Deriving Cairene Arabic from Modern Standard Arabic:

A Framework for using Modern Standard Arabic text to synthesize Cairene Arabic speech from phonetic transcription

Hannah Hassig

Universiteit Utrecht, UiL-OTS

Master's Thesis: Research Linguistics

June 2011

Supervisors: René Kager, Hugo Quené

Abstract

What is often termed 'Arabic' is in actuality a wide variety of languages, from the formal Modern Standard Arabic to many colloquial dialects. Dialects range from being extremely similar to one another (for example, Moroccan and Saharan Arabic) to unintelligible to one another (for example, Moroccan and Egyptian/Cairene Arabic). Modern Standard Arabic is the universal language of areas where these colloquial dialects are spoken, and is the language of government and formal affairs. Modern Standard Arabic is not spoken as a native language, instead it is the language learnt in a formal setting, and therefore is considered the 'second language' of its speakers.

Because of the different varieties of the Arabic language, the majority of text-to-speech efforts embrace the formal language as the medium from which to work. This is facilitated by the fact that colloquial Arabic is rarely written, and when written, usually lacks any standardized orthographic system. With this reasoning, it is possible to imagine a system that uses the text of Modern Standard Arabic to produce colloquial Arabic synthesized speech. The goal of this project is to do just that: transform Modern Standard Arabic via mapping and constraint-based Optimality Theory into Cairene Arabic (a colloquial dialect), then use the resulting output as input for a proposed speech synthesizer.

The phonological transformation begins with Modern Standard Arabic orthography. A one-to-one mapping system transforms this into the symbols of the International Phonetic Alphabet, reflecting the sounds and morphology of Cairene Arabic. Ranked constraints in the framework of Optimality Theory account for the phonological processes found in Cairene Arabic, on a lexical as well as a phrasal domain. These include lexical stress, emphasis spread, syncope, epenthesis, assimilation, and phrasal stress. Ultimately this produces fully-specified IPA for Cairene Arabic.

The output of the phonological transformation is thus the input for the speech synthesizer. Although not actually implemented in the study, a framework for such a synthesizer is provided. Using diphones as the unit of concatenative synthesis, the speech synthesizer is created using the open source software *Festival*. Additionally, the possible diphones for Cairene Arabic, as well as the nonsense strings of syllables to be recorded to obtain all the diphones are given. The actual implementation of such a synthesizer would produce Cairene Arabic speech, complete with assimilation processes and prosodic contours.

The study demonstrated that it is possible to transform Modern Standard Arabic into Cairene Arabic, although the two are usually considered different languages, rather than two forms of one universal language. This transformation was executed using different linguistic theories, including moraic theory, Optimality Theory, and rewrite rules. Additionally, the study demonstrated that, with sufficient time and resources, a speech synthesizer for a colloquial form of Arabic, Cairene Arabic, could be achieved.

Contents

1: Introduction

Throughout the countries of the world where Arabic is recognized as a national language, there exists what has been termed a diglossia. More specifically, a diglossia refers to the fact that there is a different 'language' used for daily affairs and another used for formal situations (government, education, et cetera) (Ferguson 1989). Consequently, many dialects of what will be termed Spoken Arabic (SA) are found all over the world, ranging from those of the Levant, Maghreb, and Gulf, among others. These dialects are not mutually intelligible; depending on the degree of spatial difference between the two dialects, speakers of one may be almost completely unable to understand speakers of another (compare, for example, the Spoken Arabic of Morocco with that of Syria). The dialects could be said to form a sort of continuum, where two dialects situated closer to each other on this continuum would render themselves more mutually intelligible to speakers as opposed to those dialects further apart on the continuum. However, the majority of the Arabic world has some knowledge (depending on education level) of Modern Standard Arabic (MSA), the 'formal' version found in the diglossia. MSA is taught in schools, to pupils who already speak a dialect of SA as their native language. Consequently, the status of MSA with respects to the dialects of SA has been widely debated. Most prominently, opinions differ as to whether MSA and SA are a part of the same language continuum, or constitute two separate languages.

Modern Standard Arabic and Cairene Arabic (the SA dialect of focus for this project) will be considered to be part of the same language continuum. Consequently, it should be possible to derive Cairene Arabic from Modern Standard Arabic, and vice versa, using different frameworks in linguistic theory. The phonological transformation is achieved via a combination of rewrite rules (both context-free and context-dependent) and constraint-based Optimality Theory.

Text-to-speech synthesis is a rapidly evolving technology, providing intelligible and natural-sounding synthesized voices. However, the dialects of Arabic have been left behind in the text-to-speech industry. Because of the prevalence of Modern Standard Arabic as the language used for texts, and the lack of colloquial Arabic texts, almost all efforts to create a text-to-speech synthesizer for Arabic concern Modern Standard Arabic. Using Modern Standard Arabic texts as input for a speech synthesizer is problematic when considering that the majority of Arabic texts are only partially diacriticized when written, and rely on a morpho-syntactic parser to determine the correct diacritization.

If it is possible to derive colloquial Cairene Arabic from Modern Standard Arabic via phonological frameworks, it should be possible to synthesize Cairene Arabic speech using Modern Standard Arabic text as input. Such a speech synthesizer has yet to be created. This project develops the framework for a speech synthesizer that outputs Cairene Arabic from Modern Standard Arabic text. The synthesizer uses the output of the phonological transformation from Modern Standard to Cairene Arabic as the input from which speech can be synthesizer. While the speech synthesizer is not actually implemented, how such an algorithm could be developed and implemented is discussed.

The second chapter is devoted to the phonetics, phonology and morphology of both Modern Standard Arabic and Cairene Arabic. In the third chapter, the phonological transformation from Modern Standard Arabic to Cairene Arabic is carried out, using rewrite rules and Optimality Theory. The fourth chapter discusses the history of speech synthesis, as well as the problems typically encountered when attempting to synthesize Arabic. The fifth chapter provides an overview of a hypothetical speech synthesizer using the open source program *Festival*. Finally, implications of the project as well as further research are discussed in the last chapter.

2: Phonetics, Phonology, Morphology and Syntax

1. Purpose

This chapter provides a phonological overview of both Modern Standard Arabic and Cairene, as well as motivation for the particular decisions involving the phonology. The phoneme inventories of both Modern Standard and Cairene Arabic are presented, with discussion about the particular nature of each of these inventories, and obvious differences between the two. The morphology of both Modern Standard and Cairene Arabic is also of interest: the morphology of verbs, nouns, pronouns and affixes are discussed, in addition to a brief overview of the syntax of each dialect. From there, Modern Standard Arabic is no longer discussed, as the phonological properties of interest to this project are only those of Cairene Arabic.

The syllable of Cairene Arabic is motivated as the unit on which phonological processes take place. However, the domain of on which phonological processes work is not merely the syllable, but also the prosodic word (here referred to as the utterance). This domain is then the basis for discussion on the different phonological processes of lexical stress, emphasis spread, and syllable repair processes.

Ultimately, the goal of this section is to providea background for the different phonological processes that occur in Cairene Arabic. In the second chapter, the transformation from Modern Standard Arabic to Cairene Arabic, as well as the phonological transformations that occur from the underlying to the surface forms of Cairene Arabic, will be discussed, using both mapping rules and ranked constraints using the framework of Optimality Theory. Therefore, this section is an overview of the different theories surrounding the phonological processes in Modern Standard and Cairene Arabic, as well as motivation for endorsing a particular framework, or using parts of different theories to form something entirely different.

2. Phoneme Inventory: MSA

Modern Standard Arabic is the ‘formal’ language that is the standard language across all countries of the world where Arabic is a national language. Modern Standard Arabic is the language used by the government, for educational purposes, and other formal situations; a colloquial dialect of Arabic is the language of daily affairs.

The majority of the Arabic world has knowledge, to some extent depending on education level, of Modern Standard Arabic. The language is taught in schools, to pupils who already speak a dialect of colloquial Arabic as a native language. Consequently, Modern Standard Arabic is the native language of no individual, as it is instead somewhat of a contrived language in order to facilitate communication between a large population. The contrived nature of Modern Standard Arabic is a result of Classical Arabic, the language written in the Koran and other historical Arabic texts. The majority of differences between Classical and

Modern Standard Arabic stem from the additions to Modern Standard Arabic to adjust to modern-day life (vocabulary for novel technologies, foreign borrowings into the language, et cetera), as well as some syntactic differences. However, Modern Standard Arabic is a contrived language based on Classical Arabic, the language of the Koran. Because of its contrived nature, the orthography directly reflects the pronunciation, thus resulting in a phonemic orthography.

The figures below depict the consonantal phonemes of Modern Standard Arabic, first written with orthography of the language, then using the International Phonetic Alphabet. When symbols appear in pairs, symbols to the left are voiced, those to the right are voiceless (with the exception of the emphatic consonants).

Figure 1: Consonantal Phonemes of MSA

	labial	dental	alveolar	emphatic	palatal	velar	uvular	pharyngeal	glottal
nasal	م		ن						
stop	ت		د ت	ض ط	ج	ك	ق		ء
fricative	ف	ذ ث	ز س	ظ ص	ش		غ خ	ع ح	ه
tap			ر						
approximate			ل	ل	ي	و			

	labial	dental	alveolar	emphatic	palatal	velar	uvular	pharyngeal	glottal
nasal	m		n						
stop	b		d t	dˤ tˤ	dʒ	k	q		ʔ
fricative	f	ð θ	z s	ðˤ sˤ	ʃ		ɣ x	ʕ ħ	h
tap			r						
approximate			l		j	w			

The same can be said for the vowels. The figures below depict the vowels (both long and short) and diphthongs of Modern Standard Arabic.

Figure 2: Vowels and Diphthongs of MSA

Short Vowels:

	Front	Central	Back
Close	ِ		ُ
Mid			
Open	َ		

	Front	Central	Back
Close	i		u
Mid			
Open	a		

Long Vowels and Diphthongs:

	Front	Central	Back	Diphthongs:
Close	ي		و	يَ
Mid				وَ
Open	ا			

	Front	Central	Back	Diphthongs:
Close	i:		u:	/aj/
Mid				/aw/
Open	a:			

The clear phonemic orthography of the language is demonstrated below, first in the orthography of Modern Standard Arabic, then in the International Phonetic Alphabet.

الكِتاب	البِنتُ	دَرَسَتْ
alkita:b	*albint*	*darasat*
the-book	the-girl	study-PERF

‘The girl studied the book’

The features of the phonemes of Modern Standard Arabic (using those of Watson 2002, with the notable addition of Retracted Tongue Root, RTR, for emphatic segments) are given here.

Figure 3: Features of MSA Phonemes

	m	n	b	d	t	$d^ʕ$	$t^ʕ$	dʒ	k	q	ʔ	f	ð	θ	z	s	$ð^ʕ$	$θ^ʕ$	ʃ	ɣ	x	ʕ	ħ	h	r	l	j	w
consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
sonorant	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
syllabic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
voice	+	+	+	+	-	+	-	+	-	-	-	-	+	-	+	-	+	-	-	+	-	+	-	-	+	+	+	+
spread glottis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
constricted glottis	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
continuant	+	+	-	-	-	-	-	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-
nasal	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
strident	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-
lateral	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
delayed release	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LABIAL	+		+									+																+
round	-		-									-																+
CORONAL		+		+	+	+	+	+					+	+	+	+	+	+	+						+	+	+	
anterior		+		+	+	+	+	+					+	+	+	+	+	+	+						+	+	+	
distributed		-		-	-	-	-	+					+	+	+	+	+	+	+						-	-	-	
DORSAL									+	+									+	+	+	+	+	+				
high									+	+									+	-	-	-	-	-				
low									-	-									-	+	+	+	+	+				
back									-	+									-	+	+	+	+	+				
tense									-	-									-	-	-	-	-	-				
RTR[1]	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	+	+	-	-	-	+	+	-	-	-	-	-

	i	u	a
high	+	+	-
low	-	-	+
back	-	+	-
rounded	-	+	-
tense	+	+	-

3. Phoneme Inventory: CA

Cairene Arabic is the spoken colloquial language found in Egypt's capital city and the surrounding area. Choosing Cairene Arabic as the dialectal Arabic focus of this project lies in the fact that Cairene Arabic (more loosely, Egyptian Arabic) is often considered the most widely understood dialect throughout the Arabic world. This wide range of intelligibility is a result of the dominance of Egypt in the Arabic media. For this reason, Egyptian Arabic is a popular dialect for second-language learners of Arabic (though many study only Modern Standard Arabic). Additionally, unlike most other forms of colloquial Arabic, Egyptian Arabic can be found in written format. Advertisements are often written in this form of

[1]Retracted Tongue Root. This feature is further discussed in Chapter 3.

colloquial Arabic, in addition to a few plays and novels within the past two decades. These transcriptions of Egyptian Arabic facilitate a systematic study of the phonological processes of this language, providing direct evidence for phonemes and sound patterns that would otherwise need the support of a native speaker. For these reasons, Cairene Arabic is the dialectical focus at hand.

However, because Cairene Arabic, in its written format, uses the orthography employed by Modern Standard Arabic, it is not safe to say that the orthography corresponds to the phonemes of Cairene Arabic in a one-to-one mapping. The phoneme inventory considered here for Cairene Arabic is a combination of different inventories, namely Harrel (1957), Watson (2002) and Youssef (2006). The phoneme inventories for the consonants and vowels of Cairene Arabic are as follows:

Figure 4: Phonemes of Cairene Arabic

	labial	alveolar	emphatic	palatal	velar	uvular	pharyngeal	glottal
nasal	m	n						
stop	b	d t	dˤ tˤ		g k	q		ʕ ʔ
fricative	v f	z s	zˤ sˤ	ʃ dʒ		ɣ x	ħ	h
tap		r						
approximate		l		j	w			

Short Vowels

	Front	Central	Back
Close	i		u
Mid			
Open		a	ɑˤ

Long Vowels and Diphthongs:

	Front	Central	Back	Diphthongs:
Close	i:		u:	/aj/
Mid				/aw/
Open		a:	ɑ:ˤ	

The majority of the phoneme inventory for Cairene Arabic was taken from Youssef (2006). Harrell (1957) considers /r ~ ɾ/ to both be phonemes of Cairene Arabic, depending on the individual speaker. Because the quality of the alveolar tap can vary depending on the speaker, the phoneme inventory as given here will be considered to contain only /r/.

The other discrepancies between the phoneme inventories, as given by Youssef (2006) and Harrell (1957), concern the emphatic phonemes of the language. Youssef (2006) posits the underlying emphatic phonemes of Cairene Arabic as follows:

[tˤ] [dˤ] [sˤ] [zˤ] [rˤ] [ɑˤ] [ɑ:ˤ].

Harrell (1957) posits the underlying emphatic phonemes as follows:

[tˤ] [dˤ] [sˤ] [zˤ] [(lˤ)]

From this, it is clear that [t□] [d□] [s□] [z□] are part of the phonemic inventory of Cairene Arabic; there are many minimal pairs demonstrating the existence of such phonemes, in all positions, as demonstrated:

[t] ~ [tˤ]

[ti:n] 'figs'	[tˤi:n] 'mud'

[d] ~[dˤ]

[dala:l] 'coquetry'	[dˤala:l] 'backsliding'

[s] ~ [sˤ]

[si:n] letter nameس	[sˤi:n] 'China'

[z] ~ [zˤ]

[zu:r] 'visit'	[zˤu:r] 'perjury'

(Harrell 1957: 71)

The existence of [lˤ] as a member of the phonemic inventory is doubtful; Harrell cites [lˤ] as a phoneme given that there is one word in the language (with morphological derivations) that contains this word: that for God, /ɑlˤɑh/. Indeed, all lexical items cited by Harrell and others as evidence for the existence of [lˤ] in the phoneme inventory of Cairene Arabic contain the vowel /ɑ/. Additionally, emphatic /lˤ/ does not occur in any instances with the vowels /i/ or /u/. [lˤ] can thus be considered an allophone of /l/, in the environment of another emphatic: the emphatic vowel /ɑˤ/. Taking this stance, the underlying forms of lexical items related to *God* are /ɑˤlɑˤh/, with emphatic /ɑ/ (already determined to be a phoneme of the language) instead emphatic /l/.

Finally, the presence of emphatic /r ˤ/ is controversial: Youssef considers this a phoneme, while Harrell and others do not. Youssef argues that /r ˤ/ can appear in environments other than /ɑˤ/, existing in 'the onset of a syllable containing a long non-low vowel or a short high vowel, and no other pharyngealized sounds in the word' (2006:15). However, only two examples are provided:

[rˤo:ħi] 'my soul'

[rˤu:ħi] 'go' (third person feminine singular) (Youssef 2006: 15)

Harrell provides two other examples for the existence of phonemic /r□/:

[rˤubˤ] 'quarter'

[murˤrˤi] 'my bitter' (Harrell 1957: 72)

These minimal instances of emphatic /rˤ/ do not seem to be indicative of the presence of a phoneme. Therefore, the phonemic inventory as given here does not have /rˤ/ as a phoneme, and instead [rˤ] is an allophone of /r/ in surface representations. The near-minimal pairs and the existence of /rˤ/ in lexical items without any other underlying emphatic phonemes occur as a result of free variation: an optional condition that /r/ may be emphatic has previously been posited, and will not be further discussed here.[2] Thus, lexical items such as /rˤo:ħi/ will surface as [ro:ħi], and also appear as /ro:ħi/ in the underlying form.

The features of Cairene Arabic phonemes are given below, taken mainly from Youssef (2006):

Figure 5: Features of CA Phonemes

	m	n	b	d	t	dˤ	tˤ	g	k	q	ʔ	v	f	z	s	zˤ	sˤ	ʃ	ɣ	x	ʔ	ħ	h	r	l	j	w
consonantal	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-
sonorant	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+
syllabic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
voice	+	+	+	+	-	+	-	+	-	-	-	+	-	+	-	+	-	-	+	-	+	-	-	+	+	+	+
spread glottis	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
constricted glottis	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
continuant	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-
nasal	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
strident	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-
lateral	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-
delayed release	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
LABIAL	+		+									+	+														+
round	-		-									-	-														+
CORONAL		+		+	+	+	+							+	+	+	+	+						+	+	+	
anterior		+		+	+	+	+							+	+	+	+	+						+	+	+	
distributed		-		-	-	-	-							+	+	+	+	+						-	-	-	
DORSAL								+	+	+								+	+	+	+	+	+				

[2]For more on the status of the free variation of emphatic /r/, refer to Hassig 2010: "Emphasis Spread in Cairene Arabic."

high								+	+	+								+	-	-	-	-	-				
low								-	-	-								-	+	+	+	+	+				
back								-	-	+								-	+	+	+	+	+				
tense									-	-								-	-	-	-	-	-				
RTR	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	+	+	-	-	-	+	+	-	-	-	-	-

	i	u	o	a	ɑˤ
high	+	+	-	-	-
low	-	-	-	+	+
back	-	+	+	-	+
rounded	-	+	+	-	-
tense	+	+	-	-	+
RTR	-	-	-	-	+

4. Comparison of the Phoneme Inventories

Demonstrated in the previous figures, Modern Standard Arabic and Cairene Arabic have many of the same phonemes. The phonemes that Modern Standard Arabic and Cairene Arabic share are:

/m/ /n/ /b/ /t/ /dˤ/ /tˤ/ /k/ /q/ /ʔ/ /f/ /z/ /s/ /ʃ/ /ɣ/ /x/ /ʕ/ /ħ/ /h/ /r/ /l/ /w/ /j/

/i/ /u/ /a/

However, Modern Standard Arabic has the following, not found in Cairene Arabic:

/dʒ/ /ð/ /θ/ /ðˤ/ /sˤ/

In contrast, Cairene has the following, not found in Modern Standard Arabic:

/g/ /zˤ/ /sˤ/

/ɑˤ/ /o/

Assuming Modern Standard Arabic and Cairene Arabic are considered ‘dialects’ of one language for purposes of the project. Historically speaking, it will be assumed here that Cairene Arabic (and the other dialects of Arabic), like Modern Standard Arabic, derived from Classical Arabic, deviating from the classical language and from each other on the basis of geographical location. Because Modern Standard and Cairene Arabic are both offshoots of the same historical language, Classical Arabic, it should be possible to map Cairene Arabic from Modern Standard Arabic.[3]

The following correspondences of Modern Standard Arabic phonemes to Cairene Arabic phonemes are mapped:

<u>*Modern Standard Arabic*</u> <u>*Cairene Arabic*</u>

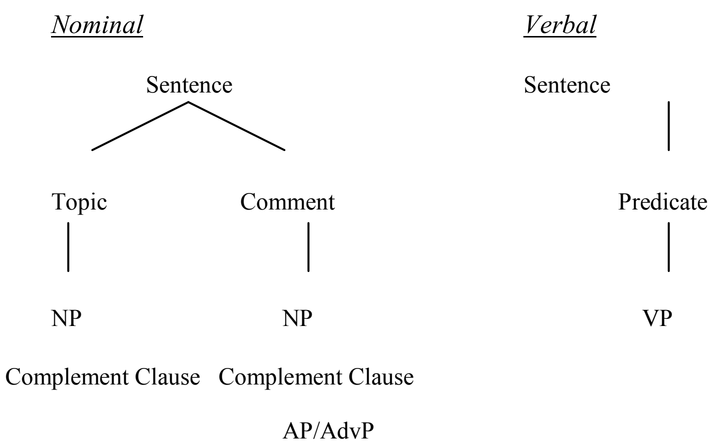
[3]The converse should also hold: it should be possible to map Modern Standard from Cairene Arabic. The validity of this does not pertain to the current project.

/dʒ/	/g/
/ð/	/z/
/θ/	/s/
/ðˤ/	/zˤ/
/θˤ/	/sˤ/

The differences in vowels will not be discussed here; instead, vowel differences between Cairene Arabic and Modern Standard Arabic are assumed to result from the surrounding consonantal environment, and therefore pronunciation is predictable by rule, discussed in the following chapter. However, looking at the phoneme differences between Cairene Arabic and Modern Standard Arabic reveals that Cairene Arabic has neutralized two pairs of phonemes of Modern Standard Arabic: Modern Standard Arabic /ð/ and /z/ are neutralized in Cairene Arabic as /z/, and /θ/ and /s/ are neutralized as /s/. Along the same line, the emphatic /ðˤ/ and /θˤ/ in Modern Standard Arabic become /zˤ/ and /sˤ/ in Cairene Arabic. Cairene Arabic transforms the Modern Standard Arabic phoneme /dʒ/ into /g/ depending on the context. This will be discussed further.

5. Morphology: MSA

Before delving into the morphology of Modern Standard Arabic, the syntax must be discussed to some extent. Modern Standard Arabic, in its fully vocalized form, employs case endings; as a result, the language possesses almost entirely free word order. However, in its spoken form, case endings are usually not pronounced (with the exception of Quranic readings). Consequently, the syntax must determine the different cases of the nouns. The language is described by Arab grammarians as containing two types of sentences: nominal and verbal. From these designations, the sentence is labeled as 'verbal' if it begins with a verb, and 'nominal' if it begins with anything else, be this element a Topic, Subject, or Object. (Ditters 1992: 110). Schematically, this looks as follows:



Within this distinction exist a variety of sentence structures. Sentences with the copula ‘to be’ are considered nominal sentences; the verb ‘to be’ is not pronounced in the present tense, leaving the nominal element (either the subject or the topic) in the first position. An example:

MSA (IPA transcription):	al-bint-u	fi:	al-bajt-i
Gloss:	the-girl-NOM	in	the-house-GEN
Translation:	The girl is at home (in the house).		

Within the verb sentence, there are also a variety of structures. The ‘unmarked’ verbal sentence is VSO—Verb, Subject, Object. Other possibilities for verbal sentences are:

Order	*Use*	
VSO	unmarked ordering	
VOS	object emphasis	
SVO	topic: subject	
SOV	topic: subject and object	
OVS	object focus	(Ditters 1992: 157-158).

Again, this free word ordering is possible through the employment of a detailed morphological system, which allows for the identification of subject or object on the basis of case endings.

The elements available as subject or object in an utterance are nouns. These elements may take the following case endings:

Case Ending:	*(Feminine)*	*Arabic Terminology:*	*Traditional Terminology:*
/u/ or /un/	/tu/ or /tun/	*al-marfuu*	nominative
/a/ or /an/	/ta/ or /tan/	*al-mansˤuub*	accusative
/i/ or /in/	/ti/ or /tin/	*al-madʒruur*	genitive

Generally speaking, the bare case ending (without the /n/) is employed when the element in question is definite (i.e. preceded by *al*, ‘the’). Otherwise, the form of the case ending with the /n/ is used. Examples of this follow:

MSA:	*Gloss:*
al-bint-u	the-girl-NOM
bint-un	a girl-NOM

Nominative elements are taken to denote the subject of a sentence. Accusative and genitive elements denote object(s) of a sentence. However, this is on the most basic level, for sentences consisting solely of Verb, Subject and Object, in any of the previously mentioned orders. When additional elements are introduced (subordinate conjunctions, complement clauses, et cetera), the case endings may denote other elements than the most typical. For example, in a copula sentence in the present, the subject is denoted by the 'accusative' case ending. This will be set aside for current purposes.

