenthall, S. (1994). *Vowel/Glide Alternation in a Theory of Constraint Interaction*. Ph.D. dissertation. University of Massachusetts Amherst.

Rosenthall, S., & Van Der Hulst, H. (1999). Weight-by-position by Position. *Natural Language & Linguistic Theory, 17(3)*, 499-540.

Saib, J. (1977). The Treatment of Geminates: Evidence from Berber. *Studies in African Linguistics, 8(3)*, 299-316.

Samek-Lodovici, V. (1993). A Unified Analysis of Crosslinguistic Morphological Gemination. In *Proceedings of CONSOLE-1*, 1-16.

Schein, B., & Steriade, D. (1986). On geminates. *Linguistic inquiry, 17(4)*, 691-744.

Selkirk, E. (1990). A Two-Root Theory of Length. *University of Massachusetts Occasional Papers 14* (pp. 123-171). Amherst, MA: GLSA publications.

Sherer, T. D. (1994). *Prosodic Phonotactics*. Ph.D. dissertation. University of Massachusetts Amherst.

Thurgood, G. (1993). Geminates: A Cross-linguistic Examination. *Papers in honor of Frederick H. Brengelman on the occasion of the twenty-fifth anniversary of the Department of Linguistics, CSU Fresno*, 129-139.

Thurgood, E., & Demenko, E. (2001). The Phonetic Realization of Phonologically Geminate Affricates in Polish: The Long and the Short of it. *Speech and Language Technology, 5*, 9-19.

Topintzi, N. (2008). On the Existence of Moraic Onset Geminates. *Natural Language and Linguistic Theory, 26(1)*, 147-184.

Tranel, B. (1991). CVC Light Syllables, Geminates and Moraic Theory. *Phonology, 8*, 291–302.

Trommer, J., & Zimmermann, E. (2014). Generalized Mora Affixation and Quantity-manipulating Morphology. *Phonology, 31(3)*, 463-510.

Tucker, M. A. (2010). Iraqi Arabic Verbs: The Need for Roots and Prosody. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project.

Tyler, S. A. (1969). *Koya: An Outline Grammar: Gommu Dialect*. Berkeley: University of California Press.

Ultan, R. (1975). Infixes and their Origins. In H. Seiler (Ed.), *Linguistic Workshop III* (pp. 157–205). Munchen: Fink, Structura 9.

Urbanczyk, S. C. (1993). Infixing and Moraic Circumscription. In T. Sherer (Ed.), *Phonological Representations: University of Massachusetts Occasional Papers, 16*, 319–57.

Ussishkin, A. (1999). The Inadequacy of the Consonantal Root: Modern Hebrew Denominal Verbs and Output-output Correspondence. *Phonology,16*, 401-442.

Ussishkin, A. (2000). *The Emergence of Fixed Prosody*. Doctoral dissertation. University of California Santa Cruz.

Ussishkin, A. (2006). Semitic Morphology: Root-based or Word-based? *Morphology, 16*, 37-40.

Ussishkin, A. (2007). Morpheme Position. In P. de Lacy (Ed.). *The Cambridge Handbook of Phonology* (pp. 458-472). New York: Cambridge University Press.

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

Yu, A. C. (2007). *A Natural History of Infixation*. New York: Oxford University Press.

Zimmermann, E. (2017). *Morphological Length and Prosodically Defective Morphemes*. New York: Oxford University Press.