The language contains two functional genders: male and female, in the most general sense. These genders surface in the language as morphological endings on adjectives, verbs, and pronouns. Concerning adjectives, the morphological distinction exists as:

Masculine:	- ø (no ending)		
Feminine:	-a / -at	*Arabic Orthography:*	ة

The adjective must also agree with the noun it modifies in terms of case endings. For example:

MSA:	al-bint-u	al-saɣir-at-u
Gloss:	the-girl-NOM	the-small-FEM-NOM
Translation:	the small girl (nominative)	

Pronouns in Modern Standard Arabic are morphologically uninflected. A complete list of the pronouns:

Singular		*IPA*	*MSA*	*Vocalized*
	First person	ʔana	انا	انا
	Second person (masculine)	ʔanta	انت	انتَ
	Second person (feminine)	ʔanti	انت	انتِ
	Third person (masculine)	huwa	هو	هُوَ
	Third person (feminine)	hiya	هي	هِيَ
Dual				
	Second person	ʔantumaa	انتما	انتُما
	Third person	humaa	هما	هُما
Plural				
	First Person	naħnu	نحن	نَحنُ
	Second person (masculine)	ʔantum	انتم	انتُم
	Second person (feminine)	ʔantunna	انتن	انتُنَّ
	Third person (masculine)	hum	هم	هُم

Third person (feminine)	hunna	هن	هنَّ

Pronouns in MSA may also be enclitics (suffixes), as direct or indirect objects, markers possession, and objects of prepositions:

Singular	*IPA*	*Arabic*	*Vocalized*
First person	-(n)i:	(ن)ي	(ن)ي
Second person (masculine)	-ka	ك	كَ
Second person (feminine)	-ki	ك	كِ
Third person (masculine)	-hu	ه	هُ
Third person (feminine)	-ha:	ه	هَ
Dual			
Second person	-kuma:	كما	كُما
Third person	-huma:	هما	هُما
Plural			
First Person	-na:	نا	نا
Second person (masculine)	-kum	كم	كُم
Second person (feminine)	-kunna	كن	كُنَّ
Third person (masculine)	-hum	هم	هُم
Third person (feminine)	-hunna	هن	هُنَّ

It is impossible to discuss the morphology of Modern Standard Arabic without an explanation of the triconsonantal root system. More specifically, a root refers to a collection of three (rarely, four) phonemes in a specific order, that denote a variety of semantically-related lexical items. The oft-cited example of a triconsonantal root is d-r-s, with the basic meaning of 'study'. This root can be modified to produce *darasa* 'he studied', *mudrasa* 'a place of study (school)', or *yudarris* 'he makes someone else study/he teaches,' among other forms. Concerning verbs, the language consists of fourteen (ten of which are actually used) forms, given with their stems (C1 C2 and C3 denote the root) as follows:

Form	Verb Stem (perfective)	Verb Stem (imperfective)	'Meaning'
I	C1 a C2 a C3	a C1 C2 a C3	to X (transitive or intransitive)
II	C1 a C2: a C3	u C1 a C2: i C3	causative or intensifying

III	C1 a: C2 a C3	u C1 a: C2 i C3	action directed at or together with another
IV	□a C1 C2 a C3	u C1 C2 i C3	causative or declarative
V	ta C1 a C2: a C3	ata C1 a C2: a C3	reflexive of form II
VI	ta C1 a: C2 a C3	ata C1 a: C2 a C3	reflexive of form III
VII	in C1 a C2 a C3	an C1 a C2 i C3	reflexive passive
VIII	i C1 ta C2 a C3	a C1 ta C2 i C3	reflexive-intransitive
IX	i C1 C2 a C3:	a C1 C2 a C3:	to turn (color)
X	ista C1 C2 a C3	asta C1 C2 i C3	reflexive of form IV

(Abu-Chacra 2007:118-122)

Not all ten forms are in use for all triconsonantal roots. Normally, the verb is used in five or six of the derived forms (Abu-Chacra 2007:116).

The morphological forms for tense are divided in Modern Standard as either perfective or non-perfective (past or non-past). The perfective for verbs is morphologically realized as follows (X refers to the verb stem):

<u>*Perfective:*</u>

Singular	*IPA*	*Arabic*	*Vocalized*
First person	X-tu	ت	تُ
Second person (masculine)	X-ta	ت	تَ
Second person (feminine)	X-ti	ت	تِ
Third person (masculine)	X-a		َ
Third person (feminine)	X-at	ت	َت
Dual			
Second person	X-tuma	تم	تُمَ
Third person (masculine)	X-a:	ا	ا
Third person (feminine)	X-ata:	تا	َتا
Plural			
First Person	X-na	نا	نا

Second person (masculine)	X-tum	تم	تُم
Second person (feminine)	X-tun:a	تن	تُنَّ
Third person (masculine)	X-u:	و	و
Third person (feminine)	X-na	ن	نَ

<u>*Imperfective:*</u>

Singular	*IPA*	*Arabic*	*Vocalized*
First person	a-X	أ-	أ-
Second person (masculine)	ta-X	ت-	تَ-
Second person (feminine)	ta-X-i:na	ت-X-ين	تَ-X-ينَ
Third person (masculine)	ya-X	ي-	يَ-
Third person (feminine)	ta-X	ت-	تَ-
Dual			
Second person	ta-X-a:ni	ت-X-ان	تَ-X-ان
Third person (masculine)	ya-X-a:ni	ي-X-ان	يَ-X- ان
Third person (feminine)	ta-X-a:ni	ت-X- ان	تَ-X-ان
Plural			
First Person	na-X	ن-X	نَ-X
Second person (masculine)	ta-X-una	ت-X-ن	تَ-X-ُنَ
Second person (feminine)	ta-X-na	ت-X- ن	تَ-X–نَ
Third person (masculine)	ya-X-una	ي-X- ن	يَ-X-ُنَ
Third person (feminine)	ya-X-na	ي-X-ن	يَ-X- نَ

In addition to an imperfect and perfect verbal tense, Modern Standard Arabic also has a marking to indicate the future. The future is specified through the addition of the particle *sa* to the imperfect form of the verb. Therefore, for the imperfect *ʔaktub* ‘I write,’ the future becomes *saʔaktub*, ‘I will write.’ There is also a full form of the particle, *sawfa*.

Consequently, ‘I will write’ can also be specified as *sawfaʔaktub.* The particle is not inflected for person or number. (Abu-Chacra 2007: 109).

Conjunctions in Modern Standard Arabic are notable in that clitics can attach to these elements. For example, the conjunction *liʔana* ‘because,’ takes the clitic form of the subject of the adjoining clause. This is demonstrated below:

IPA:	*baddalt*	*qami:sˁi:*	*liʔanahu*	*wasix*
Translation:	change-FIRST SING	shirt-MY	because-THIRD M. SING	dirty-PERF.
	‘I changed my shirt because it was dirty’			

The clitic *hu* for the third person masculine singular refers to the subject of the clause ‘it (my shirt) was dirty.’ Consequently, the clitic attaches itself to the conjunction. (Abu-Chacra 2007: 193). The same holds true for a number of other conjunctions. The conjunction *wa*, ‘and’ is a clitic that attaches itself to the following element in an utterance. This appears as *ʔaktub waʔadrus kul alyum* ‘I write and I study every day,’ where *wa* attaches to *ʔadrus* to become *waʔadrus.* Another example is the conjunction *fa* ‘then’ which works in the same was as *wa.*

Negation of the imperfective uses the particle *la*. Thus *la ʔaktub* denotes ‘I do not write.’ For the perfective, negation is achieved through the particle *ma*:. *Ma kataba* denotes ‘He did not write.’

Of course, the morphology of Modern Standard Arabic is much more complex than the introduction provided here. What has been discussed above will suffice for the current goals. For a concise and complete overview, see Abu-Chacras 2007.

6. Morphology: CA

Unlike Modern Standard Arabic, which has a syntax with almost entirely free word order, Cairene Arabic has a much more restricted syntactic structure. Cairene Arabic consistently employs the syntactic structure SVO (Subject Verb Object). Because of this strict word ordering, Cairene Arabic abandons the case endings that allow for the free word order in Modern Standard Arabic.

Cairene Arabic has a binary gender distinction. The genders are marked as morphological endings on adjective and nouns, and within the morphology of the verb and pronoun. Concerning adjectives, the morphological distinction is:

Masculine: - ø (no ending)

Feminine: -a

The pronouns in use in the language are:

Singular	IPA
First person	ʔana
Second person (masculine)	ʔinta
Second person (feminine)	ʔinti
Third person (masculine)	huwwa
Third person (feminine)	hiyya
Dual	
Second person	ʔintum
Third person	humma
Plural	
First Person	ʔiħna
Second person (masculine)	ʔintum
Second person (feminine)	ʔintum
Third person (masculine)	humma
Third person (feminine)	humma

Pronouns in Cairene Arabic occur also as enclitics (suffixes), as direct or indirect objects, markers possession, and objects of prepositions:

Singular	IPA
First person	-ni
Second person (masculine)	-ak
Second person (feminine)	-ik
Third person (masculine)	-uh
Third person (feminine)	-ha
Dual	
Second person	-kum
Third person	-hum
Plural	

First Person	-na
Second person (masculine)	-kum
Second person (feminine)	-kum
Third person (masculine)	-hum
Third person (feminine)	-hum

Because Cairene Arabic derives from the same source as Modern Standard Arabic, it also employs a triconsonantal root system. Therefore, the morphology for the verbs (stems) occur in both the perfective and the non-perfective (past and non-past). Ten forms are also used, and are presumed to have the same general meaning as those in Modern Standard Arabic.

Perfective

Form	MSA Verb Stem	Egyptian Arabic Verb Stem
I	C1 a C2 a C3	C1 a C2 a C3
II	C1 a C2: a C3	C1 a C2: a C3
III	C1 a: C2 a C3	C1 a: C2 i C3
IV	ʔa C1 C2 a C3	ʔa C1 C2 a C3
V	ta C1 a C2: a C3	it C1 a C2: a C3
VI	ta C1 a: C2 a C3	it C1 a: C2 i C3
VII	in C1 a C2 a C3	in C1 a C2 a C3
VIII	i C1 ta C2 a C3	i C1 ta C2 a C3
IX	i C1 C2 a C3:	i C1 C2 a C3:
X	ista C1 C2 a C3	ista C1 C2 a C3

Imperfective:

Form	MSA Verb Stem	Egyptian Arabic Verb Stem
I	a C1 C2 a C3	C1 C2 a C3
II	u C1 a C2: i C3	C1 a C2: a C3
III	u C1 a: C2 i C3	C1 a: C2 i C3
IV	u C1 C2 i C3	C1 C2 i C3
V	ata C1 a C2: a C3	t C1 a C2: a C3
VI	ata C1 a: C2 a C3	t C1 a: C2 i C3

VII	an C1 a C2 i C3	n C1 i C2 i C3
VIII	a C1 ta C2 i C3	C1 ti C2 i C3
IX	a C1 C2 a C3:	C1 C2 a C3:
X	asta C1 C2 i C3	sta C1 C2 a C3

The morphological inflections for the verbs (to indicate the pronoun) are as follows:

Perfective

Singular

First person	X-t
Second person (masculine)	X-t
Second person (feminine)	X-ti
Third person (masculine)	X
Third person (feminine)	X-it

Plural

First Person	X-na
Second person (masculine)	X-tu
Second person (feminine)	X-tu
Third person (masculine)	X-u
Third person (feminine)	X-u

Imperfective

Singular

First person	a-X
Second person (masculine)	ti-X
Second person (feminine)	ti-X-i
Third person (masculine)	yi-X
Third person (feminine)	ti-X

Plural

First Person	ni-X
Second person (masculine)	ti-X-u
Second person (feminine)	ti-X-u

Third person (masculine)	yi-x-u
Third person (feminine)	yi-X-u

In Cairene Arabic, the future tense is denoted through use of a particle that is prefixed to the inflected form of the verb in the imperfect. For all verbal inflections this particle remains the same: *ħa.* There is no full form of the particle.

Conjunctions in Cairene Arabic can be in some way considered clitics. The conjunctions *wa* ′and′ and *fa* ′then′ attach to the following element in the phrase. The conjunction *illi*′because′ can either stand alone or take the subject of the following clause, in cliticized pronoun form. This holds true for a variety of other conjunctions.

Interestingly, Cairene Arabic employs both a prefix and a suffix to indicate negation. The prefix, taken from Classical Arabic, is *ma*, while the suffix is *ʃi.* This method of negation is used for both the imperfect and the perfect verbal forms.

7. Discussion of Morphology

From the above, there are apparent differences between the morphology of Modern Standard and Cairene Arabic. Obviously, a major difference appears in the syntax of the languages: because of the use of case markings, Modern Standard Arabic has almost entirely free word order; Cairene Arabic has no such case markings, and therefore has a strict syntax of SVO. As previously mentioned, it will be assumed here that Modern Standard Arabic employs the word order SVO.

Modern Standard and Cairene Arabic both have a gender distinction: male and female. This gender distinction is morphologically realized on adjectives, nouns, and inflections on verbs. When considering pronouns, however, there is a large distinction between the two: Modern Standard Arabic has a ternery pronoun distinction (singular, dual, and plural), while Cairene Arabic has a binary distinction (singular and plural). Comparisons are given below:

Full Forms of Pronouns

Singular	MSA (IPA)	Cairene Arabic (IPA)
First person	ʔana	ʔana
Second person (masculine)	ʔanta	ʔinta
Second person (feminine)	ʔanti	ʔinti
Third person (masculine)	huwwa	huwwa
Third person (feminine)	hiyya	hiyya

Dual		
Second person	ʔantumaa	ʔintum
Third person	humaa	humma
Plural		
First Person	naħnu	ʔiħna
Second person (masculine)	ʔantum	ʔintum
Second person (feminine)	ʔantunna	ʔintum
Third person (masculine)	hum	humma
Third person (feminine)	hunna	humma

<u>*Bound Forms of Pronouns*</u>

Singular	MSA (IPA)	Cairene Arabic
First person	-(n)i:	-ni
Second person (masculine)	-ka	-ak
Second person (feminine)	-ki	-ik
Third person (masculine)	-hu	-uh
Third person (feminine)	-ha:	-ha
Dual		
Second person	-kuma:	-kum
Third person	-huma:	-hum
Plural		
First Person	-na:	-na
Second person (masculine)	-kum	-kum
Second person (feminine)	-kunna	-kum
Third person (masculine)	-hum	-hum
Third person (feminine)	-hunna	-hum

As seen above, Cairene Arabic uses the plural for both the Modern Standard Arabic dual and plural pronouns. The same holds true for the verbal inflections:

Perfective

	MSA	Cairene Arabic
Singular		
First person	X-tu	X-t
Second person (masculine)	X-ta	X-t
Second person (feminine)	X-ti	X-ti
Third person (masculine)	X-a	X
Third person (feminine)	X-at	X-it
Dual		
Second person	X-tuma	X-tu
Third person (masculine)	X-a:	X-u
Third person (feminine)	X-ata:	X-u
Plural		
First Person	X-na	X-na
Second person (masculine)	X-tum	X-tu
Second person (feminine)	X-tun:a	X-tu
Third person (masculine)	X-u:	X-u
Third person (feminine)	X-na	X-u

Imperfective

	MSA	Cairene Arabic
Singular		
First person	a-X	a-X
Second person (masculine)	ta-X	ti-X
Second person (feminine)	ta-X-i:na	ti-X-i
Third person (masculine)	ya-X	yi-X
Third person (feminine)	ta-X	ti-X
Dual		
Second person	ta-X-a:ni	ti-X
Third person (masculine)	ya-X-a:ni	yi-X-u

	Third person (feminine)	ta-X-a:ni	yi-X-u
Plural			
	First Person	na-X	ni-X
	Second person (masculine)	ta-X-una	ti-X-u
	Second person (feminine)	ta-X-na	ti-X-u
	Third person (masculine)	ya-X-una	yi-x-u
	Third person (feminine)	ya-X-na	yi-X-u

Additionally, Cairene Arabic lacks the gender distinction for pronouns in the plural form. Instead, the same form is used to indicate male or female plurals. Concerning the stems of the verbs, Modern Standard and Cairene Arabic are compared:

Perfective

Form	MSA Verb Stem	Cairene Arabic Verb Stem
I	C1 a C2 a C3	C1 a C2 a C3
II	C1 a C2: a C3	C1 a C2: a C3
III	C1 a: C2 a C3	C1 a: C2 i C3
IV	ʔa C1 C2 a C3	ʔa C1 C2 a C3
V	ta C1 a C2: a C3	it C1 a C2: a C3
VI	ta C1 a: C2 a C3	it C1 a: C2 i C3
VII	in C1 a C2 a C3	in C1 a C2 a C3
VIII	i C1 ta C2 a C3	i C1 ta C2 a C3
IX	i C1 C2 a C3:	i C1 C2 a C3:
X	ista C1 C2 a C3	ista C1 C2 a C3

Imperfective

Form	MSA Verb Stem	Cairene Arabic Verb Stem
I	a C1 C2 a C3	C1 C2 a C3
II	u C1 a C2: i C3	C1 a C2: a C3
III	u C1 a: C2 i C3	C1 a: C2 i C3
IV	u C1 C2 i C3	C1 C2 i C3
V	ata C1 a C2: a C3	t C1 a C2: a C3

VI	ata C1 a: C2 a C3	t C1 a: C2 i C3
VII	an C1 a C2 i C3	n C1 i C2 i C3
VIII	a C1 ta C2 i C3	C1 ti C2 i C3
IX	a C1 C2 a C3:	C1 C2 a C3:
X	asta C1 C2 i C3	sta C1 C2 a C3

The future tense is marked with different morphology in Modern Standard and Cairene Arabic, although both employ prefixes. This verbal prefix attaches to the imperfect verbal inflection, in both Modern Standard and Cairene Arabic. The prefix in Modern Standard Arabic also has a full lexicalized form, unlike Cairene Arabic.

Conjunctions tend to work the same way in both Modern Standard and Cairene Arabic. The main difference is that in Modern Standard Arabic, if there is the possibility for a conjunction to take the cliticized form of the pronoun as the subject of the following clause, this must be done. In Cairene Arabic, however, this is only an option, and conjunctions can stand alone without the cliticized form of the pronoun subject.

Negation is dealt with differently in Modern Standard Arabic and Cairene Arabic. In Modern Standard Arabic, negation of a clause depends on the tense of the clause; thus the perfect and imperfect forms of the verb are negated with a different lexical item. In Cairene Arabic, the same marking is used to negate both the perfect and the imperfect form of the verb. This marking is a prefix and a suffix, attached to the verbal form. The main difference here is that Modern Standard Arabic has a lexicalized item to indicate negation, while Cairene Arabic uses a combination of particles.

From this point onwards, the intricacies of Modern Standard Arabic will no longer be discussed. Following transformation of Modern Standard Arabic script into IPA, the IPA of Modern Standard Arabic is replaced with the IPA of Cairene Arabic, discussed in the following chapter. Therefore, only the prosody and phonology of Cairene Arabic follow.

8. The Syllable

In Cairene Arabic, syllables play an important role in different phonological processes, including emphasis spread, lexical stress, and vowel duration. Importantly, syllables canspan words. For example, the phrase *al-ragil al-kabir* ‘the big man’is syllabified as *al.ra.gi.lal.ka.bir*, with the syllable *lal* spanning across what is considered two lexical items. This tends to complicate phonological processes as they occur in Cairene Arabic; a phonological process is not constrained by lexical boundaries, but can instead spread throughout the phrase. For the moment, syllabification across lexical boundaries will be ignored; instead the focus will remain on syllabification solely within the lexical item.[4]

The syllable in Cairene Arabic takes one of the following forms:

CV	CVC	CVCC
	CV:	CV:C

A CV syllable is considered 'light,' a CVC and CV: syllable 'heavy' and a CVCC and CV:C syllable 'super-heavy'(Broselow 1976: 34).

From the syllable structure, it follows that complex onsets are prohibited in Cairene Arabic. 'Since three-consonant clusters are impossible… a syllable break will always two adjacent consonants,' with the exception of word-final CC. (Broselow 1976: 34). Additionally, Cairene Arabic follows the Maximal Onset Principal: 'a single consonant followed by a vowel always shares syllable with that (following) vowel' (Broselow 1976: 34).

There are also some restrictions within the dialect on syllable position. CV: is prohibited word-finally, a result of a history in which long final vowels were reduced to short final vowels (Watson 2002: 56). Additionally, 'the super-heavy syllables CVCC and CV:C may only occur as the final syllable of a microsegment,' defined as the prosodic word (Broselow 1976: 34). The same sentiment is echoed by Watson, who states that CV:C syllables are restricted to phonological-word-final position, and CVCC syllables are restricted to utterance-final position (2002: 58). How the language repairs instances where syllables occur in forbidden positions underlying is the result of vowel epenthesis or syncope.

Concerning restrictions on segments in specific syllable positions, no restrictions have been found. For CC clusters syllable-finally, there seems to be no obedience to any sort of sonority hierarchy, nor does there appear to be a pattern in which segments may appear in either position of the CC syllable coda.

As previously mentioned, it is possible to use moraic theory to discuss the syllable structure. According to Broselow, a light syllable is monomoraic, a heavy syllable bimoraic, and a superheavy syllable trimoraic (1976: 34). In this case, however, a trimoraic syllable would be exceptional, and there are arguments against such things, especially given the restriction of their occurrence (domain-finally).

An alternative viewpoint supported by multiple researchers (Hayes 1995, Watson 2002, Davis 2011, among others) can be considered. In this analysis, superheavy syllables are also considered to be bimoraic, and the final consonantal segment is considered to be extrametrical. By disregarding the 'special' status of superheavy syllables, lexical stress might be easier to consider. Along the same line, Davis argues that a non-geminate consonant is moraic 'if it is the syllable coda, but not if it is at the end of the word' (Davis talk 2011: 3).

[4]Syllabification across lexical boundaries will be taken into consideration in the following chapter, as this resyllabification is important for synthesizing speech of the language, as well as for phonological processes that occur across lexical boundaries, including epenthesis, emphasis spread, and phrasal stress.

By this viewpoint, it is possible to have a trimoraic segment, if a CV:C or CVCC syllable were to occur in a position other than domain-final. However, this is not the case, and it suffices to say that syllables in Cairene Arabic are maximally bimoraic.

9. The Lexical Domain and the Phonological Domain

The domain is taken to refer to the largest unit in speech that will be dealt with. As previously stated, the syllable plays an important role in the phonological processes that occur in Cairene Arabic. Therefore, one might be inclined to believe that the domain should consist of the phonological word, which is defined as the stem, any subject and object pronoun suffixes, the negative suffix, and the imperfect prefix, if applicable (Watson 2002: 61). However, because syllabification can occur across the boundaries between phonological words, the domain to be dealt with must be larger than the phonological word.

'The main domain of syllabification is the phonological phrase in Cairene Arabic' (Watson 2002: 61). Consequently, as syllables determine the phonological processes that occur, the main domain for synthesis is the phonological phrase. The lexical word, the previous domain, does not account for phonological processes including both vowel shortening and epenthesis. The phonological phrase is, according to Watson, 'generally isomorphic with the syntactic phrase and included genitive structures, noun phrases, verb phrases, and participle phrases' (Watson 2002: 62).

Eventually, syllabification must occur across the phonological phrase, as well as within the phonological word. Therefore, there will be an initial syllabification, within the phonological word; eventually all phonological words within a single phonological phrase will be resyllabified to span word boundaries. Thus the first domain to be dealt with concerns the phonological word (basically equivalent to the lexical word, with some exceptions). Following this, the domain of interest shifts to the phonological phrase. These will be dealt with in Chapter 3.

10. Syllable Repair and other Lexical Processes

Cairene Arabic, as previously discussed, has many restrictions on syllable structure. Consequently, when the underlying representation contains what would be an illegal sequence of segments or an illegal syllable, the grammar of the language is such that changes are made to 'repair' this illegal cluster.

10.1 Epenthesis

Because of the syllable structure and restrictions on the syllable positions, the language forbids CCC sequences. This is solved through epenthesis. 'Whenever three consonants are potentially juxtaposed within the utterance, epenthesis of [i] occurs between the second and third consonant. Within the word, but not across word boundaries, the epenthetic vowel is realized as [u] to the left of /u/' (Watson 2002: 64). Some examples of this is as follows:

Underlying Form	*Epenthesized Form*	*Gloss*
/kull/ + /hum/	/kulluhum/	'all of them'
/kunt/ + /hina/	/kuntihina/	'I was here'

Additionally, Cairene Arabic does not allow onset-less syllables. Consequently, when a syllable in the underlying form is vowel-initial, the language repairs this with glottal stop epenthesis. According to Watson, the epenthesis of this particular consonant is due to the fact that the glottal stop is the 'minimal consonant,' in terms of features (Watson 2002: 66). This epenthesis is a requirement for the definite article, the relative pronoun, and the second-person independent pronouns. An example is given below:

Underlying Form	*Epenthesized Form*	*Gloss*
/ilbint/	/ʔil.bint/	'the girl'

10.2 Vowel Shortening

Additionally, there is a restriction on the CV:C syllable: this syllable may only be domain-final. Consequently, when the syllable occurs at some other location within the domain, 'the long vowel is shortened to prevent the appearance of a non-phonological-phrase-final CV:C syllable' (Watson 2002: 66). An example of this follows:

Underlying Form	*Shortened Form*	*Gloss*
/ba:b/ + /kum/	/bab.kum/	'your pl. door'

10.3 Syncope

Cairene Arabic consistently employs syncope. This process does have some restrictions, however, namely that '…in Cairene the vowel affected is almost invariably high, /i/ or /u/, and syncope is blocked if the resulting structure would create an impermissible syllable' (Broselow 1992). Syncope only targets unstressed vowels; if syncope would create an illegal sequence, such as CCC, the process is blocked. Besides these restrictions, syncope has free reign throughout the language. 'Syncope of unstressed, high vowels in non-final monomoraic syllables occurs within the utterance whenever the output does not violate structure preservation. That is to say, /i/ is deleted from a monomoraic syllable if and only if flanked by vowel-final syllables—namely CV(:)CiCV(:)' (Watson 2002: 70).

However, syncope does interact with epenthesis. Namely, epenthesis must occur before syncope. This is illustrated below:

Underlying representation:	/bint/ + /kibi:ra/	'a big girl'

Epenthesis:	/bintikibi:ra/
Syncope:	/bintikbi:ra/

(Watson 2002: 71)

If the processes were reversed:

Underlying representation:	/bint/ + /kibi:ra/
Syncope:	/bintkibi:ra/
Epenthesis:	*/ bintikibi:ra/

Reversing the processes effectively blocks syncope; syncope is only able to occur as a result of epenthesis, otherwise syncope would create an illegal sequence of three consonants. Therefore, the phonological processes interact, and their order must be carefully preserved.

The vowel of epenthesis in Cairene Arabic is always /i/ (Watson 2002). Additionally, unstressed short vowel /a/ is prohibited in the language, and instead becomes /i/ (Broselow 1976). Typically, this is due to an ease of pronunciation, as a higher sonority nucleus is more harmonic than one of lower sonority (Prince and Smolensky 1993: 17). /a/ is more sonorous than /i/, but this sonorous vowel could be considered to be more difficult to pronounce in an unstressed environment, and therefore is replaced with less sonorous /i/.

10.4 j-Strengthening

There is also a process of pre-suffix lengthening, also called *j*-strengthening (Watson 2002: 203). This occurs when a vowel-initial suffix attaches to a vowel-final morpheme. In order to fill the empty syllable onset created by attaching a vowel-initial suffix to a vowel-final morpheme, the vowel (nearly always /i/) lengthens, forming the glide /j/ as the onset of the following syllable. An example follows:

Underlying form	*j-strengthening*	*Gloss*
/masˤri/ + /a/	/masˤri:ja/	'Egyptian' (fem. sing.)

10.5 Assimilation

Cairene Arabic has a widely-attested phonological phenomenon of assimilation of the definite article *al*. The *l* of the definite article assimilates with the following consonantal segment, if the consonant segmental is what has been termed a 'sun letter' (Arabic: حروفشمسية).The 'sun letters' are: /l/ /n/ /r/ /t/ /d/ /s/ /z/ /tˤ/ /dˤ /sˤ/ and /zˤ/, falling under the natural class of coronal segments. In this case, the /l/ from the definite article assimilates to become identical to the following segment. (Abu-Chacra 2007: 39) In terms of features, this is the assimilation of /l/ to a following coronal consonant. According to Watson, 'assimilation of /l/ is clearly a lexical process since /l/ within any other morpheme fails to assimilate totally to a following coronal obstruent' (2002: 218). An example is given below:

Underlying Form	Assimilation	Gloss
/al/ + /sa:fir/	/assa:fir/	'the travel'

Additionally, underlying coronal nasal /n/ in Cairene Arabic assimilate in place to the following adjacent consonant. The labial nasal /m/ also assimilates in place to the following velar plosive, but only within the phonological word. The assimilation of the coronal nasal /n/ may occur across phonological word boundaries. (Watson 2002: 235-237)

The exact ordering of these phonological processes will be motivated in the following chapter.

11. Lexical Stress

'Cairene Arabic is a language with word stress… mean[ing] that one of the syllables in a content word is perceived as prominent and receives main stress' (Watson 2002: 79). This word stress will hereby be referred to as lexical stress.

The issue of secondary stress in Cairene Arabic has been subject to debate. According to Watson, ′secondary stress is not perceived' (2002: 79). McCarthy cites evidence from Harrell 1960 and Abul-Fadl 1961, by which secondary stress is ascribed to every heavy syllable that is not primary-stressed (McCarthy 1979: 450). However, because the goal here is Cairene Arabic synthesized speech output, secondary stress is considered to be unnecessary to create speech with realistic auditory qualities. Therefore, only the primary lexical stress will be considered, which will then be used to determine primary utterance stress.

11.1 Harrell 1957

Possibly the first account of lexical stress in Cairene Arabic was by Harrell 1957. This account is summarized as:

(a) Stress the ultima if it is a superheavy syllable (CVCC or CV:C)
(b) Otherwise stress the antepenultimate syllable if the antepenult and penult are light syllables (CV), unless the preantepenult is also light.
(c) Otherwise stress the penultimate syllable.

Examples are provided for each of these patterns (Harrell 1957):

(a)	saka′ki:n	'knives'	(superheavy ultima syllable)
(b)	′buxala	'misers'	(antepenult syllable stressed)
	mayi′hibik	'he does not like'	(light preantepenult, penultimate syllable stressed)
(c)	mar′taba	'mattress'	(penultimate syllable stressed)

This account is mainly descriptive; it does not provide any logical reason for why the stress pattern of Cairene Arabic could feasibly work this way. Despite this, the account still has one important advantage: namely, there are no exceptions, and all known lexical items of Cairene Arabic fit within the pattern. It is still necessary, however, to determine the motivation behind the pattern of lexical stress in Cairene Arabic.

11.2 Mitchell 1960

Oft-cited rules for Cairene Arabic lexical stress assignment are as follows:

(a) Stress a final superheavy or CV: syllable
(b) Stress a penultimate heavy (CVC or CV:) syllable
(c) Stress the penultimate or antepenultimate syllable, whichever is separated by an even number of syllables from the closest preceding heavy syllable or, if there is no such syllable, from the beginning of the word.

(Mitchell 1960)

There are examples for each of these claims:

(a)	ka′tabt	‘I wrote’	(stress on a ultimate superheavy CVCC)
	fala′hi:n	‘peasants’	(stress on ultimate superheavy CV:C)
(b)	′be:tak	‘your house’	(stress on penultimate heavy CV:)
	da′rasni	‘he taught me’	(stress on penultimate heavy CVC)
(c)	mad′rasa	‘school’	(stress on penultimate)
	yix′talifu	‘they differ’	(stress on antepenultimate)

(Watson 2002: 80)

The above description of Cairene Arabic, at first glance, appears different than that of Harrell 1957. However, closer examination makes it clear that this is not the case: upon closer examination, it becomes clear that the two are equally accurate descriptions of lexical stress patterns in the language. Both account for the same patterns of stress, favoring stress assignment for superheavy and heavy syllables. However, the stress pattern by Mitchell as well as that by Harrell fail to provide any sort of reasoning as to why the stress pattern would work this way, and instead provide a description of the stress system.

11.3 McCarthy 1979

In his 1979 work ‘On Stress and Syllabification,’ McCarthy discusses lexical stress in Cairene Arabic, with reference to Harrell´s 1957 account. McCarthy notes that the description provided by Harrell ‘offers several notable peculiarities to an investigation of the relationship between heavy syllables and stress’ (1979: 446). This is followed by an attempt to further explain the relationship between syllables and stress.

McCarthy sets forth an argument for a ternary syllable weight distinction. ‘Word-internally, the stress rule contrasts light syllables with heavy syllables. Word-finally, stress lodges onto a superheavy syllable, but a word-final CVC syllable fails to attract the stress’ (1979: 446). This is summarized in stating that ´there are two binary syllable weight distinctions, light versus heavy word-internally, and light and heavy versus superheavy word/finally' (446).

Through use of metrical phonology to account for lexical stress in Cairene Arabic, McCarthy provides an explanatory model of lexical stress, linking lexical stress with syllable weight. In this account, binary feet are assigned from left to write to pairs of light syllables. This is then followed by a right-branching superstructure, which gathers up all feet and stray syllables in the lexical item (449). The entire tree created by this superstructure is then labeled according to the principle that a right node is strong if and only if it branches (450).If the node does not branch, then it is labeled weak. This account appears to provide a succinct explanation for the lexical stress patterns in Cairene Arabic.

11.4 Hayes 1995

In 1995, Hayes reworked the description of lexical stress patterns provided by Mitchell, following the analysis of McCarthy. According to this account, Cairene Arabic parses feet from left to right, fashioning moraic trochees, and forbids degenerate feet, described as 'subminimal feet which survive to the surface' (Hayes 1995: 87). The moraic trochee assigns bi-moraic feet, with the stress on the leftmost syllable within the foot. Foot extrametricality does not apply. The rules for lexical stress assignment are summarized as follows:

(a) Consonant Extrametricality	C →<C> /__]word
(b) Foot Construction	Form moraic trochees from left to right.
(c) Word Layer Construction	End Right Rule(assign stress to the head of the rightmost foot.)

Similar to the analysis by McCarthy, the analysis given by Hayes is explanatory, assigning lexical stress by a way of moraic promotion. As previously mentioned in the discussion of the syllable in Cairene Arabic, the idea of consonant extrametricality in word-final position seems to hold true.

For example, for the word /buxala/, lexical assignment occurs as follows:

(a) there is no final consonant.

(b) the feet are constructed, forming /(buxa)la/.

(c) stress is then assigned to the head of the foot (buxa), as this is the rightmost foot.

The end result is thus [(ˈbu.xa)la]

11.5 Al-Jarrah 2008

A 2008 account by Al-Jarrah gives a constraint-based analysis of Cairene Arabic lexical stress. Al-Jarrah posits the following constraints:

Lx = Pr	every lexical word must consist of a prosodic word.

NONFINAL	the final syllable should not be parsed into a higher prosodic structure.
PARSE- σ	a syllable must be parsed into a higher prosodic structure.
MAIN-RIGHT	align the head-foot with the word, on the right edge.
ALL-FEET-LEFT	align each foot with the word, on the left edge.
TROCHAIC	align the head-syllable with its foot, on the left edge.
FOOT-BINARY	feet are binary under a moraic analysis. (2008: 2)

PARSE is divided into two separate constraints: PARSE$\sigma_{>\mu\mu}$and PARSE$\sigma_{\leq\mu\mu}$. The first constraint refers to syllables with more than two moras, superheavy syllables. The second constraint refers to syllables with two or less moras, namely light and heavy syllables. The final constraint ranking given to account for lexical stress in Cairene Arabic is:

Lx = Pr >> TROCHAIC >> MAIN-RIGHT >> PARSE$\sigma_{>\mu\mu}$ >> NONFINAL >> PARSE$\sigma_{\leq\mu\mu}$

>> FOOT-BINARY >> ALL-FEET-LEFT (2008: 12)

An example of the constraints in action:

ʃa.ga.ra:t	Lx= Pr	TROCHAIC	MAIN-RIGHT	PARSE$\sigma_{>\mu\mu}$	NONFINAL	PARSE$\sigma_{\leq\mu\mu}$	FOOT-BINARY	ALL-FEET-LEFT
(ʃa.ga)(ˈra:t)					*		*	**
(ˈʃa.ga)ra:t			*!					
ʃa(ˈga. ra:t)				*!	*	*	**	*

The winner, (ʃa.ga)(ˈra:t), parses syllables greater than two moras: (ra:t). According to the analysis by Al-Jarrah, this syllable contains three moras, as word-final consonants are considered moraic. The syllables /ʃa/ and /ga/ are parsed to form one foot, which is bimoraic, because these syllables each contain a single mora. However, the winning candidate violates FOOT-BIN, because the final foot contains three moras. The revisal of the constraints for current purposes makes clear that word-final consonants are not moraic, and therefore there is no need for the parsing of a foot that is greater than two moras, because such a foot or syllable is illegal in Cairene Arabic.

This constraint-based account purported by Al-Jarrah works from the descriptions of Harrell and Mitchell, and the analysis provided by McCarthy. Thus far, the constraint-based analysis provides the most succinct way of accounting for lexical stress in Cairene Arabic. Additionally, the constraint-based analysis fits in with the desired framework to be used for the goals of this project.

12. Emphasis Spread

Arabic, and other related Semitic languages, contain a group of phonemes described as ‘emphatic.’ The properties of these phonemes vary across the Semitic languages; here the focus lies only on Cairene Arabic. However, emphatic phonemes clearly change the featural properties of neighboring segments, both progressive and regressive assimilation.

12.1 Definition of Emphasis

What exactly emphasis is has been the subject of a long debate. Physically speaking, emphasis has been defined alternately as a ‘spreading and raising of the tongue’ or an ‘elevation of the dorsum,’ and acoustically as a ‘thickness [or] heaviness’ (Lehn 1963: 29). Another study proclaimed emphasis to be the result of a different place of articulation: i.e. an emphatic /$s^ʕ$/ has a place of articulation that is further back in the mouth than its non-emphatic counterpart, the alveolar /s/ (Laufer and Baer 1988: 182).

Emphasis has also been considered a realization of a secondary articulation: velarization. This consideration has lingered over centuries, beginning with the Arab grammarian Sibawayhi in the eighth century, and continuing to the present. However, consider that ‘many of these researchers used the term ‘velarization’ loosely, and probably did not mean to imply that the sounds were velarized rather than pharyngealized… instead merely using a phonetic term rather than the phonemic term, emphatic’’(Laufer and Baer 1988: 183). The argument against emphasis as a realization of a velar secondary articulation is discovered when examining the phonemic inventory Arabic: velarization as the defining feature of emphasis ‘creates a problem for the uvular consonant [q], namely that the consonant suddenly has a primary articulation (uvular) and a secondary articulation (velar) in the same general area’ (Laufer and Baer 1988: 184). Laufer and Baer provide two different solutions, provided that velarization as a secondary articulation is the realization of emphasis. The first solution does not consider the uvular consonant [q] an emphatic segment, thus creating no issue for velarization as the secondary articulation involved in emphasis. The second solution considers [q] as an emphatic segment, but instead considers emphasis as a realization of either a secondary articulation of velarization, or a place of articulation difference, making [q] the emphatic counterpart to the velar [k], not considered an emphatic segment. Note that emphasis as a secondary articulation of velarization has not been experimentally proven, but still stands as a possible definition of emphasis.

What appears to be the most convincing definition of emphasis is a realization of a secondary pharyngeal articulation. In 1957, Jakobson claimed that emphasis was realized by a contraction of the pharynx. Emphasis has also been described as a realization of one of more of the following articulatory features:

> ‘(1) slight retraction, lateral spreading, and concavity of the tongue and raising of its back
>
> (2) faucal and pharyngeal constriction (*pharyngealization)*
>
> (3) slight lip protrusion or rounding (labialization)
>
> (4) increased tension of the entire oral and pharyngeal musculature’
>
> (Lehn 1963: 30-31).

The above four articulatory features overlap to some degree: the pharyngeal constriction in (2) seems to be almost, if not entirely, the same as the increased tension of the pharyngeal musculature described in (4). However, it is emphasized in the description of these four processes that emphasis may be the result of any one of the processes single-handedly, or a combination of two or more of the processes. As phonetic features, this cannot be the case: lip protrusion or rounding cannot be the sole feature of emphasis, as this defines an entirely different process: labialization as a secondary articulation, indicated by the /C^w/. Thus emphasis cannot be equivalent to labialization; emphasis must therefore occur as a result of pharyngeal activity.

Emphasis as the result of pharyngealization has also been proven experimentally. In a study examining the articulatory strategies used by Arabic and Hebrew speakers when producing both pharyngeal and emphatic segments, Laufer and Baer determined that all the subjects 'used a similar articulatory strategy when producing emphatics and pharyngeals: a constriction in the lower pharynx when compared with their non-pharyngeal or non-emphatic counterparts' (190). Therefore, as it stands, pharyngealization appears to be the best explanation for the realization of emphasis.

In terms of features, pharyngealization is an articulatory process affecting the root of the tongue. This stance is aligned with that put forth by McCarthy: emphasis 'is identified with the distinctive features [RTR] *(retracted tongue root)* which is present underlyingly on certain consonants' (1997: 232). The emphatic segments of Cairene Arabic are thus considered to have the feature [+RTR].

Emphasis spread is 'a process of assimilation by which a phonological feature, i.e. pharyngealization, extends over more than one segment through a regular pattern' (Youssef 2006: 12). The secondary articulation of the emphatic segment extends over a number of linguistically predetermined segments. In Cairene Arabic, emphasis spread involves bidirectional assimilation, both progressive assimilation and regressive assimilation. However, this does not indicate that the assimilation is equivalent on either side of the emphatic segment; rather, there are interesting differences in behavior depending on the direction of the spread' (Youssef 2006: 15). Emphasis can spread throughout a syllable, phonological word, and even span word boundaries. Due to its ability to assimilate segments to a specific secondary articulation, emphasis can be described as 'long distance assimilation of the pharyngealization feature throughout the phonological word domain, triggered by a segment that bears this feature contrastively' (Youssef 2006: 15). The exact nature of emphasis spread in Cairene Arabic has been subject to a bit of a debate. As previously mentioned, however, the emphatic segments in the phoneme inventory of Cairene Arabic are considered to be: [tˤ] [dˤ] [sˤ] [zˤ] and [ɑ(:)ˤ].

Emphasis is never realized as the feature of a single segment. Instead, in Cairene Arabic, 'the minimum domain of emphasis is the sequence CV' (Broselow 1976: 32). CV is also the minimum length of a syllable in Cairene Arabic, as previously discussed. Thus, 'emphasis is a function of the entire syllable' (Harrell 1957: 78).

12.2 Leftward Emphasis Spread

Leftward emphasis spread is the result of regressive assimilation. That is, segments assimilate to the secondary pharyngeal articulation found on a neighboring following segment.

According to a variety of sources, leftward emphasis spread in Cairene Arabic is absolute. Emphasis spread from the emphatic segment leftwards to the beginning of the phonological word. This has been reported by Harrell (1957) and Youssef (2009). This absolute spread of emphasis to the beginning of the word indicates that there are no leftward blockers; the feature [+RTR] thus spreads from the emphatic segment and attaches itself to every phoneme between itself and the beginning of the phonological word.

For example, the underlying /jibu:zˤ/, *malfunctions*, with the emphatic phoneme /z□/, is realized as follows:

Underlying form:	/jibu:zˤ/
Syllables:	/ji. bu:zˤ/
Emphasis spread:	/jˤiˤ. bˤu:ˤzˤ/
Surface form:	[jˤiˤ. bˤu:ˤzˤ]

The feature [+RTR] is clearly unblocked from spreading throughout the word, to the syllable containing the emphatic, [bˤu:ˤzˤ], and the word-initial syllable, [jˤiˤ].

12.3 Rightward Emphasis Spread

Rightward emphasis spread is the result of progressive assimilation. Namely, segments assimilate, and receive a secondary pharyngeal articulation, found on a neighboring preceding segment. Rightward emphasis spread, unlike leftward emphasis spread, is not absolute in Cairene Arabic. According to Youssef (2009), rightward emphasis spread is blocked by [i:], [I] and [e:]. This indicates that any syllable of which one of these vowels is the nucleus, will not be made emphatic.

For example, the underlying /sˤe:fi/, *my summer*, with the emphatic phoneme /s□/, is realized as follows:

Underlying form:	/sˤe:fi/
Syllables:	/sˤe: . fi/
Emphasis spread:	/sˤe: ˤ.fi/
Surface form:	[sˤe: ˤfi]

As evidenced above, the feature [+RTR] is blocked from spreading to the second syllable, /fi/, due to the presence of the high vowel [i].

Emphasis spread in Cairene Arabic (and other Semitic languages) has previously been accounted for using a variety of frameworks: descriptive rules, phonological rules, parallel structures, and feature geometry (Youssef 2006). Additionally, optimality theory has also been used to account for emphasis spread, and will be the focus for the goal at hand.

12.4 Optimality Theory and Emphasis Spread

Optimality Theory has previously been criticized for its use of strict constraints, considered by opponents to be unable to account for process-specific constraints. In his 1995 article using the framework of Grounded Phonology to account for emphasis spread in Palestinian Arabic, Davis states that his analysis of the data at hand 'present at least two challenges for [Optimality Theory]' (Davis 1995: 495). The first challenge involves process-specific constraints: there is no way using the framework of Optimality Theory to produce a constraint ranking that explains the derivations of emphasis spread. Additionally, there would be no feasible solution to explain how one direction of emphasis spread in Cairene Arabic is subject to grounded conditions, while the other direction is entirely unblocked (495). The second challenge lies in the explanation of the blocking segments in the process of emphasis spread: Davis states that 'given that pharyngealization is arguably phonological… it is unexpected in Optimality Theory for a feature of a vowel… to have the effect of blocking the spread of pharyngealization'(496). According to this account, Optimality Theory is unable to account for emphasis spread, both in its framework and in its theoretical explanatory value. That is, Optimality Theory cannot sufficiently explain why emphasis spread occurs in one direction but not the other, or why there is variation between the bidirectionality of emphasis spread.

A response to Davis by McCarthy (1997) demonstrated that Optimality Theory is indeed capable of explaining emphasis spread: emphasis spread can be accounted for by a simple ranking where markedness constraints outrank faithfulness constraints. These markedness constraints can be aligned in terms of directionality to the edges of prosodic words (thus accounting for spreading throughout the domain of the prosodic word domain). Therefore, Optimality Theory is perfectly capable of accounting for emphasis spread, as will be further demonstrated.

In an unpublished paper, Youssef explains emphasis spread in Cairene Arabic using Optimality Theory (MA Thesis, Tromsoe 2009). Youssef's constraints were proposed as follows:

ALIGN-V-PLACE[DOR]-L (*Pwd*)

> The left edge of a V-place [dor] feature must be aligned to the left edge of the prosodic word.

ALIGN-V-PLACE[DOR]-R (*Pwd*)

> The right edge of a V-place [dor] feature must be aligned to the right edge of the prosodic word.

*V-PLACE[COR,DOR]

> No segment should have both V-place [cor] and V-place [dor] features.

ALIGN-V-PLACE[DOR]-L (*syllable*)

The left edge of the V-place [dor] feature must be aligned to the left edge of the syllable.

ALIGN-V-PLACE[DOR]-R (*syllable*)

The right edge of the V-place [dor] feature must be aligned to the right edge of the syllable.

MAX V- PLACE[DOR]

Every V-place [dor] feature in the input has a correspondent V-place [dor] feature in the output.

(2009: 20-21)

Youssef implemented these constraints using the place [DOR] to indicate DORSAL. The place as indicated was a feature of Vowels (marked as V-PLACE), as formulated by the Parallel Structures Model. However, that particular model will not be implemented in the study at hand; consequently, the V-PLACE and DORSAL markings require revision, to be discussed in the following chapter.

13. Phrasal Stress

Because Cairene Arabic syllabifies across phonological word boundaries, and over the entire phonological phrase, phrasal (utterance) stress must be considered in addition to lexical stress. However, there are differing accounts for where main stress should be placed. It is certain that whichever syllable receives main stress in the utterance should be a syllable that received primary lexical stress.

According to Hellmuth 2004, phrasal stress occurs to the right of the phonological phrase. Therefore, the main stress should be placed on one of the rightmost elements that received lexical stress. This could be achieved through branching and promotion of the rightmost branch with lexical stress, similar to the analysis discussed for lexical stress by McCarthy 1979. Using constraints in the framework of Optimality Theory, this would be written as MAINSTRESSRIGHT.

An alternative view states that ‘main stress is placed on the second element of the phrase, and syllable repair processes take place where necessary to produce a fully syllabifiable string’ (Watson 2002: 62). According to this view, the elements with main lexical stress to the left of the utterance are promoted to receive main stress within the utterance.

Between these two completely opposite theories, the current focus will employ the idea purported by Hellmuth. The theory that phrasal stress occurs on the rightmost element of the prosodic phrase has been found in many of the world’s languages. Consequently, it seems more logical that Cairene Arabic follows the same pattern as the majority of the world’s languages, rather than promoting phrasal stress on a leftmost element in the phonological phrase.

14. Conclusions

In this chapter, the phonology and morphology of both Modern Standard Arabic and Cairene Arabic were discussed, as well as the different phonological processes, both lexically and across the utterance, in Cairene Arabic. Different theories have been used to motivate the phonological processes in Cairene Arabic, and this chapter attempted to select the motivated theory that best suited the current goals. The subsequent chapter provides an account to derive Cairene Arabic, in its surface form (written in IPA), from Modern Standard Arabic orthography, using mapping rules and Optimality Theory by way of ranked constraints.

3: Transformation from MSA to CA

1. Purpose

The previous chapter discussed both the phonetics and the morphology of Modern Standard and Cairene Arabic. This chapter provides motivation for and discussion of the phonological transformation from Modern Standard to Cairene Arabic.

The differences in syntax between Modern Standard and Cairene Arabic are not discussed; assuming an artificial standard word order, as in the previous chapter, for both languages solves potential difficulties. Phoneme substitutions occur from Modern Standard to Cairene Arabic by a series of rewrite rules, using the International Phonetic Alphabet. This is followed by lexical substitution—namely of the pronouns, by rewrite rules that involve full substitution of the Modern Standard Arabic form of the pronoun with that of the Cairene Arabic form.

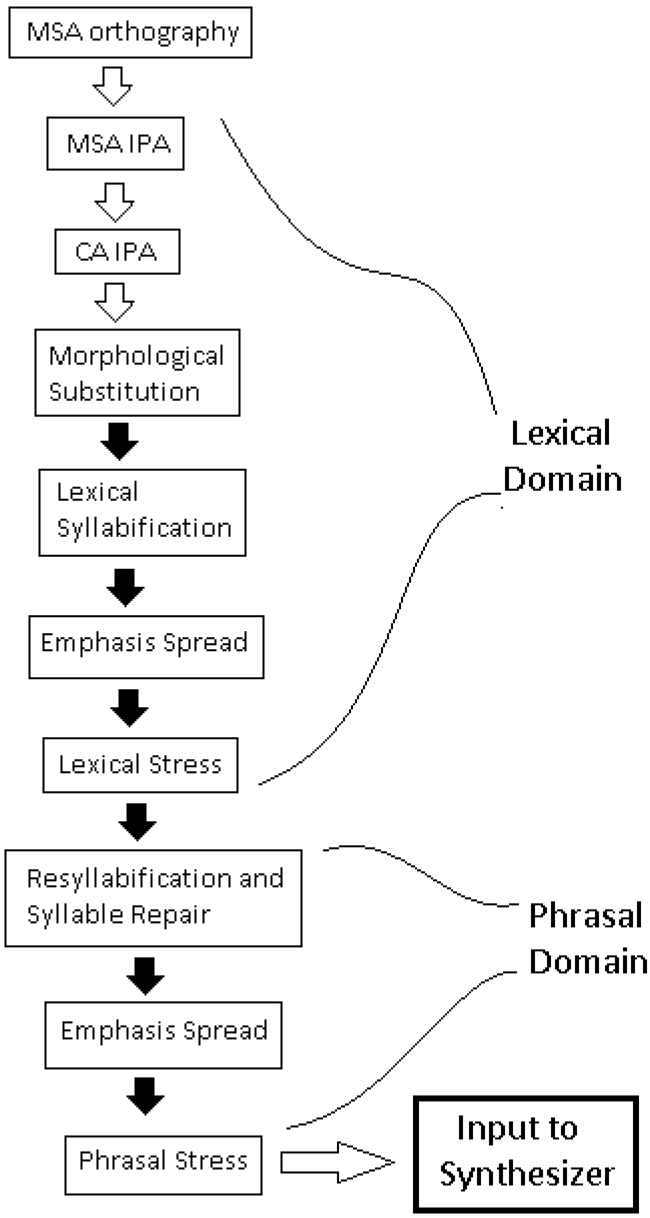
After these transformations via rewrite rules, Optimality Theory was used as the method of transformation. These transformations involved lexical syllabification, emphasis spread, and lexical stress assignment. On the phrasal domain, these transformations resyllabified the input and repaired syllables, emphasis spread, and stress assignment on the phrasal level.

Following this, the phonology of Cairene Arabic is discussed, beginning with syllabification on the lexical level. This is followed by stress on the lexical level. Resyllabification across the domain of the utterance occurs. Utterance-level phonology is then accounted for, including utterance stress (both primary and secondary) as well as syllable repairs. Finally, post-lexical phonological processes take place, most namely the process of emphasis spread.

In addition to the logic and motivation behind the theories presented, examples are provided to demonstrate each individual step, from Modern Standard Arabic orthography to the symbols of the IPA complete with all necessary phonological processes. Fully specified examples demonstrate the complete transformation from Modern Standard to Cairene Arabic, step-by-step through the algorithm. It is then shown that it is possible to implement the algorithm by a computer, using the open source program PRAAT to depict resyllabification and syllable repair. This demonstrated that it is possible to automaticize the algorithm.

2. From Mapping to Optimality Theory: A Diagram

The following diagram depicts the transformation from Modern Standard Arabic orthography to the input to the speech synthesizer. From the diagram, it is clear that the output of one step is the input to the next step. Each different level is diagrammed. The white arrows indicate that rewrite rules are used to achieve the following step, while black arrows indicate a use of constraint-based Optimality Theory from one step to another. In actuality, Optimality Theory does not occur in a step-wise fashion, and instead occurring as a parallel process. Therefore, the black arrows in the following diagram are used mainly for convenience of descriptive purposes.



The ultimate output of the transformation is the input to the speech synthesizer, to be discussed in the third and fourth chapters. The diagram depicts the shift in domain: the lexical domain is used for the transformation from MSA IPA up until the assignment of lexical stress. From then on, the phrasal domain is used for the processes of resyllabification and syllable repair, emphasis spread, and phrasal stress. The reasoning behind the different domains is described in the fifth section of this chapter. Each of the steps is further exemplified in the following sections, concluding with a discussion of the implementation of the various steps in an algorithm.

3. Phoneme Mapping

The initial input is Modern Standard Arabic written in Arabic orthography, and fully vocalized (sans case endings), with SVO word ordering, as previously discussed. The initial step is to map this orthography to the International Phonetic Alphabet. As stated in the previous chapter, this is a one-to-one mapping. The correspondence is given below can be found in Chapter 2, section 4.

The output of the mapping, beginning from Modern Standard orthography, is thus IPA, which will serve as the input for the next transformation. This transformation has Modern Standard Arabic in IPA as its input, and Cairene Arabic in IPA as its output. The mapping of this is as follows:

MSA	→	*CA*	*MSA*	→	*CA*
a:		a:	b		b
t		t	θ		s
dˤ		g	ħ		ħ
x		x	d		d
ð		z	r		r
z		z	s		s
θˤ		sˤ	sˤ		sˤ
dˤ		dˤ	tˤ		tˤ
ðˤ		zˤ	ʕ		ʕ
ʔ		ʔ	f		f
q		q	k		k
l		l	m		m
n		n	h		h
w		w	y		y
i		i	u		u
a		a	u:		u:
i:		i:			

As previously discussed, Cairene Arabic neutralizes phoneme distinctions in Modern Standard Arabic. The phonemes /ð/ and /z/ are neutralized in Cairene Arabic as /z/, and /θ/ and /s/ are neutralized as /s/. Along the same line, the emphatic /ðˁ/ and /θˁ/ in Modern Standard Arabic become /zˁ/ and /sˁ/ in Cairene Arabic. Cairene Arabic transforms the Modern Standard Arabic phoneme /d□/ into /g/, depending on the context.

Two examples of this transformation are given below:

(1) *MSA orthography*:	الكِتاب	دَرَسَت	ةليمجلا	البنت
(2) *MSA IPA*:	alkita:b	darasat	aldʒami:l	albint
(3) *CA IPA*:	alkita:b	darasat	algami:la	albint
(4) *Gloss*:	'The beautiful girl studied the book.'			

(1) *MSA orthography*:	الكِتاب	دَرَسَت	ةليمجلا	البنت
(2) *MSA IPA*:	alkita:b	darasat	aldʒami:l	albint
(3) *CA IPA*:	alkita:b	darasat	algami:la	albint
(4) *Gloss*:	'The beautiful girl studied the book.'			

The transformation, at this step, begins with Modern Standard Arabic orthography as input, which is then transformed via a one-to-one mapping to the symbols provided by the International Phonetic Alphabet. These symbols are then transformed via a mapping to the symbols of the International Phonetic Alphabet of Cairene Arabic. This is optimally achieved using context-free rewrite rules, as demonstrated above. Therefore, the ultimate output at this point is pseudo-Cairene Arabic, directly mapped from Modern Standard Arabic.

4. Morphology

Following the direct transformation from Modern Standard Arabic orthography to pre-Cairene Arabic, the morphology need be acknowledged. The transformation between Modern Standard Arabic to Cairene Arabic does not involve any phonemic changes within the pronouns, either full or bound. Consequently, the forms of the pronouns can thus be identified as Modern Standard Arabic, then transformed into their Cairene Arabic counterparts, via a mapping.

First, the full pronouns need to be identified. Full pronouns can only occur in subject position, indicating the first element in the phrase. Therefore, identification as Modern Standard Arabic pronouns must be done. This transformation is as follows:

Modern Standard Arabic	→	*Cairene Arabic*
ʔana		ʔana
ʔanta		ʔinta
ʔanti		ʔinti

huwwa	huwwa
hiyya	hiyya
ʔantumaa	ʔintum
humaa	humma
naħnu	ʔiħna
ʔantum	ʔintum
ʔantunna	ʔintum
hum	humma
hunna	humma

An example of this is given below:

(1)	*Input*:	naħnu	darasna:
(2)	*Identification*:	S = naħnu = first person plural pronoun	
(3)	*Transformation*:	ʔiħna	darasna:
(4)	*Gloss*:	'We studied.'	

Following this, the bound pronouns must be identified. Bound pronouns may occur as the Object in an utterance, or as Subject following a conjunction. Therefore, bound pronouns are always in the position X-*bound pronoun*, where X refers to either a verb or a conjunction. These elements need to be identified as the Modern Standard Arabic form, and then transformed to the Cairene Arabic form. The transformation is as follows:

Modern Standard Arabic	→	*Cairene Arabic*
-(n)i:		-ni
-ka		-ak
-ki		-ik
-hu		-uh
-ha:		-ha

-kuma:	-kum
-huma:	-hum
-na:	-na
-kum	-kum
-kunna	-kum
-hum	-hum
-hunna	-hum

Examples of this are given below:

(1)	*Input*:	albint	fa:rathu
(2)	*Identification*:	O = hu = third person masculine singular	
(3)	*Transformation*:	albint	fa:ratuh
(4)	*Gloss*:	'The girl won it.'	

(1)	*Input*:	albint ara:dat tezhib ila albeit liʔanaha: ka:nat mutaʔba
(2)	*Identification*:	S = ha: = third person feminine singular
(3)	*Transformation*:	albint ara:dat tezhib ila albeit liʔanaha ka:nat mutaʔba
(4)	*Gloss*:	'The girl wanted to go home because she was tired.'

The transformation of pronouns, from those of Modern Standard Arabic to those of Cairene Arabic, is again achieved by rewrite rules. In this case, however, the rewrite rules are no longer context-free. Instead, the context is determined by position: a full pronoun may only be found in the Subject position, while a bound pronoun may be found in either Subject position (after a conjunction) or Object position. Thus the rewrite is determined by the context, which would identify the position of the element and whether a bound or full pronoun is possible in this position.

After the identification (and transformation) of the pronouns, it is thus necessary to transform the verbs from their Modern Standard Arabic inflected forms to those of Cairene Arabic. The comparison below shows the differences between Modern Standard Arabic and Cairene

Arabic verb stems, where C1 refers to the first segment in the root, C2 to the second, and C3 to the third.

Perfective

Form	Modern Standard Arabic	Cairene Arabic
I	C1 a C2 a C3	C1 a C2 a C3
II	C1 a C2: a C3	C1 a C2: a C3
III	C1 a: C2 a C3	C1 a: C2 i C3
IV	ʔ a C1 C2 a C3	ʔ a C1 C2 a C3
V	ta C1 a C2: a C3	it C1 a C2: a C3
VI	ta C1 a: C2 a C3	it C1 a: C2 i C3
VII	in C1 a C2 a C3	in C1 a C2 a C3
VIII	i C1 ta C2 a C3	i C1 ta C2 a C3
IX	i C1 C2 a C3:	i C1 C2 a C3:
X	ista C1 C2 a C3	ista C1 C2 a C3

Imperfective

Form	MSA Verb Stem	Egyptian Arabic Verb Stem
I	a C1 C2 a C3	C1 C2 a C3
II	u C1 a C2: i C3	C1 a C2: a C3
III	u C1 a: C2 i C3	C1 a: C2 i C3
IV	u C1 C2 i C3	C1 C2 i C3
V	ata C1 a C2: a C3	t C1 a C2: a C3
VI	ata C1 a: C2 a C3	t C1 a: C2 i C3
VII	an C1 a C2 i C3	n C1 i C2 i C3
VIII	a C1 ta C2 i C3	C1 ti C2 i C3
IX	a C1 C2 a C3:	C1 C2 a C3:
X	asta C1 C2 i C3	sta C1 C2 a C3

However, only identification of the stem is not enough. The inflection of the stem must also be taken into account. Comparisons of the inflections are given below:

Perfective

Singular	MSA	Cairene Arabic
First person	X-tu	X-t
Second person (masculine)	X-ta	X-t
Second person (feminine)	X-ti	X-ti
Third person (masculine)	X-a	X
Third person (feminine)	X-at	X-it
Dual		
Second person	X-tuma	X-tu
Third person (masculine)	X-a:	X-u
Third person (feminine)	X-ata:	X-u
Plural		
First Person	X-na	X-na
Second person (masculine)	X-tum	X-tu
Second person (feminine)	X-tun:a	X-tu
Third person (masculine)	X-u:	X-u
Third person (feminine)	X-na	X-u

Imperfective

Singular	MSA	Cairene Arabic
First person	a-X	a-X
Second person (masculine)	ta-X	ti-X
Second person (feminine)	ta-X-i:na	ti-X-i
Third person (masculine)	ya-X	yi-X
Third person (feminine)	ta-X	ti-X
Dual		

	Second person	ta-X-a:ni	ti-X
	Third person (masculine)	ya-X-a:ni	yi-X-u
	Third person (feminine)	ta-X-a:ni	yi-X-u
Plural			
	First Person	na-X	ni-X
	Second person (masculine)	ta-X-una	ti-X-u
	Second person (feminine)	ta-X-na	ti-X-u
	Third person (masculine)	ya-X-una	yi-x-u
	Third person (feminine)	ya-X-na	yi-X-u

An example of the changes of verbal inflection:

MSA: /tataʕalum aldars/

CA: /titaʕalum aldars/

Gloss: *She teaches the lesson.*

To properly determine the verb and verbal inflection as they are transformed from Modern Standard to Cairene Arabic, the algorithm must formulate some sort of variable rule. In this case, variables would be used to assign C1, C2 and C3 to a particular verbal form. From these variables and the surrounding inflection, the algorithm could rewrite the Modern Standard Arabic form as the associated Cairene Arabic form, using a mapping. An example would be as follows, for the perfective of verbal form V:

MSA:	ata C1 a C2: a C3
Identify	C1 C2 C3
Map	ata → t
Map	C1 → C1
Map	a → a
Map	C2: → C2:
Map	a → a
Map	C3 → C3
Output:	t C1 a C2: a C3

In the case of future tense, this must also be dealt with. This can be done using a rewrite rule, where /sa/ or /sawfa/ are transformed to /ħa/, shown below.

MSA:	/sawfa tataʕalum aldars/
CA:	/ħatitaʕalum aldars/
Gloss:	*She will teach the lesson.*

Finally, sentences with negation must also be dealt with. This can also be done with a rewrite rule, by which the different lexical items used to indicate negation in Modern Standard Arabic are rewritten as the prefix and suffix (infix) on the verbal form of inflection in Cairene Arabic. This is indicated as follows, where X indicates the verbal form in Cairene Arabic.

Modern Standard Arabic		→	*Cairene Arabic*
ma	X		ma-X-ʃi
la	X		ma-X-ʃi
leisa	X		ma-X-ʃi
lan	X (+ future)		ma-X-ʃi (+ future)

An example of this rewrite rule:

MSA:	/ma ðaħabtu ila beiti:/
CA:	/mazaħabtuʃi ila beiti:/
Gloss:	*I did not go home.*

As is clearly demonstrated above, the lexical item used to indicate negation in Modern Standard Arabic, /ma/, is not a prefix of the verbal inflected form. In contrast, negation is indicated in Cairene Arabic by an infix to the verbal inflected form, /ma – ʃi/.

This identification of pronouns, inflections and negation could cause difficulties for the algorithm. Because the algorithm merely identifies the sequence of segments, if the same sequences coincidentally appeared in an utterance, the algorithm could be supposed to take this as the pronoun or inflection or negation. Therefore, there must be positional rules for the identification of such sequences. Because the syntax is presumed to be SVO, as previously discussed, the full pronoun can only appear in the S position, while bound pronouns may appear in either O or S position, attached to a verb or conjunction. The verbal inflections may only occur in V position. This should prevent the algorithm from improperly identifying a sequence as a pronoun or inflection.

Concerning the morphological parser as a whole, there are theoretically two ways in which the orthography of Modern Standard Arabic can be parsed. The simpler idea of the two employs a context-free rewrite rule. In this case, the algorithm merely identifies pre-specified strings of phones, and replaces these strings with the corresponding string of phones for Cairene Arabic. As discussed above, this could lead to improper identification: if a string of phones appears by coincidence that does not indicate a specific morphological entity, this string would be incorrectly replaced. However, this can be prevented by specifying where the replacement is allowed to occur.

A more theoretical analysis first determines the morphological components of the input. This is somewhat similar to the idea that replacement of strings requires some sort of specification as to where it can occur, as in the ad hoc rewrite rule. However, in this case, the input is analyzed, determining the morphological elements that occupy the subject and object positions in the input phrase. Following determination of the specific pronouns, the conversion to the corresponding pronouns of Cairene Arabic can be easily achieved. This is preferable to the context-free rewrite rule, as it involves both an analysis as well as a form of rewrite rules. One drawback to this analysis is that it would require more processing, when compared to the context-free rewrite rule. However, the theoretical benefits that result from determining the role of each element in the input phrase indicate that, most probably, this analysis would be subject to less incorrect replacements than the context-free rewrite rule. Therefore, the morphological component of the algorithm for the current purposes requires both a morphological analyzer as well as a converter.[5]

5. Stratal Optimality Theory

The ranked constraints to be used for current purposes can in no way be ranked so that there is only one level of constraint ranking. Instead, stratal Optimality Theory is employed, in which "morphology and phonology are stratified and interleaved…but the strata (Stem, Word, Postlexical) are characterized by systems of parallel constraints' (Kiparsky 2010: 1).

Two strata are employed: the lexical (word) and the postlexical (phrase). In the lexical stratum, the lexical items are individually subjected to the ranked constraints. These lexical items are then combined as the input to the phrasal strata (postlexical). Different strata are necessary to account for phonological opacity, of which prosodic morphology is a massive source (Kiparsky 2010: 2). Therefore, it is possible to have constraints on different strata that would, if on the same strata, cause phonological feeding or bleeding that would result in non-optimal outputs for the given language.

[5] The morphological analyser and converter at this point only consists of identification of S V and O objects in pre-set positions. A full discussion of morpho-syntactic parsers can be found in Chapter 4.

Although two different strata are defined for the current project, these strata each contain sub-strata. For the lexical stratum, there are two sub-strata: lexical syllabification/emphasis spread, and lexical stress. For the post-lexical stratum, there are two sub-strata: resyllabification and syllable repair/emphasis spread, and phrasal stress. Importantly, the sub-strata are not parallel. i.e. they do not occur simultaneously. Instead, stress assignment must follow any sort of syllabification processes. However, as a rule, as previously discussed Optimality Theory does not occur in a step-wise manner, but rather simultaneously, processing multiple levels at the same time. Therefore, the specific strata laid out that feed into one another are separated mainly for descriptive convenience.

Consequently, there are in total four strata to be used. The first strata consists of lexical syllabification and emphasis spread over the lexical domain, the second of stress over the lexical domain, the third of resyllabification, syllable repair and emphasis spread over the phrasal/post-lexical domain, and the fourth of stress over the post-lexical domain. The different strata are discussed in the following sections.

6. Lexical Syllabification

Lexical syllabification occurs prior to lexical stress, in order to allow stressing of syllables at the lexical level. Therefore, it is imperative that syllables be assigned within lexical boundaries, as opposed to syllabification at the utterance level (which will occur later, as resyllabification). This is due to the interaction between syllabification and moraic theory, which plays a role in the assignment of stress on a lexical level.

In this case, syllables will be determined by weight on the basis of moracity, in combination with the idea that, at this point, no segments may be added or deleted. Moras are assigned to syllables as follows:every vowel receives a mora; a syllable may consist maximally of two moras (Hayes 1995); word-final consonants are not moraic, while word-medial consonants that occur in a syllabic coda are moraic, in line with Hayes (1995) and Davis (2011). This mora-assignment is not directly a part of the constraints of the Optimality Theory at hand. Instead, it is rather artificially assumed that the algorithm somehow ‘knows’ how and where to assign moras, and syllables are created from this mora assignment. In addition, at this stage, above all, segments may not be added or deleted, creating an interaction of ranked constraints. This is discussed below.

The standard Optimality Theory constraints employed at this level are as given:

MAX-V	each V in the input must have a correspondent in the output.
DEP-V	each V in the output must have a correspondent in the input.
MAX-C	each C in the input must have a correspondent in the output.
DEP-C	each C in the output must have a correspondent in the input.
ONSET	syllables must have an onset (C) (Prince & Smolensky 1993: 17)

*COMPLEX	an onset or coda may not consist of more than one C.
*CCC	no sequences of three consonants (in or across syllables).
$*\sigma_{\mu\mu\mu}$	no syllable may contain more than two moras. This prevents closed syllables with long vowels or dipthongs word-medially.

As previously mentioned, at this point in the transformation, both vowel and consonantal segments may not be deleted or inserted. Consequently, the faithfulness constraints are highly-ranked: MAX-V, DEP-V, MAX-C and DEP-C. However, the markedness constraints ONSET, *COMPLEX and *CCC are the lowest ranked, as they are markedness constraints. At this point, *COMPLEXconsists of two separate constraints that have been merged: *COMPLEXONSET and *COMPLEXCODA, as they do not play a role. Similarly, *CCC does not play a role, since any candidate which did not violate *CCC would have had a faithfulness violation, and would therefore have already lost to other candidates before *CCC played any role. The highest-ranked constraint, $*\sigma_{\mu\mu\mu}$, essentially prevents complex codas from occurring non-word-finally.

Examples of the interactions of constraints are provided[6]:

albint	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
al.bint						*	*	
al.bi.nit			*!			*		
alb.int	*!					*	**	
a.lbint						*	**!	
al.bin				*!		*		
alb.ʔint	*!				*	*	**	*

In the above, the optimal candidate is /al.bint/. Any candidate that deletes a segment from the input is immediately eliminated via the highly-ranked faithfulness constraints. All of the candidates violate ONSET, as all of the candidates have initial onset-less syllable /al/. The candidates which do not divide the consonant cluster /lb/, and instead maintain the consonant cluster as an onset or coda, are eliminated with the constraint *COMPLEX. Thus, the optimal candidate becomes /al.bint/.

sakaki:n	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC

[6] The winning candidate is indicated by **bold** text.

sa.ka.ki:n								
sak.ki:n		*!						
sak.a.ki:n						*!		

In the above tableau, the optimal candidate is /sa.ka.ki:n/. The candidate that deletes a segment, /sak.ki:n/, violates a faithfulness constraint. Candidate /sak.a.ki:n/ is eliminated due to the onset-less syllable /a/.

At this point, no assignment of stress or higher moraic structure is discussed. Assignment of higher moraic structure (namely, the foot), in addition to assignment of lexical stress, appear in the seventh section of the current chapter.

7. Emphasis Spread

The constraints used for emphasis spread were first proposed by Youssef (2009), as discussed in the previous chapter:

ALIGN-V-PLACE[DOR]-L (*Pwd*)	The left edge of a V-place [dor] feature must be aligned to the left edge of the prosodic word.
ALIGN-V-PLACE[DOR]-R (*Pwd*)	The right edge of a V-place [dor] feature must be aligned to the right edge of the prosodic word.
*V-PLACE[COR,DOR]	No segment should have both V-place [cor] and V-place [dor] features.
ALIGN-V-PLACE[DOR]-L (*syllable*)	The left edge of the V-place [dor] feature must be aligned tothe left edge of the syllable.
ALIGN-V-PLACE[DOR]-R (*syllable*)	The right edge of the V-place [dor] feature must be aligned tothe right edge of the syllable.
MAX V- PLACE[DOR]	Every V-place [dor] feature in the input has a correspondent V-place [dor] feature in the output.

(2009: 20-21)

However, as previously mentioned, these constraints were implemented using the place [dor] to indicate DORSAL, and the place as a feature of Vowels (V-place), following the Parallel Structures Model. Because the current project does not include the Parallel Structures Model, the constraints are revised:

ALIGN-[+RTR]-L (*Pwd*)	The left edge of the feature [+RTR] must be aligned to the left edge of the prosodic word.
ALIGN-[+RTR]-R (*Pwd*)	The right edge of the feature [+RTR] must be aligned to the right edge of the prosodic word.

*[+FRONT, +RTR]	No segment should have both [+front] and [+RTR] features.
ALIGN-[+RTR]-L (*syllable*)	The left edge of the feature [+RTR] must be aligned to the left edge of the syllable.
ALIGN-[+RTR]-R (*syllable*)	The right edge of the feature [+RTR] must be aligned to the right edge of the syllable.
MAX[+RTR]	Every secondary articulatory place [dor] feature in the input has a correspondent place [dor] feature in the output.

The main change between the constraints posited by Youssef and the constraints above is the notation of secondary articulatory place features. Instead of V-place, the constraints were formed using the feature specification in chapter one: emphatic segments have a secondary articulatory place [dor] feature, [+RTR].

The directionality of the spread of the phonological feature of emphasis is accounted for within the framework of Optimality Theory through the use of alignment constraints. The alignment constraints concern the feature [+RTR] with respect to either the syllable or the prosodic word (*Pwd*). The alignment constraints each refer to a direction, left or right, that concerns the edge of either the syllable or the prosodic word. Therefore, directionality is accounted for within the prosodic word.

The ranking of the constraints determines the spread of the feature [+RTR]. In Cairene Arabic, the faithfulness constraint MAX[+RTR] must be highly ranked, as there is no evidence of deletion of underlying emphatic features of segments (that is, an underlying /s□/ is never realized as /s/). Thus the faithfulness constraint is undominated.

ALIGN-[+RTR]-R (*syllable*) and ALIGN-[+RTR]-L (*syllable*) must follow the faithfulness constraint in the strict ranking. Because the domain of emphasis spread is the syllable, each syllable must be either emphatic or non-emphatic—there is no evidence of a syllable with an emphatic onset and non-emphatic rime. These two constraints ensure that a syllable is either entirely emphatic or not, and are therefore highly ranked. The ranking of L >> R is arbitrary, and therefore the two are freely ranked with respect to one another.

*[+FRONT, +RTR] must be outranked by ALIGN-[+RTR]-L (*Pwd*), as leftward emphasis spread in Cairene Arabic is unblocked. However, *[+FRONT, +RTR] must outrank ALIGN-[+RTR]-R (*Pwd*), since rightward emphasis spread is blocked by vowels that have the feature [+FRONT].

The constraints are ranked as follows:

MAX[+RTR] >>

ALIGN-[+RTR]-L (*syllable*), ALIGN-[+RTR]-R (*syllable*) >>

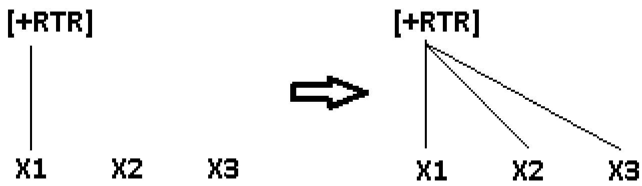
ALIGN-[+RTR]-L (*Pwd*) >>

*[+FRONT, +RTR] >>

ALIGN-[+RTR]-R (*Pwd*) >>

DEP [+RTR]

From the above constraints, the feature [+RTR] originates from one segment and spreads onto adjacent segments. Therefore, it is not assumed that the feature [+RTR] is ‘added’ onto any segments. Instead, this can be diagrammed as below, using an autosegmental representation to demonstrate that the feature [+RTR] is not duplicated:



Examples of the constraints in action from Cairene Arabic are given below:

/sˤe: . fi/	MAX [+RTR]	ALIGN-[+RTR]-L (*syllable*)	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN-[+RTR]-R (*Pwd*)	DEP - [+RTR]
sˤe: . fi			*!			*	
sˤe: ˤ . fi					*	*	*
se: . fi	*!						
s ˤe: ˤ . fˤi			*!		*	*	**
sˤe: ˤ . fˤiˤ					**!		***

In the tableau above, the feature [+RTR] originates from the word-initial segment. Because emphasis spread is unblocked within syllables, the feature spreads onto the adjacent segment. However, the feature [+RTR] is prevented from spreading onto the adjacent syllable, /fi/, because of the high front vowel, /i/. Therefore, the optimal candidate is /s□e:□.fi/.

/jɪ . bu:zˤ/	MAX [+RTR]	ALIGN-[+RTR]-L (*syllable*)	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN[+RTR]-R (*Pwd*)	DEP [+RTR]
jɪ . bu:zˤ		*!*		*			
jɪ . bu:z	*!						
jɪ . bu:ˤzˤ		*!		*			*
jɪ . bˤu:ˤzˤ				*!			**
jɪˤ . bˤu:ˤzˤ		*!		*	*		***
jˤɪˤ . bˤu:ˤzˤ					*		****

The optimal candidate is /jˤɪˤ.bˤu: ˤzˤ/. The feature [+RTR] originates from the word-final segment. The emphatic feature spreads left through the syllable, and, as there is nothing to prevent the spreading, continues throughout the word.

Thus emphasis spread is a process that can maximally affect the entire lexical item. Assuming that the lexical strata for the Optimality Theory framework consists of the minimal number of possible sub-strata, it is possible to combine the constraints for emphasis spread with those of lexical syllabification. In this case, merging the two causes no problem, as the constraints do not interact. Therefore, the combined ranking is:

$*\sigma_{\mu\mu\mu}$, MAX[+RTR] >>

MAXV, DEPV, MAXC, DEPC, ALIGN-[+RTR]-L(*syllable*), ALIGN-[+RTR]-R(*syllable*)>>

ONSET , ALIGN-[+RTR]-L (*Pwd*) >>

*COMPLEX, *CCC, *[+FRONT, +RTR] >>

ALIGN-[+RTR]-R (*Pwd*) >>

DEP [+RTR]

8. Lexical Stress

Lexical stress assignment must happen prior to syllable repair. The reasoning behind this lies in the fact that syllable repair works across the entire domain of the utterance. Consequently, to repair syllables using phonological processes to be further discussed creates a single domain, the utterance, by combining the lexical items. Therefore, lexical items are no longer relevant when considering the domain of the utterance/phrase, and lexical stress must be assigned while the domain is still that of the lexical item, not of the entire utterance/phrase.

Al-Jarrah (2008) proposed the following constraint ranking to account for lexical stress, as discussed in the previous chapter:

Lx = Pr >>

TROCHAIC >>

MAIN-RIGHT >>

PARSEσ>$\mu\mu$>>

NONFINAL >>

PARSEσ $\leq\mu\mu$. >>

FOOT-BINARY >>

ALL-FEET-LEFT (2008: 12)

However, PARSEσ in this account distinguishes between syllables with more than two moras and those with one or two moras. According to the view taken in the current project, syllables are maximally bimoraic (as demonstrated at the level of lexical syllabification), and the final

C in a CV:C or CVCC syllable is non-moraic. The constraints have been revised accordingly, defined as follows:

Lx=Pr	every lexical word must consist of a prosodic word. (Prince and Smolensky 1993: 51)
TROCHAIC	align head syllable with its foot, on the left edge. (Prince and Smolensky 1993: 56)
MAIN-RIGHT	align the head foot with prosodic word, on the right edge. (Prince and Smolensky 1993: 60)
PARSE	all syllables must be parsed into feet. (Prince and Smolensky 1993: 51).
NONFINAL	the final syllable should not be parsed into higher prosodic structure. (Prince and Smolensky 1993: 56)
FOOTBIN	feet are binary under a moraic analysis. (Prince and Smolensky 1993: 51)[7]
ALL-FT-L	align each foot with prosodic word on the left edge. (Prince and Smolensky 1993: 32)
ALL-FT-R	align each foot with prosodic word on the right edge. (Prince and Smolensky 1993: 32)
MAIN-LEFT	align head foot with prosodic word, on left edge. (Prince and Smolensky 1993: 60)

The constraints are ranked:

Lx = Pr >>

TROCHAIC>>

FOOTBIN>>

ALL-FT-L >>

MAIN-RIGHT>>

PARSE>>

NONFINAL>>

ALL-FEET-R >>

[7] In the original constraint by Prince and Smolensky, a violation is incurred for each foot that is not bimoraic. For example, a monomoraic foot receives one violation, as does a foot with four moras. For current purposes, the constraint is modified so that a foot receives violations as follows: one violation if the foot is monomoraic, and one violation *per* additional mora (thus a foot with four moras would receive two violations).

MAIN-LEFT

An example of the constraints for lexical stress follows:

/jˤɪˤ.bˤu:ˤzˤ/	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
jˤɪˤ.(ˈbˤu:ˤzˤ)					*	*	*		*
(jˤɪˤ. ˈbˤu:ˤzˤ)		*!	*				*		
(ˈjˤɪˤ.bˤu:ˤzˤ)			*!				*		
(ˈjˤɪˤ).bˤu:ˤzˤ			*!	*		*		*	

In the above tableau, the candidate j□I□.(□b□u:□z□) is the optimal candidate. The constraint TROCHAIC rules out any candidate with improper stress and foot relationships. The constraint FOOTBIN determines the optimal candidate. The optimal candidate contains one foot, which is contains precisely two moras, as the final consonantal segment in the coda, /z□/, is non-moraic.

/yix.ta.lɪ.fu/	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈyix.ta.lɪ.fu)			*!				*		
(yix.ta.)(ˈlɪ.fu)			*!		**		*	**	*
(yix.)(ˈta.lɪ.)fu				*	*	*		****	*
(yix. ˈta.)lɪ.fu		*!	*	*		**		**	
yix.(ˈta.lɪ.)fu				*	*	**!		*	*
(yix.ˈta)(lɪ.fu)		*!	*		**		*	**	
(yix.ta)(ˈlɪ.fu)			*!		**		*	**	*

In the above tableau, the optimal candidate is (yix.)(ˈta.lɪ.)fu. Here, stress occurs on the antepenultimate syllable. The optimal candidate violates the constraint ALL-FEET-L, as the second foot is not aligned to the left of the prosodic word. However, the only other candidate still possible at this point in the grammar also violates this constraint. Additionally, the optimal candidate violates the constraint MAIN-RIGHT, as the foot with the lexical stress is not aligned to the right of the prosodic word. Again, however, the only other candidate still possible also violates this constraint. The optimal candidate is determined via the constraint PARSE, as the candidate contains two feet, as opposed to only one.

As depicted in the above tableaus, the constraints in the determined hierarchy determine lexical stress placement in Cairene Arabic, via the methodology purported by McCarthy and Hayes (see Chapter 2).

9. Resyllabification and Syllable Repair

Following lexical stress assignment, it becomes necessary to work with a larger domain than the lexical item, as previously discussed. From this point onward, the domain in question consists of the phrase. With this new domain, syllabification must be reconsidered. Therefore, at this point, previous syllabification boundaries are ignored, and the domain is considered to be a string of segments without any sort of previous syllabification. However, string of segments is still considered to have stress, determined on the lexical level. In this case, stress is now a property of vocalic elements, as it is possible that consonantal elements are syllabified as the onset or coda of a syllable with a different nucleus, that does not have stress.

Switching from the lexical domain to the phrasal domain indicates a use of stratal Optimality Theory (Kiparsky 2010). In this case, the first stratum consists of the lexical domain, and all the processes which work on this domain, including syllabification, stress, and morphological substitutions. The second domain, which will be discussed from this point onward, is the phrase. Therefore, the output of the strata on the lexical domain serves as the input for the strata on the phrasal domain.

In this case, the resyllabification must also contain a repair mechanism, producing fully-formed syllables of Cairene Arabic. Syllables are determined on the basis of moracity. Moras are assigned to syllables as described. Every vowel receives a mora, as in the lexical syllabification. Syllables are again maximally of two moras (bimoraic). Word-final consonants are not moraic, while word-medial consonants that occur in a syllabic coda are moraic. The constraints to be used are as follows:

MAX-V	each V in the input must have a correspondent in the output.
DEP-V	each V in the output must have a correspondent in the input.
MAX-C	each C in the input must have a correspondent in the output.
DEP-C	each C in the output must have a correspondent in the input.
$*\sigma_{\mu\mu\mu}$	no syllable may contain more than two moras
ONSET	syllables must have an onset (C).
*COMPLEXONSET	an onset may not consist of more than one C.
*COMPLEXCODA	an coda may not consist of more than one C.
*CCC	no sequences of three consonants (in or across syllables).
*VVV	no sequences of three vowels (in or across syllables).[8]
*V: [-STRESS]	prohibits long unstressed vowels.
*[+high] [+low]	'*high-low' prohibits /ia/.[9] * [+back]

[8]This constraint prohibits sequences of three vowels. Sequences of two vowels are still allowed, in order to account for dipthongs.

*[+coronal] [+coronal] ‘l-assimilation’: assimilates /l/ to the following segment.
[+lateral] [-lateral]

*[+coronal] [-coronal] [+nasal] [-nasal]	‘n-assimilation’: assimilates coronal nasal to following segment’s place.
*[+labial] [+velar] [+nasal] [-nasal]	‘m-assimilation’: assimilates /m/ to following velar to become /ŋ/.[10]

Thus, the constraints in the level of resyllabification across the domain of the utterance are different than those across the domain of the lexical item, as syllable repair must also be accounted for.

In addition to these constraints, there are other phonological processes that take place within Cairene Arabic that are not so much about syllable repair, but instead concern ease of pronunciation. The first of these constraints prohibits the sequence /ia/. Instead, this sequence becomes /i:ja/. Additionally, Cairene Arabic has a process of l-assimilation, whereby /l/ assimilates to an adjacent coronal segment (regressive assimilation). Similar are processes of nasal assimilation: /n/ changes place of articulation depending on the adjacent segment (regressive assimilation), and /m/ changes place of articulation to [+velar] when followed by a [+velar] segment (regressive assimilation).

The constraints are ordered as follows:

*CCC, *VVV, ONSET, *V: [-STRESS] >>

MAX-C, DEP-C>>

$*\sigma_{\mu\mu\mu,}$>>

l-assimilation, n-assimilation, m-assimilation, *high-low >>

MAX-V, DEP-V

Examples of the constraints in action are provided below:

---

[9] In this case, this sequence becomes /i:ja/. While the usual candidate for onset insertion is the minimal consonant, the glottal stop, in this case the /j/ is inserted as the minimal consonant taking the place value from the preceding segment, /i/. Because the glottal stop is, in Cairene Arabic, always inserted in the utterance-initial position, there is no preceding consonantal segment from which to obtain a place value, and therefore the minimal consonant is used.

[10]The three assimilation constraints are merely shorthand notations for the interaction between markedness constraints and faithfulness constraints. Rather than spell out the entire phonological processes necessary for these to occur, it is simpler to prohihit a sequence with a markedness constraint that leads, artificially, to a specific phonological result.

al.ˈra.gil.yi. ˈsa.fir.fi:al. tˁɑ. ˈʔi.ra	*CCC	*VVV	ONSET	*V: [-STRESS]	DEP-C	MAX-C	$*\sigma_{\mu\mu\mu}$,	l-assim.	n-assim	m-assim	*high-low	MAX-V	DEP-V
al.ˈra.gil.yi. ˈsa.fir.fi:al. tˁɑ. ˈʔi.ra			*!*					**					
ʔal. ˈra.gil. yi. ˈsa.fir. fi:.ʔal. ˈtˁɑˁ.ʔi.ra					*! *			**					*
ʔal. ˈra.gil. yi. ˈsa.fir.fi:l. tˁɑˁ.ˈʔi.ra					*		*	*!*				*	
ʔar. ˈra.gil. yi.ˈsa.fir.fi:tˁ .tˁɑˁ.ˈʔi.ra					*		*					*	

In the tableau, the optimal candidate is /ʔar.ˈra.gil.yi.ˈsa.fir.fi:tˁ.tˁɑˁ.ˈʔi.ra/.The optimal candidate does not contain any onset-less syllables, nor any consonant or vowel clusters greater than two. Importantly, the optimal candidate assimilates the lateral /l/ to neighboring coronal consonants.

alˈbintkaˈbira	*CCC	*VVV	ONSET	*V: [-STRESS]	DEPC	MAXC	$*\sigma_{\mu\mu\mu}$,	l-assim.	n-assim	m-assim	*high-low	MAXV	DEPV
alˈbin. tka. ˈbi.ra	*!		*										
ʔal.ˈbin. ka.ˈbi.ra					*	*!							
ʔal. ˈbint. ka. ˈbi.ra	*!				*								
ʔal. ˈbin. ka. ˈbi.ra					*	*!							
ʔal. ˈbin. tik. ˈbi.ra					*							*	*

In the tableau above, the optimal candidate is /ʔal.ˈbin.tik.ˈbi.ra/. The optimal candidate does not contain any clusters of three consonants or vowels, nor does it contain any onset-less syllables. Importantly, the optimal candidate does not violate the faithfulness constraint MAX-C, although it does violate the three other faithfulness constraints.

10. Emphasis Spread Revisited

Emphasis spread was previously discussed in the sixth section of this chapter, but following resyllabification and syllable repair, a return to the phonological process is necessary. Because resyllabification can syllabify across word boundaries, it is possible that a syllable is created that has an emphatic segment as well as non-emphatic segments. Thus emphasis spread need be reconsidered.

Not all the previous constraints discussed in the sixth section need to be used in this section. At this point, emphasis spread as a phonological process can be accounted for with the following constraints:

ALIGN-[+RTR]-L (*syllable*)	The left edge of the feature [+RTR] must be aligned to the left edge of the syllable.
ALIGN-[+RTR]-R (*syllable*)	The right edge of the feature [+RTR] must be aligned to the right edge of the syllable.
MAX[+RTR]	Every secondary articulatory place [dor] feature in the input has a correspondent place [dor] feature in the output.
DEP[+RTR]	Every secondary articulatory place [dor] feature in the output has a corresponding place [dor] feature in the input.

The constraints concerning the prosodic word are no longer necessary at this stage: because the domain is the utterance, rather than the prosodic word, only the syllable need be considered as the unit of interest. Additionally, the constraint preventing segments with the feature [+FRONT] from also having the feature [+RTR] is no longer needed as there is no blocking of emphasis spread, as emphasis spread here will never go beyond the syllable. The alignment constraints for syllables account for this by ensuring that a syllable is either uniformly [+RTR] or [-RTR]. The faithfulness constraint ensures that no emphatic segments are changed to non-emphatic.

The constraints are ranked as follows:

MAX[+RTR] >>

ALIGN-[+RTR]-R (*syllable*), ALIGN-[+RTR]-L (*syllable*) >>

DEP[+RTR]

The faithfulness constraint that deletes the feature [+RTR] is inviolable. The alignment constraints are ranked on the same level because, at this stage, there is no blocking of emphasis spread in either direction. Instead it is only crucial that emphasis be a feature of the entire syllable.

ʔar.ra.gil.yi.'sa.fir.fi:tˤ.tˤɑ.'ʔi.ra	MAX [+RTR]	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*syllable*)	DEP [+RTR]
ʔar.ra.gil.yi.'sa.fir.fi:t.tˤɑ.'ʔi.ra	*!			
ʔar.ra.gil.yi.'sa.fir.fˤi: ˤtˤ.tˤɑ.'ʔi.ra		*!		**
ʔar.ra.gil.yi.'sa.fir.fi:tˤ.tˤɑˤ.'ʔi.ra			*!	*
ʔar.ra.gil.yi.'sa.fir.fˤi:ˤtˤ.tˤɑˤ.'ʔi.ra				***

The optimal candidate is /ʔar.ra.gil.yi.'sa.fir.fˤi:ˤtˤ.tˤɑˤ.'ʔi.ra /. In this candidate, the feature [+RTR] spreads from the segment /tˤ/ in the syllable /fi:t/ to give /fˤi: ˤtˤ/. The faithfulness constraint DEP[+RTR] is violated, but as it is lowest ranked, this does not prevent the candidate from winning.

Importantly, the constraints regarding emphasis spread can be easily added to the constraints involving resyllabification and syllable repair. In combination with the constraints from the previous section, the full ranking would be:

*CCC, *VVV, ONSET, *V: [-STRESS], MAX[+RTR] >>

ALIGN-[+RTR]-R (*syllable*), ALIGN-[+RTR]-L (*syllable*),>>

MAX-C, DEP-C>>

*$\sigma_{\mu\mu\mu}$,>>

l-assimilation, n-assimilation, m-assimilation, *high-low >>

MAX-V, DEP-V, DEP[+RTR]

Here, the constraint MAX[+RTR] is part of the highest-ranked group of constraints. Because MAX[+RTR] does not interact with any of the other constraints on this level, further ranking is unnecessary. The alignment constraints are ranked just below the constraint MAX[+RTR], and since these two constraints also do not interact with BIMORAICITY, they share a level in the hierarchy. The constraint DEP[+RTR] is among the lowest ranked, and was therefore grouped among the lowest-ranked constraints for convenience, as this constraint could have just as easily been ranked on the same level as*$\sigma_{\mu\mu\mu}$,.

11. Phrasal Stress

The domain is the utterance, and therefore it no longer makes sense to leave the stress on the lexical level. Therefore, this section deals with phrasal stress, in the post-lexical domain. At the level of the utterance, there is one primary stress, and one or more secondary stresses.

Because lexical stresses have already been assigned, the utterance, as it currently stands, has multiple primary stresses. It is thus logical to create a constraint IDENT(STRESS), that prohibits changing the value of the feature [STRESS] from input to output. Therefore, syllables that were stressed as a result of the lexical stress assignment must remain stressed, and no unstressed syllable may be assigned stress. However, an important distinction must be made between primary and secondary stress. As the utterance currently stands, all stresses have the same value, simply [+STRESS], with no distinction between primary and secondary.

According to Hellmuth 2004, phrasal stress occurs to the right of the prosodic phrase. This will be considered the primary stress. For this to happen, the rightmost lexical stress must be promoted to become the primary stress of the utterance, and the rest of the lexically stressed syllables have a secondary stress.

The following constraints are posited to account for phrasal stress:

MAINSTRESSRIGHT	Primary stress occurs on the rightmost foot.
MAX-HEIGHT	Prohibits changing height of vowels from input to output.
MAX-STRESS	Prohibits changing value of [STRESS] from input to output.
*CLASH	Prohibits adjacent stressed syllables.
*a C [-STRESS]	Prohibits unstressed /a/ before a consonant. (See following paragraphs for a description, on the basis of HNUC. Note that this constraint does not affect emphatic /a/ with the feature [+RTR])
PARSE	Each syllable must be parsed into a foot.

The constraints are thus ordered as follows:

PARSE >>

*a C >>
[-STRESS]

MAINSTRESSRIGHT >>

*CLASH >>

MAX(STRESS) >>

MAX(HEIGHT)

The constraint prohibiting unstressed /a/ can be considered an offshoot of the constraint HNUC proposed by Prince and Smolensky (1993: 17). This constraint states that a higher sonority nucleus (vocalic segment) is more harmonic than one of lower sonority. Based on the sonority hierarchy, the vocalic segment /i/ is not more sonorous than /a/. Because of its position in an unstressed syllable, a more sonorous vowel could be considered more difficult to produce. Therefore, the sonorous vowel /a/ is replaced with a less sonorous vowel, /i/.

The faithfulness constraint with respect to vowel height needs to be ranked lowest, because in order to change the vowel /a/ to /i/, this results in a change of vowel height. Therefore, the markedness constraints must be ranked higher, to prevent unstressed /a/ from occurring, as well as preventing any unstressed long vowels. Unstressed /a/ is only allowed to occur word-finally, as indicated by the specification in the constraint that the unstressed /a/ must be followed by a consonantal segment in order to be eligible for the constraint. *CLASH is ranked directly under MAINSTRESSRIGHT, to prevent two stressed syllables to occur in adjacent position. These two constraints must be ranked above the faithfulness constraint concerning stress, in order to allow stress to change with respect to the optimal output. The faithfulness constraint MAX(STRESS) only prevents the deletion of the feature [+STRESS]. The constraint does not prevent changing a primary lexical stress to a secondary stress at the level of the utterance.

Importantly, foot boundaries are still needed in the assignment of phrasal stress. Because phrasal stress must occur on the rightmost foot, the violations of this constraint are taken as ‘any foot with a main stress that is not the rightmost foot receives a violation.’ However, because feet have already been assigned, the constraint FOOTBIN is no longer relevant. Instead, what is important is that no syllable is unfooted.

An example of the ranked constraints is as follows:

ʔal.'bin.tid.'ra.si.tal.ki.'ta:b	*a C [-stress]	MAIN STRESS RIGHT	*CLASH	MAX (STRESS)	MAX (HEIGHT)
(ʔal.)('bin.tid.)('ra.si.)(tal.ki.)('ta:b)	*!*	**			
(ʔil.)('bin.tid.)('ra.si.)(tal.ki.('ta:b)	*!	**			*
(ʔil.)('bin.tid.)('ra.si.)(til.ki.)('ta:b)		*!*			**
(ʔil.)(bin.tid.)(ra.si.)(til.ki.)('ta:b)				*!*	
(ʔil.)('bin.tid.)('ra.si.)(til.ki.)('ta:b)					**

In the above tableau, the optimal candidate is (ʔil.)('bin.tid.) ('ra.si.)(til.ki.)('ta:b). In this candidate, all occurrences of unstressed /a/ have become /i/. The mainstress occurs on the right-most foot, /ta:b/. No stresses have been added or deleted, and secondary stress is also accounted for. The optimal candidate violates the faithfulness constraint concerning the height of the vowels, but as this constraint is ranked low, violations here do not determine the optimal candidate.

12. Examples

In this section, full examples of the transformation from Modern Standard Arabic to Cairene Arabic will be given.

12.1 Example 1

Gloss			My friend (*masc.*) arrives on time.
MSA Orthography			صديقي يصل في الميعاد
Orthography → IPA			sˤadi:qi: yasˤal fi: almi:ʕa:d
MSA IPA → CA IPA			sˤadi:qi: yasˤal fi: almi:ʕa:d
Morphology	Full Pronouns		none
	Bound Pronouns		sˤadi:qi: →sˤadi:qi
	Verb	Perfective	---
		Imperfective	yasˤal → yisˤal
		Future	---
		Negation	---

Lexical Syllabification

sˤadi:qi	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
sˤa.di:.qi								
sˤad.qi			*!					

yisˤal	*$\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
yisˤ.al						*!		
yi.sˤal								*
yisˤ.sˤal					*!			

fi:	*$\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
fi:								
fi		*!						

almi:ʕa:d	*$\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
al.mi:.ʕa:d						*		
ʔal.mi:.ʕa:d					*!			

The output is [sˤa.di:.qi] [yi.sˤal] [fi:] [al.mi:. ʕa:d]

Emphasis Spread

sˤa.di:.qi	MAX [+RTR]	ALIGN-[+RTR]-L (*syllable*)	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN-[+RTR]-R (*Pwd*)	DEP [+RTR]
sˤa.di:.qi			*!				
sa.di:.qi	*!						
sˤaˤ.dˤi:.qi			*!			*	**
sˤaˤ.dˤi:ˤ.qi					*!	*	***
sˤaˤ.dˤi:ˤ.qiˤ					*!		****
sˤaˤ.di:.qi						*	*

yi.sˤal	MAX [+RTR]	ALIGN-[+RTR]-L (*syllable*)	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN-[+RTR]-R (*Pwd*)	DEP - [+RTR]
yi.sˤal			*!*	*			
yi.sal	*!						
yi.sˤaˤlˤ				*!			**
yˤiˤ.sˤaˤlˤ					*		****
yˤiˤ.sˤal			*!*		*		**

fi:	MAX [+RTR]	ALIGN-[+RTR]-L (*syllable*)	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN-[+RTR]-R (*Pwd*)	DEP - [+RTR]
fi:							
fˤi: ˤ					*!		**

al.mi:.ʕa:d	MAX [+RTR]	ALIGN-[+RTR]-L (*syllable*)	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN-[+RTR]-R (*Pwd*)	DEP - [+RTR]
al.mi:.ʕa: ˤdˤ		*!		**			**
al.mi:.ʕa: ˤdˤ				*!*			***
al.mi:. ʕa:d							

The output is [sˤaˤ.di:.qi] [yˤiˤ.sˤaˤlˤ] [fi:][al.mi:.ʕa:d].

Lexical Stress

sˤaˤ.di:.qi	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈsˤaˤ.di:.qi)			*!*				*		
ˈsˤaˤ.(di:.qi)			*!	*	*	*	*		*
(ˈsˤaˤ.di:.)qi			*!	*		*		*	
(sˤaˤ.ˈdi:.)qi		*!		*		*		*	
sˤaˤ.(ˈdi:.qi)			*!		*	*	*		*
sˤaˤ.(ˈdi:.)qi				*	*	**		*	*

yˤiˤ.sˤaˤlˤ	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈyˤiˤ.sˤaˤlˤ)							*		
(yˤiˤ.ˈsˤaˤlˤ)		*!					*		
yˤiˤ.(ˈsˤaˤlˤ)			*!		*	*	*		*
(ˈyˤiˤ.)sˤaˤlˤ			*!	*		*		*	

fi:	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈfi:)							*		
ˈfi:	*!					*			

al.mi:.ʕa:d	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈal.mi:.ʕa:d)			*!****				*		
(ˈal.mi:.) ʕa:d			*!	*				*	
(al. ˈmi:.) ʕa:d		*!	*	*				*	
al.(mi:.ˈʕa:d)		*!	**				*		*
al.(ˈmi:.ʕa:d)			*!*		*		*		*
al.mi:.(ˈʕa:d)					**	**	*		**

The output is [sˤaˤ.(ˈdi:.)qi] [(ˈyˤiˤ.sˤaˤlˤ)] [(ˈfi:)] [al.mi:.(ˈʕa:d)].

Resyllabification and Syllable Repair

ˈsˤaˤdi:qiˈyˤi ˤsˤaˤlˤˈfi:alˈm i:ʕa:d	*CCC	*VVV	ONSET	*V: [-STRESS]	DEP-C	MAX-C	*$\sigma_{\mu\mu\mu}$,	l-assim.	n-assim	m-assim	*high-low	MAX-V	DEP-V
sˤaˤ.ˈdi:.qiˈ.yˤiˤ .sˤaˤlˤ.ˈfi:.al. ˈmi:.ʕa:d		*	*	*!									
sˤaˤ.ˈdi:.qiˈ.yˤiˤ .sˤaˤlˤ.ˈfi:.ʔal. ˈmi:.ʕa:d				*!	*	*							
sˤaˤ.ˈdi:.qiˈ.yˤiˤ .sˤaˤlˤ.ˈfi:l. ˈmi:.ʕa:d				*!								*	
sˤaˤ.ˈdi:.qiˈ.yˤi ˤ.sˤaˤlˤ.ˈfi:l. ˈmi:.ʕad												*	

The output is [sˤaˤ.ˈdiː.qiˈ.yˤiˤ.sˤaˤlˤ.ˈfiːl. ˈmiː.ʕad].

Emphasis Spread Revisited

sˤaˤ.ˈdiː.qiˈ.yˤiˤ.sˤaˤlˤ.ˈfiːl. ˈmiː.ʕad	MAX [+RTR]	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*syllable*)	DEP[+RTR]
sˤaˤ.ˈdiː.qi. ˈyi.sal. ˈfiːl. ˈmiː.ʕad				*!****
sˤaˤ.ˈdiː.qi. ˈyˤiˤ.sˤaˤlˤ.ˈfiːl. ˈmiː.ʕad				

The output is [sˤaˤ.ˈdiː.qi. ˈyˤiˤ.sˤaˤlˤ.ˈfiːl. ˈmiː.ʕad].

Phrasal Stress

sˤaˤ.ˈdiː.qi. ˈyˤiˤ.sˤaˤlˤ.ˈfiːl. ˈmiː.ʕad	*a C [-stress]	MAIN STRESS RIGHT	*CLASH	MAX (STRESS)	MAX (HEIGHT)
sˤaˤ.ˈdiː.qi. ˈyˤiˤ.sˤaˤlˤ.ˈfiːl.ˈmiː.ʕad	*!	**	*		
sˤaˤ.diː.qi. ˈyˤiˤ.sˤaˤlˤˌfiːl ˈmiː.ʕad	*!				
sˤaˤ.ˌdiː.qi. ˈyˤiˤ.sˤaˤlˤˈfiːl. ˈmiː.ʕid			*!		*
sˤaˤ.ˌdiː.qi.▯yˤiˤ.sˤaˤlˤˌfiːl. ˈmiː.ʕid		*!	*		*
sˤaˤ.ˌdiː.qi.ˌyˤiˤ.sˤaˤlˤ.fiːl.ˈmiː.ʕid					*

The ultimate output is thus [sˤaˤ.ˌdiː.qi.ˌyˤiˤ.sˤaˤlˤ.fiːl.ˈmiː.ʕid].

12.2 Example 2

Gloss The two of you went to the restaurant.

MSA Orthography أنتما ذهبتم الى المتم

Orthography → IPA ʔantumaa ðahabtuma ila: mataʕam

MSA IPA → CA IPA ʔantumaa zahabtuma ila: mataʕam

Morphology Full Pronouns ʔantumaa → ʔintum

Bound Pronouns none

Verb Perfective zahabtuma→ zahabtu

Imperfective

Future

Negation

Lexical Syllabification

ʔintum	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
ʔin.tum								
ʔint.um							*!	

zahabtu	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
za.hab.tu								
zah.btu							*!	*

ila:	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
il.a:						**!		
i.la:						*		
la:		*!						

mataʕam	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
ma.ta.ʕam								
ma:.ʕam				*!				

The output is [ʔin.tum] [za.hab.tu] [i.la:] [ma.ta.ʕam].

Emphasis Spread

There are no segments with the feature [+RTR]

Lexical Stress

ʔin.tum	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ʔin.ˈtum)		*!	*				*		
(ˈʔin.tum)			*				*		
ʔin (ˈtum)			*		*!	*	*		

za.hab.tu	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
za.(ˈhab.tu)			*!		*	*	*		*
(ˈza.hab.)tu			*!	*		*		*	
za. (ˈhab.)tu				*	*	**		*	*
(ˈza.hab.tu)			*!*				*		

i.la:	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈi.la:)			*!				*		
i.(ˈla:)					*	*	*		*
(i. ˈla:)		*!					*		

ma.ta.ʕam	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈma.ta.ʕam)			*!				*		
(ˈma.ta.)ʕam				*!		*		*	
ma.(ta. ˈʕam)		*!			*	*	*		
(ma. ˈta.)ʕam		*!				*		*	
ma.(ˈta.ʕam)					*	*	*		

The output is [(ˈʔin.tum)] [za.(ˈhab.)tu] [i. (ˈla:)] [ma.(ˈta.ʕam)].

Resyllabification and Syllable Repair

ˈʔintumzaˈhabtu iˈla:maˈtaʕam	*CCC	*VVV	ONSET	*V: [-STRESS]	DEP-C	MAXC	$*\sigma_{\mu\mu\mu}$,	l-assim.	n-assim	m-assim	*high-low	MAXV	DEPV
ˈʔin.tum.za. ˈhab.tui. ˈla:maˈta. ʕam			*!	*									
ˈʔin.tum.zaˈha b.ti. ˈla.maˈta. ʕam												*	
ˈʔin.tum.zaˈha b.tu. ˈla.maˈta. ʕam												*	

In this case, there are two optimal candidates. However, because Cairene Arabic favors /i/ over other vowels, the candidate [ˈʔin.tum.zaˈhab.ti.ˈla:maˈta.ʕam] is the optimal candidate.

Emphasis Spread Revisited

There are no segments with the feature [+RTR]

Phrasal Stress

ʔˈin.tum.zaˈhab.ti. ˈla.maˈta.ʕam	*a C [-stress]	MAIN STRESS RIGHT	*CLASH	MAX (STRESS)	MAX (HEIGHT)
ʔˈin.tum.zaˈhab.ti. ˈla.maˈta.ʕam		*!**			****
ˌʔin.tum.zi.ˌhab. ˌtil.mi. ˈta. ʕim			*!		****
ˌʔin.tum.zi.ˌhab.ti. ˌla.mi.ˈta.ʕim				*	****

The ultimate output is thus [ˌʔin.tum.zi.ˌhab.ti. ˌla.mi.ˈta.ʕim].

12.3 Example 3

Gloss — We will not eat breakfast tomorrow.

MSA Orthography — نحن لن نتناولون الافطار غدا

Orthography → IPA — naħnu lan natana:walun alift^{ʕ}a:r ɣadan.

MSA IPA → CA IPA — naħnu lan natana:walun alift^{ʕ}a:r ɣadan.

Morphology

- Full Pronouns — naħnu →ʔiħna
- Bound Pronouns — none
- Verb
  - Perfective
  - Imperfective — natana:walun→ nitana:walun
  - Future — lan→ħa
  - Negation — lannatana:walun→manitana:walunʃi
  - Output — ħamanitana:walunʃi

Lexical Syllabification

ʔiħna	$*\sigma_{\mu\mu\mu}$	Max-V	Dep-V	Max-C	Dep-C	Onset	*Complex	*CCC
ʔiħ.na								
ʔi.ħna							*!	

ħamanitana:walunʃi	$*\sigma_{\mu\mu\mu}$	Max-V	Dep-V	Max-C	Dep-C	Onset	*Complex	*CCC
ħa.ma.ni.ta.na:.wa.lun.ʃi								
ħam.nit.na:.wa.lun.ʃi		*!*						

alift^{ʕ}a:r	$*\sigma_{\mu\mu\mu}$	Max-V	Dep-V	Max-C	Dep-C	Onset	*Complex	*CCC
a.lif.t^{ʕ}a:r						*		
al.if.t^{ʕ}a:r						**!		

ɣadan.	$*\sigma_{\mu\mu\mu}$	MAX-V	DEP-V	MAX-C	DEP-C	ONSET	*COMPLEX	*CCC
ɣa.dan.								
ɣad.an.						*!		

The output is [ʔiħ.na] [ħa.ma.ni.ta.na:.wa.lun.ʃi] [a.lif.tˤa:r] [ɣa.dan.].

Emphasis Spread

a.lif.tˤa:r	MAX [+RTR]	ALIGN- [+RTR]-L (*syllable*)	ALIGN- [+RTR]-R (*syllable*)	ALIGN- [+RTR]-L (*Pwd*)	*[+FRONT] [+RTR]	ALIGN- [+RTR]- R (*Pwd*)	DEP - [+RTR]
a.lif.tˤa:r			*!	**		*	
a.lif.tˤa:ˤrˤ				*!*			*
a.lif.ta:r	*!						
aˤ.lˤiˤfˤ.tˤaˤ:rˤ					*		*****

The output is [ʔiħ.na] [ħa.ma.ni.ta.na:.wa.lun.ʃi] [aˤ.lˤiˤfˤ.tˤaˤ:r] [ɣa.dan.].

Lexical Stress

ʔiħ.na	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN- RIGHT	ALL- FEET-L	PARSE	NON FINAL	ALL- FEET-R	MAIN- LEFT
(ˈʔiħ.na)			*!				*		
(ʔiħ.ˈna)		*!	*				*		
(ˈʔiħ.)na				*		*		*	

ħa.ma.ni.ta.na: .wa.lun.ʃi	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN- RIGHT	ALL- FEET-L	PARSE	NON FINAL	ALL- FEET-R	MAIN- LEFT
(ˈħa.ma.)(ˈni.ta.) (ˈna:.wa.)(ˈlun.ʃi)			**	*!**	***		*	***	***
(ˈħa.ma.)(ˈni.ta .)(ˈna:.wa.)(ˈlu n.ʃi)			**		***		*	***	***
(ˈħa.ma.)(ˈni.ta.) (ˈna:.wa.)(ˈlun.ʃ i)			**	*!	***		*	***	**

aˤ.lˤiˤfˤ.tˤaˤ:rˤ	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈaˤ.lˤiˤfˤ.)tˤaˤ:rˤ)			*!	*		*		*	
aˤ.(ˈlˤiˤfˤ.tˤaˤ:rˤ)			*!		*	*	*		*
aˤ.(lˤiˤfˤ.ˈtˤaˤ:rˤ)		*!			*	*	*		*
aˤ.(lˤiˤfˤ.)(ˈtˤaˤ:rˤ)					***	*	*	**	**

ɣa.dan	Lx = Pr	TRO CHAIC	FOOT BIN	MAIN-RIGHT	ALL-FEET-L	PARSE	NON FINAL	ALL-FEET-R	MAIN-LEFT
(ˈɣa.dan)		*!					*		
(ɣa. ˈdan)							*		
ɣa.(ˈdan)			*!		*	*	*		*

The output is [(ˈʔiħ.)na] [(ˈħa.ma.)(ˈni.ta.)(ˈna:.wa.)(ˈlun.ʃi)] [aˤ.(lˤiˤfˤ.)(ˈtˤaˤ:r)] [(ɣa. ˈdan)].

Resyllabification and Syllable Repair

ˈʔiħnaˈħamaˈnitaˈna:waˈlunˈiaˤlˤiˤfˤtˤaˤ:rˤaɣdan	*CCC	*VVV	ONSET	*V: [-STRESS]	DEP-C	MAX-C	*σμμμ	l-assim.	n-assim	m-assim	*high-low	MAX-V	DEP-V
ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈna:.wa. ˈlun.ʃi.aˤ.ˈlˤiˤf ˤ.tˤaˤ:.ɣa. ˈdan			*!	*							*		
ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈna:.wa. ˈlun.ʃi.aˤ.ˈlˤiˤf ˤ.tˤaˤ.ɣa. ˈdan			*!								*	*	
ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈna:.wa. ˈlun.ʃi:.yaˤ.ˈlˤ iˤfˤ.tˤaˤ.ɣa. ˈdan					*								

The output is [ˈʔiħ.na.ˈħa.ma.ˈni.ta.ˈna:.wa.ˈlun.ʃi:.yaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa.ˈdan]

Emphasis Spread Revisited

ˈʔiħ.na.ˈħa.ma.ˈni.ta.ˈnaː.wa.ˈlun.ʃiː. yaˤ.ˈlˤiˤfˤ.tˤaˤrˤ.ɣa.ˈdan	MAX[+RTR]	ALIGN-[+RTR]-R (*syllable*)	ALIGN-[+RTR]-L (*syllable*)	DEP[+RTR]
ˈʔiħ.na.ˈħa.ma.ˈni.ta.ˈnaː.wa.ˈlun.ʃiː. yaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa.ˈdan			*!	
ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈnaː.wa. ˈlun.ʃiː.yˤaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa. ˈdan				*
ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈnaː.wa. ˈlun.ʃiː.ya. ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa. ˈdan	*!			

The output is [ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈnaː.wa. ˈlun.ʃiː.yˤaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa. ˈdan]

Phrasal Stress

ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈnaː.wa. ˈlun.ʃiː.yˤaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa. ˈdan	*a C[-stress]	MAIN STRESS RIGHT	*CLASH	MAX (STRESS)	MAX (HEIGHT)
ˈʔiħ.na. ˈħa.ma. ˈni.ta. ˈnaː.wa. ˈlun.ʃiː.yˤaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣa. ˈdan	*!****	***			
ˈʔiħ.ni. ˈħa.mi. ˈni.ti. ˈnaː.wi. ˈlun.ʃiː.yˤaˤ.ˈlˤiˤfˤ.tˤaˤ rˤ.ɣi. ˈdan		*!**			****
ˌʔiħ.ni.ˌħa.mi.ˌni.ti.ˌnaː.wi.ˌlun. ʃiː.yˤaˤ.ˌlˤiˤfˤ.tˤaˤ rˤ.ɣi. ˈdan					****

The ultimate output is thus [ˌʔiħ.ni.ˌħa.mi.ˌni.ti.ˌnaː.wi.ˌlun.ʃiː.yˤaˤ.ˌlˤiˤfˤ.tˤaˤ rˤ.ɣi. ˈdan].

13. Implementation

Because the ultimate goal is synthesized Cairene Arabic speech, it is necessary that all the phonological processes are able to be completed by an automatic process, to facilitate the transformation of Modern Standard Arabic orthography to Cairene Arabic speech. Therefore, an algorithm that can carry out context-free rewrite rules, context-dependent rewrite rules, and constraint-based Optimality Theory must be developed in order to implement the transformation to Cairene Arabic written in IPA.

Ideally, it would be possible to have one algorithm that would work through the entire process, from context-free to context-dependent rewrite rules, and two strata of Optimality Theory (lexical and phrasal). However, that would require the ability of a programmer, as well as more time than the current project permits.

There is a program already in existence, PRAAT, that can be used to demonstrate constraint-based Optimality Theory in an automatized setting.. Although PRAAT would not be the end mechanism for the entire algorithm if it were implemented, PRAAT does provide some groundwork for the programming that would need to be done. In this section, a demonstration of how such an algorithm could work (in Optimality Theory) is detailed.

Only the phonological process of syllable repair will be implemented in PRAAT, for demonstration purposes. For the project to work, it would be necessary to implement all phonological processes in PRAAT. Using PRAAT, the constraints can be ranked in relation to one another, creating either a strict ordering of ranked constraints, or a stochastic ordering of ranked constraints.

This section will demonstrate how the constraints used for resyllabification and syllable repair across the phrasal domain can be implemented in PRAAT. The constraints to be used are as follows (with ranking):

> *CCC, *VVV, ONSET, *V: [-STRESS] >>
>
> MAX-C >>
>
> l-assimilation, n-assimilation, m-assimilation, *high-low >>
>
> MAX-V, DEP-V, DEP-C

This is an instance of stratal Optimality Theory: constraints are ranked in relation to each other, but some constraints may have the same ranking. The text file for the ranking, in PRAAT, appears as follows:

```
"ooTextFile"
"OTGrammar"
12 constraints
constraint [1]: "*CCC" 100 100 !  *CCC
constraint [2]: "*VVV" 100 100 ! *VVV
constraint [3]: "O\s{nset}" 100 100 ! ONSET
constraint [4]: "*V:[-stress]" 100 100 ! *V:
constraint [5]: "M\s{AX}C" 90 90 ! MAX-C
```

```
constraint [6]: "L\s{assim}" 80 80 ! L-assim
constraint [7]: "N\s{assim}" 80 80 ! N-assim
constraint [8]: "M\s{assim}" 80 80 ! M-assim
constraint [9]: "*H\s{igh}L\s{ow}" 80 80 ! *HIGH-LOW
constraint [10]: "M\s{AX}V" 70 70 ! MAX-V
constraint [11]: "D\s{EP}V" 70 70 ! DEP-V
constraint [12]: "D\s{EP}C" 70 70 ! DEP-C
     0 fixed rankings
     2 tableaus
"albintkabira" 5
     "al.bin.tka.bi.ra"   1 0 1 0 3 0 0 0 0 0 0 0 0
     "?al.bin.ka.bi.ra"   0 0 0 0 3 1 0 0 0 0 0 0 1
     "?al.bint.ka.bi.ra"  1 0 0 0 0 0 0 0 0 0 0 0 1
     "?al.bin.ka.bi.ra"   0 0 0 0 3 1 0 0 0 0 0 0 1
     "?al.bin.tik.bi.ra"  0 0 0 0 2 0 0 0 0 0 1 1 1
"alragilyisafirfi:alt`a`ira" 4
     "al.ra.gil.yi.sa.fir.fi:al.t`a`.?i.ra"
          0 0 2 0 7 0 2 0 0 0 0 0 0
     "?al.ra.gil.yi.sa.fir.fi:.?al.t`a`.?i.ra"
          0 0 0 0 6 0 2 0 0 0 0 1 2
     "?al.ra.gil.yi.sa.fir.fi:l.t`a`.?i.ra"
          0 0 0 0 6 0 2 0 0 0 1 0 0
     "?ar.ra.gil.yi.sa.fir.fi:t`.t`a`.?`i.ra"
          0 0 0 0 6 0 0 0 0 0 1 0 0
```

When run via PRAAT, this produces the following:

	ranking value	disharmony	plasticity
*CCC	100.000	103.641	1.000000
*VVV	100.000	101.041	1.000000
*V:[-stress]	100.000	100.028	1.000000
Onset	100.000	99.254	1.000000
MAXC	90.000	87.949	1.000000
*HighLow	80.000	83.373	1.000000
Lassim	80.000	70.780	1.000000
Massim	80.000	68.543	1.000000
Nassim	80.000	67.725	1.000000
DEPC	70.000	67.428	1.000000
DEPV	70.000	59.207	1.000000
MAXV	70.000	58.486	1.000000

albintkabira	*CCC	*VVV	*V:[-stress]	Onset	MAXC	*HighLow	Lassim	Massim	Nassim	DEPC	DEPV	MAXV
al.bɪn.tka.bi.ra	*!			*								
ʔal.bɪn.ka.bi.ra					*					*		
ʔal.bɪnt.ka.bi.ra	*!									*		
ʔal.bɪn.ka.bi.ra					*					*		
☞ ʔal.bɪn.tik.bi.ra										*	*	*

alragilyisafirfi:altˤɑʔira	*CCC	*VVV	*V:[-stress]	Onset	MAXC	*HighLow	Lassim	Massim	Nassim	DEPC	DEPV	MAXV
al.ra.gil.yi.sa.fir.fi:al.tˤɑˤ.ʔɪ.ra				*|*			**					
ʔal.ra.gil.yi.sa.fir.fi:.ʔal.tˤɑˤ.ʔɪ.ra							*|*			**	*	
ʔal.ra.gil.yi.sa.fir.fil.tˤɑˤ.ʔɪ.ra							*|*					*
☞ ʔar.ra.gil.yi.sa.fir.fitˤ.tˤɑˤ.ʔɪ.ra												*

PRAAT produces a clear ranking of the constraints, pre-specified by the ranking value. The top-ranked constraints have a ranking value of 100, and this decreases by a value of 10 for each strata below the top-ranked strata.

Therefore, PRAAT can be used to successfully implement the grammar imposed by Optimality Theory. By implementation of a number of input forms to the ranked constraints, the grammar can eventually be learned by the computer. Eventually, then, the constraints can be used to determine the optimal output forms from any input, which is the entire purpose of the algorithm.

14. Conclusions

In this chapter, the transformation from Modern Standard Arabic orthography to the input to what will be the speech synthesizer. In this case, there were two domains: the lexical domain and the phrasal domain.

The transformation begun with Modern Standard Arabic orthography as input, which was transformed into the symbols of the IPA, first of Modern Standard Arabic, then of Cairene Arabic. This was followed by a morphological transformation, of both verbal inflections and pronouns.

After these transformations via rewrite rules, Optimality Theory was used as the method of transformation. These transformations involved lexical syllabification, emphasis spread, and lexical stress assignment. On the phrasal domain, these transformations resyllabified the input and repaired syllables, emphasis spread, and stress assignment on the phrasal level.

Examples of the transformation demonstrated the algorithm in action, from Modern Standard Arabic orthography to the symbols of the IPA complete with all necessary phonological processes. It was then shown that it is possible to implement the algorithm by a computer. The open source program PRAAT was used to depict resyllabification and syllable repair. This demonstrated that it is possible to automaticize the algorithm.

At this point, phonological processes will no longer be discussed. The following chapters provide an overview of how a speech synthesizer could be implemented using the open source program *Festival*.

Interlude: Phonology to Synthesized Speech

Overview

The ultimate goal of this project is to provide the framework for a speech synthesizer for Cairene Arabic, using Modern Standard Arabic text as input. Chapters 2 and 3 dealt with the phonology, transforming Modern Standard Arabic orthography into Cairene Arabic, in the orthography prescribed by the International Phonetic Alphabet.

The output of the these two chapters is thus the input for the following chapters. The project thus has two distinct parts: the first portion provided an overview of Modern Standard Arabic, its orthography, morphology, and phonology. This was extended into an overview of Cairene Arabic, its morphology and phonology. The two were then compared, which was then followed by the process of how to bridge the two: phonological rules and constraints were given to transform Modern Standard Arabic orthography into Cairene Arabic IPA.

This IPA Cairene Arabic is thus the input to the text-to-speech system to be described in the following chapters. The third chapter provides an overview of the text-to-speech technology. An overview of different types of speech synthesis is given, as well as the reason for choosing a particular type of speech system. This is followed by a survey of the work that has been done concerning Modern Standard Arabic and dialectical Arabic speech synthesis. The fourth chapter lays the framework for a diphone speech synthesis of Cairene Arabic.

Because of the use of IPA as the input for the text-to-speech synthesizer, multiple problems are avoided. These problems are discussed extensively in the following chapter. The majority of these problems concern the implementation of an automatic diacriticization generator for the Arabic orthography, normally written without diacritics. The use of the International Phonetic Alphabet as the input avoids the issue of diacritic generation, as the input contains both vocalic and consonantal segments. Additionally, the use of the International Phonetic Alphabet makes letter-to-sound mapping much easier, as each of the symbols of the International Phonetic Alphabet can be described with its own feature matrix.

The International Phonetic Alphabet is used for implementation in a text-to-speech synthesizer under the framework of the speech synthesis program *Festival*. This is detailed in the fourth chapter. In order for this to be properly implemented, the Speech Assessment Methods Phonetic Alphabet (SAMPA) is used in place of the International Phonetic Alphabet, as SAMPA is computer-readable, as opposed to IPA. The full framework for this speech synthesizer is provided in the fourth chapter.

The final output of the project is, hypothetically, synthesized Cairene Arabic speech.

4: Arabic Text-to-Speech Synthesis

1. Purpose

This chapter provides an overview of text-to-speech synthesis, and more specifically the text-to-speech synthesis that uses the diphone as the unit of concatenative synthesis. The chapter begins with a description of diphone synthesis, followed by argument in favor of diphone synthesis as the unit for concatenative synthesis.

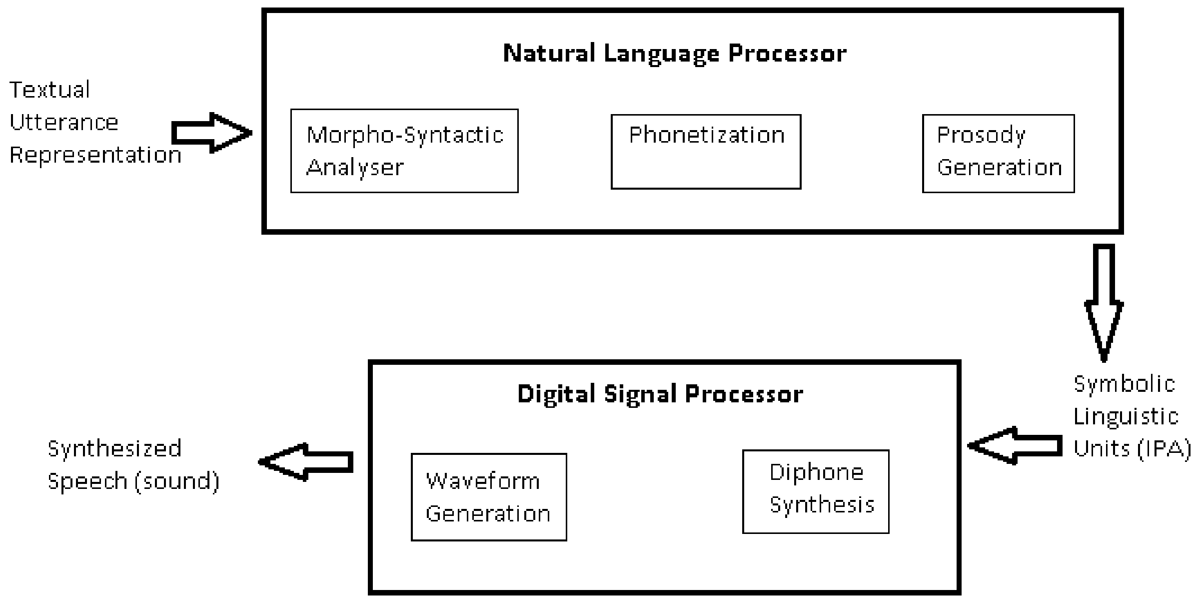
This is followed by a discussion of the problems often encountered within speech synthesis of Arabic, with reference to how these problems will be solved for the current purposes. These difficulties include the lack of diacritical marks in Modern Standard Arabic texts, as well as the different problems encountered when attempting speech synthesis of a colloquial dialect of Arabic, as opposed to the formalized standard language. Previous efforts in both Modern Standard and Cairene Arabic are expounded upon, using a variety of methods with diphones as the main unit of concatenation.

The goal of this chapter is to provide background in the field of speech synthesis, to be used as a foundation on which a theoretical speech synthesizer will be built, in the following chapter.

2. Speech Synthesis

Text-to-speech synthesis has a wide variety of applications in the natural world. Speech synthesizers using text as input first appeared in the middle of the 1980s. Since then, speech synthesis technology has expanded for use in telecommunication services, foreign language education, aid to the disabled, children's books and toys, and scientific research (Dutoit 1995: 2-4).

A general text-to-speech synthesizer consists of a Natural Language Processor and a Digital Signal Processing module. The Natural Language Processor, in the most general sense, is a text-analyzer that produces a phonetic transcription of input text, in format recognizable by a computer. In addition, the Natural Language Processer produces the desired prosody for the text, along with a letter-to-sound module (Dutoit 1995: 62). In turn the Digital Signal Processing module transforms this transcription with prosodic information into recognizable speech. (Dutoit 1995: 35) In this case, this is done by use of the diphones from audio files, concatenated to form recognizable synthesized speech. The schematic below provides a functional diagram of the speech synthesizer, from input (textual utterance representation) to output (synthesized speech).



The Natural Language Processor in the above diagram contains three parts: morpho-syntactic analysis, phonetization, and generation of prosody. These can either be sequential, where the output of the morpho-syntactic analyzer feeds into the phonetization, which feeds into prosody generation, or parallel, where the morpho-syntactic analyzer, phonetization and generation of prosody simultaneously work on the same string. The output of the Natural Language Processor is symbolic linguistic units, which are in this study fully-specified IPA. The Digital Signal Processing module contains two parts: diphone synthesis and waveform generation. The diphone synthesis assigns diphone files to the string, while the waveform is generated as a result of combining these diphones. As with the Natural Language Processor, diphone synthesis and waveform generation can occur in sequence, first diphone synthesis and then waveform generation, or in parallel, with the diphone assignment occurring simultaneously with waveform generation.

A high-quality text-to-speech synthesizer contains a morpho-syntactic analyzer within the Natural Language Processor. This analyzer is important for its contribution to the prosodic structure: in language, prosody is dependent on the syntactic structure. (Dutoit 1995: 40) This tends to present a large problem for text-to-speech synthesizers, as an automatic analyzer of the morphology and syntax to determine the prosody is not necessarily perfect. Here, the morpho-syntactic analyzer will not be a subject of interest, as the input to the proposed synthesizer was morphologically and syntactically analyzed in order to determine relevant phonological processes in Chapter 2.

The Natural Language Processor usually contains a letter-to-sound module, under the heading 'Phonetization,' which is responsible for 'the automatic determination of the phonetic transcription of the incoming text' (Dutoit 106). When using the orthography of a language as the input to the text-to-speech synthesizer, this is possibly problematic, as most, if not all, of the world's language do not have a one-to-one mapping between orthography and sound. In order to account for this, many speech synthesizers employ both dictionary-based and rule-based strategies (Dutoit 1995: 111). Rule-based strategies consist of a series of grapheme-to-phoneme rules, and often come with a list of exceptions, in a dictionary-like format. Dictionary-based strategies store maximum phonological knowledge, usually in the form of morphemes, as well as a list of inflections and derivations from which lexical items are formed (Dutoit 1995: 111).

However, for current purposes, the letter-to-sound module avoids many of the frequently encountered problems, most notably that of one character corresponding to more than one phoneme (Dutoit 1995: 106). Because the input to the Natural Language Processor is, in the current project, the International Phonetic Alphabet, each 'letter' corresponds to precisely one phoneme. This makes the task of the Natural Language Processor much less taxing than having to determine letter-to-sound rules on the basis of spelling or syntax. Additionally, the use of the International Phonetic Alphabet reduces the workload for the Natural Language Processor in terms of prosody as well—the IPA input to the processor contains all relevant prosodic information, including both phrasal and lexical stress. Therefore, the Natural Language Processor does not need to assign prosodic structure such as phrasal and lexical stress to the string via the morpho-syntactic analyzer.

Concerning the Digital Signal Processing module, there are two commonly-employed methods. The first isknown as a rule-based synthesizer. According to Dutoit, such synthesizers are mostly favorable with phoneticians and phonologists. Rule-based synthesizers always appear in the form of formant synthesizers, which employ multiple parameters related to formant frequencies and glottal waveforms. One such synthesizer employed as many as sixty parameters. (Dutoit 1995: 93) However, these different parameters tend to complicate matters, producing analysis errors in speech synthesis, leading to the desire for an alternative method of synthesis.

The favored alternative to rule-based synthesis is concatenative synthesis. In contrast to rule-based synthesizers, concatenative synthesizers embed the particularities of the language within the segments as they are processed. Concatenative synthesizers employ different units of concatenation. Examples of such units include phones, diphones, half-syllables, and triphones (Dutoit 1995: 141). Phones are single units each equivalent to one sound of the language. Diphones begin in the middle of the stable state of a phone and end in the middle of a following phone. Triphones begin in the middle of a stable state of a phone, include the entire following phone, and end in the middle of the next phone. Each of these methods of concatenative synthesis has its benefits and drawbacks further discussed in the following section.

3. Concatenative Speech Synthesis

There are many different units to be used in concatenative speech synthesis, including phones, diphones, triphones and syllables. As exemplified below, diphones are the best choice for the current purposes.

As previously mentioned, diphone-based speech synthesis is a form of unit selection. In this case, the unit of interest is a diphone, which consists of the second half of one phone and the first half of the adjacent phone. In this case, a phone may include silence, to indicate the beginning or end of the phrase. For example, a diphone may consist of a consonant and a vowel, or of a consonant and silence (end of the phrase).

Diphone synthesis is an effective method of concatenative speech synthesis in that it is more natural than single phone speech synthesis. Because diphone synthesis employs units of two phones, the transitions between phones remain intact in the synthesized speech, providing a more natural sound.

One advantage of diphone synthesis is that less memory is needed than if larger units were used for concatenation. Examples of larger units include disyllabic synthesis or even lexical (word) synthesis. Previously, diphone synthesis also has an advantage over syllable synthesis, namely because although 'the number of different syllables in each language is considerably smaller than the number of words, but the size of unit database (of syllables) is still usually too large for TTS systems' (Elshafei et al. 2002: 258). However, this is no longer a problem, as memory space for a TTS system is not difficult to come by. Additionally, it is difficult to control prosodic contours during utterances using syllables as well as diphones, as each syllable (from which diphones are extracted) contains its own prosodic contour (2009: 259). Diphone synthesis is here the preferred method of concatenative synthesis, due to the relative ease of imagining all possible diphones in a language, as opposed to all possible syllables. If, however, syllable synthesis were used, the database would look much different than that of diphones. While a diphone consists of, ideally, the second half of one phone and the first half of the adjacent phone, a syllable can take multiple forms, from CV to CV:CC in Cairene Arabic. Thus a syllable synthesizer would run from a database that contained differently-sized units, unlike a diphone synthesizer, which runs from a database in which all the units are around the same size.

The definition of a diphone is an extension to from the central point of the steady-state part of the phone to the steady-state part of the adjacent phone (Elshafei et al. 2002: 259). Therefore, diphones are concatenated at the steadiest points of phones, leading to a relatively stable and fluent synthesized speech. Additionally the use of diphones causes coarticulation rules to become obsolete: all the coarticulation processes are found in the transition between the two phones, and are therefore already encoded within the diphone (2002: 260).

However, a major drawback to using diphones as units for speech synthesis concerns coarticulatory effects that would affect multiple phones. Often, these effects do not extend over more than the adjacent phone (260). This is especially problematic for a language such as Cairene Arabic, where the coarticulatory effects of an emphatic segment can extend well beyond the adjacent phone. Therefore, consideration of these phonological processes is necessary when constructing a diphone database for Cairene Arabic. The ideal set would have to take into account all possible phonemes as well as allophones that could occur as a result of a non-adjacent emphatic segment.

A final drawback to using diphones as units for speech synthesis comes from the imprecise definition of the boundaries of a diphone. The diphone is defined to extend from the stable point of one phone to the stable point of the adjacent phone. However, what exactly determines when a phone reaches a stable point is obscure and at the discretion of the researcher, especially when highly transient sounds are involved. (Dutoit 1995: 140)

Despite possible problems with using diphones as the unit of concatenation for the speech synthesis, diphones appear to be the best unit for concatenated synthesis. A diphone database requires less storage space than a database that uses larger units for synthesis (syllables, lexical items, et cetera), leading to faster processing of information. Additionally, diphones are relatively painless to extract from pre-recorded speech, as opposed to units such as syllables or lexical items. The necessary recording for a fully-specified database of diphones is much easier to obtain than that of a database of syllables or even sub-syllables, due to the finite number of diphones.A fully-specified database accounts for allophones that occur as a result of coarticulation, even the effects of coarticulation on a non-adjacent phone. This specification is difficult to come by when using units of syllables or triphones. Concatenation by single phone is ideal in terms of database size, but speech synthesized from single phones is in no way an improvement on the intelligibility that comes from diphone speech synthesis. Consequently, the goal to create a text-to-speech synthesizer of Cairene Arabic will be synthesized using diphone-based concatenated synthesis.

4. Problems with Arabic Speech Synthesis

While there have been previous attempts to develop an Arabic text-to-speech synthesis program, these attempts have run into many problems. A few of these issues are discussed further. The solutions to these problems, if applicable, are also presented.

4.1 The Absence of Diacritical Marks

The majority of Modern Standard Arabic texts are written without diacritical marks. Exceptions to this general rule include diacritical marks to avoid a possible misinterpretation or to indicate the proper pronunciation of a proper name, especially if the name in question is foreign. The diacritical marks in use in Modern Standard Arabic are the short vowels, *a*, *i,* and *u,* and the gemination mark, the sukkun, to indicate the doubling of a consonantal segment (for instance, *rr*). Those who can read Arabic fluently do not require diacritical marks in order to fully understand the text. However, diacritical marks are often used in children's books as well as books for second-language learners of Arabic, in order to facilitate better understanding of the text.

However, when attempting to create a speech synthesizer that uses Modern Standard Arabic text as its input, this prominent absence of diacritical marks in the majority of texts presents a large problem. When diacritical marks are used, the text-to-speech synthesizer has a relatively easy task mapping the orthography of Modern Standard Arabic to phonemes, which are then synthesized into speech. The one-to-one orthography to phonememapping that allows for this ease in Modern Standard Arabic was discussed in Chapter 2.

Adding diacritical marks to a text in Modern Standard Arabic orthography via a computerized algorithm is not a simple process. Because of the complexity of the Arabic language, which is based on triconsonantal roots with patterns of vowels to indicate certain meanings or inflections, automatic generation of these forms becomes rather tricky. For any particular unvocalized item written in Modern Standard Arabic orthography, there may be a number of possibilities for vocalization. Therefore, in order to properly identify the diacritical marks necessary, the algorithm must employ a morpho-syntactic analyzer. This analyzer would thus determine, on the basis of probability, which vocalization pattern is applicable. However, as previous attempts have demonstrated, such algorithms are not perfect.

Several attempts have been made for automatic generation of diacritical marks, with relatively high degrees of success. For automatic generation of diacritical marks to be a success, the computer must take into account a variety of different morphological factors, including role the element plays in the utterance, verbal inflection due to different subjects, and case markings that may or may not be of importance, depending on the word ordering (previously discussed in Chapter 2).

Elshafei, al-Muhtaseb and Alghamadi (2006) attempted to provide machine generation of Arabic diacritical marks. Using a Hidden Markov Model, their model considered the word sequence of unvoweled Arabic text as an observation sequence, and the possible diacriticized expressions of the words as differentoutput states (2006: 2). The model was trained using a fully diacriticized Arabic corpus (5). Following training, the model was then tested, first using randomly selected sentences (without diacritics) from the training corpus, resulting in an error rate of less than 0.5%. Next, the model was tested with text from outside the training corpus, again without diacritics. This resulted in an error rate eleven times as much as the previous, at 5.5% (6). However, this is still a relatively low error rate, considering the numerous possibilities that could have led to an error.

The model developed, tested and trained by Elshafei, al-Muhtaseb and Alghamdi (2006) was not without problems. The model can only work with completely undiacriticized text, something which is a bit unrealistic, considering that the majority of texts in the Arabic language consist of partially diacriticized lexical items. Ideally, the model should be able to work with such partial diacriticizations in order to decrease its word error rate. Additionally, 5.5% could still be considered a high error rate, although it is quite good for a computerized algorithm. The model did not vocalize cases, which could be seen as a problem, especially if the model employed free word order found in the syntax of Modern Standard Arabic. Whether this was problematic was not discussed. As it stands, the model is a relatively good method to automatic generation of diacritical marks.

In another model, Habash and Rambow (2007) attempted diacriticization of Arabic via full morphological tagging. In this model, an automatic tagger attempted all possible combinations of diacritical marks. These were then entered into a lexeme database, used to determine the most likely set of diacritics for the given lexical item in an utterance. Compared to the 2006 model by Elshafei, al-Muhtaseb and Alghamdi, the error rate is quite high: 10.9%. (2007: 4). Simply examining the error rates (10.9% compared to 5.5% by Elshafei, al-Muhtaseb and Alghamdi) is not enough to determine that the automatic diacriticization by Elshafei, Muhtaseb and Alghamdi was more successful than that by Habash and Rambow. The model by Elshafei, al-Muhtaseb and Alghamdi was trained, then tested, while the model by Habash and Rambow was not trained. Therefore, the 10.9% error rate by the Habash and Rambow model leaves much room for improvement, and if trained, perhaps would be less than the 5.5% error rate by Elshafei, al-Muhtaseb and Alghamdi. Without this knowledge, it is impossible to determine which model has a better success rate.

Examining different methods to generate diacritical marks for Arabic texts clarifies that generation of diacritical marks causes many problems for a text-to-speech synthesizer. For this project, this problem has already been solved, without the aid of a morpho-syntactic analyzer. It was previous assumed in Chapter 2 that the input was Arabic orthography with all relevant diacritics. Additionally, the input to the text-to-speech synthesizer is text in the form of the International Phonetic Alphabet, rather than Arabic orthography. Consequently, there is no need for any sort of automatic generation of diacritical marks, as these are no longer an issue when assuming that the input to the system is text in the International Phonetic Alphabet.

4.2 Modern Standard Arabic as the Language of Synthesized Speech

The majority of text-to-speech systems designed for the Arabic language are designed for the standardized form of the language: Modern Standard Arabic. However, what these text-to-speech systems do not take into account is the relative scarcity of Modern Standard Arabic as an everyday language within the general population of the countries in which Modern Standard Arabic is the language of the government and other formal purposes. In the majority of these countries, the illiteracy rate is quite high, especially when compared with nations in other parts of the world (notably Europe and North America). According to the United Nations, the majority of these people thus speak only the colloquial Arabic of the region, rather than the standardized form of Arabic across the world.

In an ideal world, of course, the illiteracy rate would be 0. However, since the world is not ideal, a different situation must be taken into consideration: a text-to-speech system that synthesizes Modern Standard Arabic speech from Modern Standard Arabic text is not helpful to a large portion of the population. Therefore, what would be more helpful is a synthesized speech system that synthesized the colloquial speech of the population, so the population could thus have access to written texts in Modern Standard Arabic orthography. This is exactly what this project intends to do: bridge the gap between Modern Standard Arabic and colloquial forms of the language, using Cairene Arabic as a prototype. For actual use in the world, perhaps this is not the most realistic goal, especially considering the expense that would be incurred were a speech synthesizer to be supplied to all members of a population. In theory, however, a speech system that takes Modern Standard Arabic input and outputs a colloquial Arabic dialect could be considered a good idea.

5. *Synthesis of Modern Standard Arabic*

Despite the problems encountered by text-to-speech synthesizers for Arabic, there have been numerous attempts to devise a fully-functioning synthesizer for the standardized form of the language. A few of these attempts are discussed in the following sections.

5.1 *2002: Synthesis with Diphones and Sub-syllables*

Elshafei, Al-Muhtaseb and Al-Ghamdi (2002) proposed a synthesizer for Modern Standard Arabic that used diphones and sub-syllables as its unit from which to concatenate speech segments. Importantly, before the synthesizer can work, automatic vowelization (vocalization) must first be applied (256). This could be seen as a drawback, since, as previously mentioned, automatic generation of diacritical marks leads to morphological errors.

This synthesizer employed a Natural Language Processor, the specifications of which were outlined in the second section of this chapter. This processor produced a phonetic transcription of the text, which was originally in the orthography of Modern Standard Arabic. The morphosyntactic analyzer used to generation the automatic diacritics was contained within the Natural Language Processing Engine. (256-257) The synthesizer then employed a Letter-to-Sound module, based on the phonetic transcriptions provided by the Natural Language Processor. This module was both rule- and dictionary-based, in order to reduce the amount of articulation errors by the synthesizer.

The unit of concatenation for this synthesizer was both diphones and sub-syllables (usually triphones). As a result of this unit of concatenation, the rules of coarticulation need not be formulated, as discussed in the second portion of this section. The possible diphones were divided into four classes: vowel-vowel, consonant-vowel, vowel-consonant, and consonant-consonant (260). To generate all the possible diphones, the phoneme set of Modern Standard Arabic was used, in addition to all the allophones found within the language. These diphones were then recorded by a native speaker of Arabic, as the middle syllable of a nonsense word. The diphones differ in terms of stress and nasalization, and vowels in accentuation and dilution. (261)

The success of the synthesizer was relatively high. Because 55% of the syllables within Modern Standard Arabic are CV or CV: structure, a diphone-based synthesizer is quite intelligible, since the majority of the syllables themselves are not much larger (in terms of segments) than the diphones. The syllables are either two or three segments, CV or CV:, while the corresponding diphones and sub-syllables would thus consist of the second half of a consonantal segment, and the first half of a vocalic segment. These diphones thus provide a good representation of the language. However, because of the use of diphones that occasionally cut across syllable boundaries, the synthesizer has a significant effect of coarticulation, which possibly distracts from the intelligibility and naturalness of the speech synthesizer. Finally, in order for the synthesizer to correctly work for all possibilities within

Modern Standard Arabic, a number of diphones must be added to account for loan word, especially those that have a CCCV string, not usually a legal sequence in the language. Generally speaking, the synthesizer can be considered an advancement in terms of quality Arabic speech synthesis.

5.2 2005: Diphone Synthesis

Ghattas and Nour (2005) implemented an unlimited vocabulary text-to-speech synthesis using diphones to generate Modern Standard Arabic speech. In this synthesizer, as in the previously discussed synthesizer from 2002, units of prerecorded speech were concatenated to synthesize the language. Again, the synthesizer used text in the orthography of Modern Standard Arabic, written with full diacritics, as input.

The diphones relevant for the study were obtained from the permutation of phonemes and allophones within Modern Standard Arabic, and recorded by a selected speaker (159). The diphones were recorded in nonsense words that took into account the relevant prosodic aspects of the language (162). Overall, 238 sentences were recorded, covering only 53% of the possible diphones of Modern Standard Arabic, but all the phonemes.

Following the recording of the diphones and the creation of the synthesized speech, the synthesizer was then tested on six native listeners of Arabic. The average percentage of sentence accuracy was 67%, while word accuracy was a higher 87%. However, there was only an average of three words per sentence, a relatively low average when compared to that of sentences in natural speech. (164) The high degree of word accuracy seems to indicate that the speech synthesizer is a relative success, but when compared to the lower sentence accuracy (and the low number of items in a sentence), the success does not seem as significant.

5.3 2009: Diphone Synthesis using Festival

In her 2009 thesis, Assaf developed a prototype of an Arabic diphone speech synthesizer using the open source program Festival. Again, the language synthesized is Modern Standard Arabic. This speech synthesizer works from a basis of two hundred read sentences, recorded by a native speaker of Arabic. The lower number, according to Assaf, demonstrates that only a small bit of recorded text is necessary to generate recognizable synthesized speech (2). Based on the programming of Festival, the speech synthesis again used diphones as the unit of concatenation.

The synthesized speech was then evaluated by a group of native Arabic speakers. This was tested using the Diagnostic Rhyme test and the Modified Rhyme test, in order to test for intelligibility. In the first listening, 85% of the words were heard accurately (39). However, only 45% of the sentences were heard accurately in the second listening, compared to 30% in the first listening (41). The speech synthesizer developed here runs into the same problem as that of the 2005 synthesizer: word accuracy is much higher than sentence accuracy. The accuracy rates for the synthesizer developed in Festival correctly indicate that a somewhat recognizable diphone speech synthesizer can be developed using Festival and a minimal number of recorded phrases. However, Assaf's work does not account for the automatic diacritization of Arabic, and instead assumes that the input to the synthesizer contains all relevant diacritics. In this respect, it is perhaps illogical to compare the results of this work to the work done in 2002 and 2005, both of which used an algorithm to generate diacritics automatically.Because the diacritics were generated automatically, there is much room for error. In the study by Assaf, the diacritics were already given, and thus there was no possible error. The study by Assaf is thus comparable to the current project, in which diacritics are pre-specified leaving no room for error.

6. Synthesis of Cairene Arabic

There has been one major attempt at dialectical speech synthesis of which I am aware: Tomokiyo, Black and Lenzo (2003) described a synthesizer for Egyptian Arabic, also described by Waibel and others (2003). The data used to create the synthesizer was taken from recorded spoken Egyptian Arabic, which was transcribed by native speakers in order to facilitate accurate synthesis. The influence of Modern Standard Arabic was removed from the transcriptions, in order to achieve transcription of realistic Egyptian Arabic, as opposed to a form of Egyptian Arabic derived from the orthography of Modern Standard Arabic. (Tomokiyo et al. 2003: 2050)

The synthesizer employed letter-to-sound rules, in a manual diphone-based system. The synthesizer was 'trained' using a 'fully diacriticized set of phonetically balanced sentences' (2003: 2050). Intended to be a full speech-to-speech translation on a handheld device, the synthesizer requires an input of speech, which must then be transcribed into the system via a speech-to-text algorithm. This transcription must then be run through the algorithm of text-to-speech, with a final output of Egyptian Arabic speech (Tomokiyo et al. 2003: 2050). The input to the system can be either English or Arabic, with the output as Egyptian Arabic (2003: 2051). The translation used explicit language-independent formalism, based on speaker intention (request, statement, offer, exclamation) as opposed to literal meaning. The synthesizer for the Arabic language was built using Festival Speech Synthesis System, constructed from around 7500 sentences. (Waibel et al. 2003)

Unfortunately, while the synthesizer was then tested, the methods by which the synthesis was conducted are not detailed. What would have been invaluable for the current purposes would be the algorithm behind the transformation from Modern Standard Arabic text (transcribed from the speech) to Egyptian Arabic speech.

The synthesizer was evaluated (in terms of its output of Egyptian Arabic speech) using the Diagnostic Rhyme and Modified Rhyme tests. Additionally, a sentence-level test was used in which listeners were asked 'to mark any words that sounded bad' (Tomokiyo et al. 2003: 2052). The Diagnostic Rhyme Test had an accuracy rate of 78.3%, the Modified Rhyme Test a rate of 72.0%, and the sentence-level test an accurate rate of 84.7% (Tomokiyo et al. 2003: 2052). That the sentence-level test had a noticeably higher accuracy rate than either of the word-level tests indicate a reliance of the listeners on the context for intelligibility of individual words, as would be expected.

Despite the problems encountered by the synthesizer, the algorithm developed by Tomokiyo, Black and Lenzo demonstrates that it is possible to create a synthesis of a colloquial form of Arabic. Additionally, that the algorithm can have English or Modern Standard Arabic as input demonstrates that it is possible to transform Modern Standard Arabic in such a way that Egyptian Arabic is the output of the synthesized speech. This is important for this project, which aims to do the same, only from Modern Standard Arabic text rather than speech, to Egyptian Arabic speech. In this project, instead of using the algorithm developed by Tomokiyo, Black and Lenzo, the open source program *Festival* will be used to develop a speech synthesizer. This was previously done by Assaf (2009) for Modern Standard Arabic, discussed in the following section.

7. Speech Synthesis in Festival

As previously mentioned, in her 2009 thesis Assaf demonstrated that it was possible to develop an Arabic speech synthesizer using the framework provided by the open source program Festival. Assaf's work will be used a foundation for the work done in this project, namely the framework for a speech synthesizer of Cairene Arabic using Festival.

The first step to developing a Cairene Arabic speech synthesizer in Festival is to specify the phone set. Additionally, Festival requires a specification of the articulatory features for each phoneme. Following this, the necessary diphones for a speech synthesizer need to be discussed, as well as the most efficient way to record these diphones. This will further be laid out in the following chapter, employing features similar to those specified by Assaf 2009. Additionally, Festival requires letter-to-sound rules in order to synthesize speech. From this, the letter input to Festival are mapped to corresponding phones, which in turn correspond to a specific sound, specified with features. Assaf matched the phones of Modern Standard Arabic using a modified version of the International Phonetic Alphabet, in order to fit within the formatting system allowed by Festival.

While the exact diphone set for Cairene Arabic, has not been determined as of yet, Assaf provided a basis for this by specifying the following different categories for diphones in Modern Standard Arabic:

(1) vowels – vowels
(2) vowels – consonants— vowels
(3) vowels – emphatic consonants – vowels
(4) consonants – consonants
(5) emphatic consonants – consonants
(6) consonants – emphatic consonants

(7) silence – vowels — carrier word — vowels — silence
(8) silence — consonants — carrier word – consonants – silence
(9) silence – emphatic consonants – carrier word – emphatic consonants – silence
(10) double emphatic consonants

(Assaf 2009: 35)

These sets of diphones were then recorded by a native speaker of Arabic, in the form of nonsense words within two hundred utterances. From these recordings, extractions were made of the different diphones, which were then added to a database to be employed by the *Festival* program.

Many of the same tactics will be employed in the following chapter to develop a foundation for a possible text-to-speech synthesis of Cairene Arabic using *Festival*.

8. Prosodic Contours

In order to synthesize realistic, natural speech, prosodic contours (pitch and accent) must be provided. There are multiple approaches to generating the prosodic contours of the fundamental frequency (F_0).

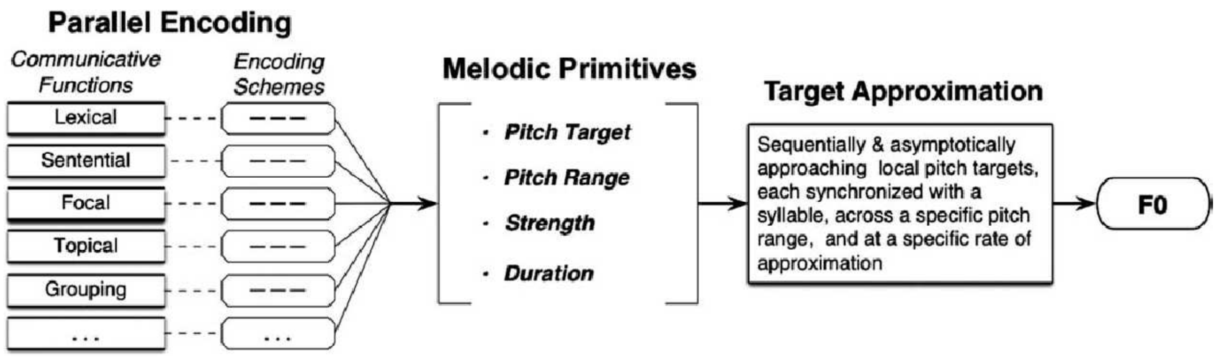
The first approach uses Fujisaki's model. This model, proposed during the 1970s and 1980s, 'assumes that the F_0 contour is the superposition of two contributions: a phrase component and an accent component, obtained by filtering two signals' (Rossi et al. 2004: 1019). The phrase component accounts for the declination of the fundamental frequency, and is characterized by a fast rise in the fundamental frequency, followed by a slow fall. The accent component is modeled by small-scale prosodic variations. The model uses logarithmic formulas to account for these prosodic variations. (1020)

Although Fujisaki's model is 'theoretically compatible with different linguistic descriptions… in practice, it is hardly compatible with any' (Dutoit 1995: 166). Because the model is acoustically-based, it is often difficult to establish a relationship between the natural F_0 contours that occur within a recorded utterance and the syntactic-prosodic organization (166). Therefore, the model usually is reduced to imposing artificial prosodic contours, which often lead to unnatural-sounding synthesized speech.

There exist other acoustic models for generating prosodic contours. However, these models generally encounter the same difficulties as Fujisaki's model: unable to establish a relationship between the recorded $F_0$and the syntactic organization, the models artificially impose prosodic contours. One such model is termed a narrow-stylization model, which employs top-down methodology. In this case, prosodically meaningful segments are organized into a hierarchy. Each point on the hierarchy is then associated with a specific melodic contour (Dutoit 1995: 167). However, the phrase as a whole has a baseline fundamental frequency, to which the model constantly returns, leading to a more artificial sound (Xu 2005: 245). This automatic organization into the model's hierarchy can lead to a number of prosodic difficulties. One such difficulty concerns the F_0 contours that already exist within the concatenated units. When combining the contours of the fundamental frequency within the unit with those added via automatization, there is a discrepancy. The model will either compromise, giving an average of the two different fundamental frequencies, or use both in tandem. Both choices lead to unnatural sounding speech, and therefore a solution for how to generate prosodic contours is still necessary.

A more appealing option for generation of prosodic contours is through a sequence of tones. Tone-based models in text-to-speech systems produce descriptions of the intonation in terms of tone sequences, based on the type of sentence (declarative, interrogative, imperative), the number of syntactic components, and the stress structure (Dutoit 1995: 168). This model assigns specific tones, which correspond to specific F_0 values, to certain elements within the utterance. For instance, the neutral declarative intonation of English was described by Pierrehumbert (1981) as follows: 'pitch accents are all **H**, phrase-final tones are typically **L-L** for a terminal phrase and **L-H** for a non terminal one' (Dutoit 1995: 168). After assignment of specific F_0 values to certain points, the F_0 contour can be created by modulation over these specific points. However, a major downside to the use of this model is the requirement of a tone grammar. Arabic (both Modern Standard and Cairene) does not employ lexical tone within its grammatical structure, and is thus ineligible for this model. Instead, Arabic employs phrasal tone, remarkably similar to that of English, where a declarative statement rises in tone until the phrasally stressed element, after which tone falls.

A third option to generate prosodic contours involves the PENTA model (Xu 2005). This model was developed through the study of pitch contours in both tonal and non-tonal languages. These studies demonstrated, among other things, that 'male and female speakers do not differ much in maximum speed of pitch change...tak[ing] about 100ms … to change pitch by even the smallest amount' (Xu 2005: 224). Additionally, pitch was determined to be a syllabic property, where 'an underlying tonal target starts at the onset of the host syllable and end at the offset of the syllable' (224). Over the entire phrase, 'the basic intonational units are pitch accents, which are liked to words rather than syllables' (230). The prosodic contour of the utterance is derived when interpolating the F0 of unaccented syllables between accented syllables, something that can be done automatically. The PENTA model is diagrammed below:



(Xu 2005: 243)

The PENTA model determines the prosodic contours based on the communicative functions of the utterance, depicted in the far left column of the model. Multiple communicative functions may be used at the same time, and each is encoded differently, to be determined depending on the specifications of the particular model. These encoding schemes specify the values of the melodic primitives: pitch target, which can be static or dynamic; pitch range, the scope of the pitch across targets; strength, the speed at which a pitch target is approached; and duration, the length of time in which the pitch target is approximated. These melodic primitives are combined to provide the complete target approximation for the phrase, which is used to developed the prosodic contour of the fundamental frequency over the phrase. (Xu 2005: 243)

The PENTA model is preferable over the previously discussed models, as it takes into account different communicative factors that determine the overall prosodic contour of an utterance. Additionally, the PENTA model assumes syllable-synchronized sequential target approximation, accounting for the fact that pitch is a feature of an entire syllable, rather than a single segment (Xu 2005: 245). The interpolation of pitch over the syllable, in combination with the pitch change in 100 ms intervals, results in a more natural-sounding prosody, which can be used to communicate different functions. These functions include interrogatives, declaratives, and focus. The current project focuses only on interrogative and declarative (including exclamatory) phrases, in combination with lexical stress.

In addition to the F_0 contour, the text-to-speech synthesizer must take into account stress. An utterance built from a diphone database, if nothing is done to account for stress, will excise some diphones from accented syllables, which may lead to disfluency in the synthesized voice. These diphones taken from accented syllables may 'have phonetic properties that are typical of this accented context only... their temporal and spectral characteristics are close to the target values that one finds in the production of carefully spoken, isolated, accented syllables' (Drullman and Collier 1991: 1766).

To account for this, Drullman and Collier (1991) extracted accented and unaccented VC and VC diphones from isolated nonsense words. These diphones were then used to synthesize different conditions: isolated words, pairs or triples of words, and five or six word sentences. Via perceptual experiments, Drullman and Collier studied whether the quality of the synthesized speech improved when unaccented and accented diphones in the database were used to synthesize unaccented and accented syllables. From these experiments, it was determined that there is no reason to choose accented diphones to indicate secondary stress in synthesized speech (1771). Additionally, the results demonstrated that the choice of accented or unaccented diphones is influential in the perception of syllables with long vowels only (1774). Therefore, these experiments show evidence that underlying stress of a diphone in the database does not play a role in the overall perception of the synthesized speech, and therefore phrasal stress may be added artificially.

In the case in which stress is added artificially, stress is the result of a higher fundamental frequency on the stressed syllable (Dutoit 1995: 169). As discussed in Chapter 3, the stressed syllable in an utterance falls closer to the end of the utterance. Therefore, this syllable receives a higher fundamental frequency, which, according to the PENTA model, will be used to interpolate the fundamental frequency of the remaining unaccented syllables in the utterance. However, interrogatives and declaratives have different prosodic patterns over the phrase. Both begin with practically the same fundamental frequency contour, and the distinction between an interrogative and a declarative is made in the second half of the prosodic contour, with a declarative resulting in a lower fundamental frequency after the high fundamental frequency of the most prominently stressed syllable, and an interrogative resulting in a higher fundamental frequency after the prominently stressed syllable (Xu 2005: 236).

This combination of accent (stress) and prosodic contours of the fundamental frequency lead to more natural-sounding synthesized speech. Of the models discussed above, the PENTA model is favored, due to its ability to account for prosodic contours indicating the function of the phrase, as well as its interpolation of the fundamental frequency over the entire phrase to produce natural-sounding prosody.

9. Conclusions

Based on the above analysis, diphone concatenation is the most applicable method by which to synthesize speech. In addition to having benefits over another form of concatenative synthesis, such as syllable or monophone synthesis, diphone synthesis can be done with relatively little recorded data, as demonstrated in Assaf's 2009 thesis. The prosodic contours of the synthesized speech will be accounted for using the PENTA model (Xu 2005). The PENTA model also accounts for phrasal stress, around which the fundamental frequency is interpolated. Therefore, the PENTA model in combination with diphone concatenation could be used to synthesize Cairene Arabic speech.

The obvious problems concerning Arabic speech synthesis are remedied for the current goals. Instead of working from Modern Standard Arabic orthography as input, the current system uses the International Phonetic Alphabet as input (the output of chapters 2 and 3 here). Consequently, the problems with vocalizations do not exist, as the International Phonetic Alphabet, when written, is entirely vocalized. Therefore, it is only necessary to have letter-to-sound rules.

The following chapter details how a Cairene Arabic speech synthesis program could be implemented, using diphones in combination with the open source program *Festival*.

5: Framework for a Cairene Arabic Speech Synthesizer

1. Introduction

This chapter provides the framework for a text-to-speech synthesizer of Cairene Arabic using the open source software *Festival*. What is discussed is only hypothetical, and does not involve any of the computational work necessary for actual implementation of such a synthesizer.

The form of speech synthesis, as previously discussed, is concatenation using diphones as the unit of concatenation. Thus the chapter begins with an exploration of the different diphones found in Cairene Arabic. This is followed by what exactly would need to be recorded in order to obtain all the different diphones.

Following the hypothetical recording to obtain a complete diphone set for Cairene Arabic, the framework for the program *Festival* is discussed. This framework includes a letter-to-sound mapping, which employs SAMPA rather than IPA, and therefore the IPA input of Cairene Arabic is mapped to the symbols of SAMPA. Additionally, the features for the different sounds are provided in accordance with the guidelines of the software. In combination, all of these different pieces set up the framework for a hypothetical speech synthesizer of Cairene Arabic.

2. Letter-to-Sound Rules

As discussed in the Introduction to this chapter, the International Phonetic Alphabet is not a very convenient system from which computers can synthesize speech. Instead, ASCII codes are preferred, as these represent text in computers in a way that was designed specifically for coding with computers, unlike the International Phonetic Alphabet. For these reasons, the standardized coding of SAMPA is preferred. (Dutoit 1995: 107).

The following describes the transformation (substitution) of the International Phonetic Alphabet to SAMPA. From this point onward, SAMPA will be used to describe in the phonetic input. The correspondence from IPA to SAMPA can be found in the appendix.

Through the use of SAMPA, it is possible to use ASCII codes, making computer processing easier. However, Festival requires pre-specified letter-to-sound rules, in which phones map to themselves with specific features indicated. These are given below, including allophones:

consonant	manner	place	voicing
m	nasal	labial	+
n	nasal	alveolar	+
b	stop	labial	+
d	stop	alveolar	+

t	stop	alveolar	-
d`	stop	alveolar-pharyngeal	+
t`	stop	alveolar-pharyngeal	-
g	stop	velar	+
k	stop	velar	-
q	stop	uvular	-
?	stop	glottal	-
v	fricative	labio-dental	+
f`	fricative	labio-dental	-
z	fricative	alveolar	+
s	fricative	alveolar	-
z`	fricative	alveolar-pharyngeal	+
s`	fricative	alveolar-pharyngeal	-
Z	fricative	palatal	+
S	fricative	palatal	-
R	fricative	velar	+
x	fricative	velar	-
?`	stop	pharyngeal	-
X\	fricative	pharyngeal	+
h	fricative	pharyngeal	-
r	trill	alveolar	+
l	lateral	alveolar	+
j	vocalic	palatal	+
w	vocalic	labial	+
m`	nasal	labial-pharyngeal	+
n`	nasal	alveolar-pharyngeal	+
b`	stop	labial-pharyngeal	+
g`	stop	velar-pharyngeal	+
k`	stop	velar-pharyngeal	-
v`	fricative	labio-dental-pharyngeal	+
f``	fricative	labio-dental-pharyngeal	-
Z`	fricative	palatal-pharyngeal	+
S`	fricative	palatal-pharyngeal	-
R`	fricative	velar-pharyngeal	+
x`	fricative	velar-pharyngeal	-
r`	trill	alveolar-pharyngeal	+
l`	lateral	alveolar-pharyngeal	+
j`	vocalic	palatal-pharyngeal	+
w`	vocalic	labial-pharyngeal	+

The specifications for vowels are given as follows:

vowel	length	height	front	round	pharyngeal
i	short	high	front	-	-
u	short	high	back	+	-
a	short	low	front	-	-
A`	short	low	back	-	+
i:	long	high	front	-	-
u:	long	high	back	+	-
a:	long	low	front	-	-
A:`	long	low	back	-	+
i`	short	high	front	-	+
u`	short	high	back	+	+
i:`	long	high	front	-	+
u:`	long	high	back	+	+
ay	dipthong	low	front	-	-
aw	dipthong	low	back	+	-
ay`	dipthong	low	front	-	+
aw`	dipthong	low	back	+	+

The specific letter-to-sound rules must be determined for each speaker. However, in general form they appear as:

```
(lts.ruleset
; Name of rule set
Cairene Arabic
)
; Rules
(
  ( [ a ] = a )
( [ b ] = b )
( [ t ] = t )
  ( [ t` ] = t` )
```

These letter-to-sound rules specify the which sound (as given in the above charts) corresponds to a letter, given in the SAMPA input).

3. *Diphone Database*

The number of possible diphones in any language is the square of the number of phones. The majority of languages do not make use of all possible diphones. However when taking into consideration possible diphones that can occur across lexical boundaries, the number of diphones is much greater than when solely considering diphones that can occur word-internally.

Therefore, the number of diphones in Cairene Arabic is almost the maximum number of diphones. The possible restrictions on diphones are discussed below.

Additionally, the diphone database must take into account allophonic variations of the different phones. As discussed in the third chapter, almost every phone in the inventory of Cairene Arabic has an allophone as the result of emphasis spread. Thus, every phone (with the exception of /q/) has an allophone with the notation s`. The exceptions to this include q, s, t, d, z, s`, t`, d`, and z`. This results from the fact that s`, t`, d` and z` are already phones within Cairene Arabic (in fact, the very phones that cause the emphatic allophones to occur), and therefore these are also the allophones of the phones s, t, d, and z. Additionally, the vowels /a/ and /a:/ do not have emphatic allophones, as these allophones are /□□/ and /□:□/, already a part of the phone inventory of Cairene Arabic.

Both long vowels and diphthongs are considered single phones. For Cairene Arabic, there are thus the follows sounds which need to be considered in the diphone database:

24 unemphatic consonantal segments

22 emphatic allophones of consonantal segments (excluding /q/, /?/)

4 emphatic consonantal segments

8 unemphatic vocalic segments (2 diphthongs, 3 long vowels, 3 short vowels)

6 emphatic allophones of vocalic segments (2 diphthongs, 2 long and 2 short vowels)

2 emphatic vocalic segment

The diphones to be used can be classified into the following categories=

(1) Consonant – Consonant
(2) Consonant – Vowel
(3) Vowel – Consonant
(4) Vowel – Vowel
(5) Silence – Consonant
(6) Consonant – Silence
(7) Vowel – Silence

The only categories excluded from this is that of Silence – Vowel and Vowel – Vowel.Because of the prohibition in Cairene Arabic against onset-less syllables, it is impossible to begin an utterance with a vowel. Therefore, there are no members of the category Silence – Vowel. Additionally, the prohibition in Cairene Arabic against onset-less syllables prevents any occurrences of Vowel—Vowel, as there will always be (minimally) a glottal stop between the two. Diphthongs and long vowels are, as previously mentioned, not considered to be a part of the category Vowel—Vowel, as these are considered single phones. All other categories are found within the language.

Using these categories in combination with the above sounds, the diphone database is constructed as follows:

(1) Consonant – Consonant
   a. unemphatic consonantal segment – unemphatic consonantal segment
   b. unemphatic consonantal segment – emphatic consonantal segment
   c. emphatic consonantal segment – unemphatic consonantal segment
   d. emphatic consonantal segment – emphatic allophone consonantal segment
   e. emphatic allophone consonantal segment – emphatic consonantal segment
   f. emphatic consonantal segment – emphatic consonantal segment
   g. emphatic allophone consonantal segment – emphatic allophone consonantal segment
   h. emphatic allophone consonantal segment – unemphatic consonantal segment
   i. emphatic consonantal segment – unemphatic allophone consonantal segment

(2) Consonant – Vowel
   a. unemphatic consonantal segment – unemphatic vocalic segment
   b. emphatic consonantal segment – emphatic vocalic segment
   c. emphatic consonantal segment – emphatic allophone of vocalic segment
   d. emphatic allophone of consonantal segment – emphatic vocalic segment
   e. emphatic allophone of consonantal segment – emphatic allophone of vocalic segment

(3) Vowel – Consonant
   a. unemphatic vocalic segment – unemphatic consonantal segment
   b. unemphatic vocalic segment – emphatic allophone of consonantal segment
   c. unemphatic vocalic segment – emphatic consonantal segment
   d. emphatic vocalic segment – emphatic consonantal segment
   e. emphatic vocalic segment – emphatic allophone of consonantal segment
   f. emphatic allophone of vocalic segment – emphatic allophone of consonantal segment
   g. emphatic allophone of vocalic segment – emphatic consonantal segment
   h. emphatic vocalic segment – unemphatic consonantal segment
   i. emphatic vocalic segment – emphatic allophone of consonantal segment
   j. emphatic allophone of vocalic segment – unemphatic consonantal segment
   k. emphatic allophone of vocalic segment – emphatic allophone of consonantal segment

(4) Silence – Consonant
   a. silence – unemphatic consonantal segment
   b. silence – emphatic allophone of consonantal segment
   c. silence –emphatic consonantal segment

(5) Consonant – Silence
   a. unemphatic consonantal segment – silence
   b. emphatic allophone of consonantal segment – silence

    c. emphatic consonantal segment – silence
(6) Vowel – Silence
    a. unemphatic vocalic segment – silence
    b. emphatic vocalic segment – silence
    c. emphaticallophone of vocalic segment – silence

Restrictions on the phones of these categories are given below:

(1) Consonant – Consonant
    a. No restrictions: two adjacent phones across a word boundary are considered a possible diphone, leaving the possibilities unlimited.
    b. Total number of diphones: 2008
(2) Consonant – Vowel
    a. An emphatic consonant (or emphatic allophone) is always followed by an emphatic vowel (or emphatic allophone), with the exception of /q/ and /?/.
    b. Total number of diphones: 393
(3) Vowel – Consonant
    a. No restrictions; if an emphatic vowel ends a word, the adjacent phone to begin the next word may be either emphatic or unemphatic.
    b. Total number of diphones: 856
(4) Silence – Consonant
    a. No restrictions.
    b. Total number of diphones: 45
(5) Consonant – Silence
    a. No restrictions
    b. Total number of diphones: 45
(6) Vowel – Silence
    a. Only these vowels occur word-finally: /i/, /a/, /A`/.
    b. Total number of diphones: 3

This brings the total number of diphones to 3350.

In Festival, a diphone database 'consists of a dictionary file, a set of waveform files, and a set of pitch mark files.'(*Festival* 21.1) The dictionary files are entries in rows. Each row contains the diphone name (Phone 1 – Phone 2), the filename, start position in milliseconds, midpoint (phone transition) in milliseconds, and end position in milliseconds. A small example of the dictionary files follows:

i—m	im	405.35	440.54	491.23
i—n	in	302.05	350.11	400.31
i—b	ib	109.58	180.45	203.71
i—d	id	599.34	633.21	671.54

The complete list of dictionary files (sans start, midpoint and end positions) can be found in the appendix.

4. Prosody Generation

Simply a diphone database is not enough to produce natural-sounding synthesized speech. In addition to the ordered segments (diphones), pitch contours and accents need to be taken into consideration. As previously discussed in Chapter 4, these prosodic contours of the F_0 will be generated using the PENTA model.

A pre-recorded voice (to generate the diphones) has specified relevant formants (F1, F2, F3 and F4) to produce specific segments. However, a voice has a fundamental frequency that determines the overall pitch. Variations within this fundamental frequency cause the prosodic contours that makes speech sound natural, as well as appropriate for the intention (declarative, imperative, inquisitive, et cetera). Therefore, synthesized speech needs to take into account both the form of the phrase (number of syllables, length of syllables, et cetera) and the intended conversational purpose (function) of the phrase.In the case of this project, predetermined changes in the fundamental frequency will determine the overall prosodic contours of the utterance.

Using the PENTA model, the prosodic contours can be implemented. The PENTA model involves prespecification of each type of communicative functions. The possible communicative functions are declarative, exclamatory, and interrogative. Each are encoded: declarative begins with a low pitch, rises slowly to the pitch accent, with a fast fall at the end; exclamatory begins with a low pitch, rises slowly to the pitch accent, and ends with a slight pitch decrease; interrogative begins with a low pitch, rises slowly to the pitch accent, and then continues to rise. In the scheme of the PENTA model, the melodic primitives are specified depending on the distinctions of the voice recorded for the diphones, but generally appear as follows:

Pitch Target	[high] [low] [rise] [fall] [mid]
Pitch Range	different F_0 values, depending on the prerecorded voice.
Strength	strong, weak, normal
Duration	long, short, normal (value dependent on speed of prerecorded voice)

Therefore, the PENTA model accounts for the prosodic contours as well as the phrasal accent. By assuming an artificial prosodic contour of the fundamental frequency, three different types of communicative functions can be achieved: declarative, indicated by the full stop; interrogative, indicated by the question mark; and exclamatory, indicated by the exclamation mark.

In *Festival*, prosodic contours are generally generated by rule. In this case, a function is used for each syllable in the utterance, and assigns target F0 points within the syllable. The function checks if the syllable is accented, and then returns a list of fundamental frequency targets. For example, a value at the start of the syllable could be 110Hz, at the middle 140Hz, and at the end, 100Hz. (*Festival* 72). Using this, the PENTA model is easily implemented, with a few minor adjustments. The PENTA model is designed in terms of syllables, so that each syllable has a pitch, and pitch changes only occur over spans of 100 ms, as discussed in Chapter 4. Therefore, the function would simply need to be specified as such: only changes over 100 ms, and every syllable has a target pitch. In reality, this is not so different from the function as it exists: small pitch modulations are only natural over the course of a syllable, and therefore it is only pertinent that the target pitch be reached.

Another method of prosody generation in *Festival* uses linear regression. In this case, the prosody is generated by 'an appropriate F_0 target value for each syllable based on available features by training from data' (*Festival* 72). However, because the diphone database that this project would employ is generated from recorded nonsense words, and therefore extracting the F_0 from these syllables as a reliable indication of prosodic is inconceivable. Therefore, a function is necessary, in the style of the PENTA model, to generate the prosodic contours necessary for natural-sounding synthesized speech.

What could be considered problematic for prosody generation is the differentiation between stress and accent. Stress is a property of the syllable, while accent is a property of the lexical item. However, word accent is the choice of the speaker, as 'main lexical stress does not always predict the location of accent within a word' (Shattuck-Hufnagel and Turk 1996: 237). Because word accent is determined by the speaker, the speech synthesizer will disregard word accent. Instead, stress is differentiated into primary and secondary stress. There is only one primary lexical stress within each phrase, and thus this will be taken as the word accent as well. These stresses, as previously discussed, will be synthesized by a change in the fundamental frequency.

5. Recording the Diphone Database

The most common way to record diphones is via nonsense words, in which the diphones are encoded. In this case, the diphones need not be recorded in their accented and unaccented forms, as discussed in Chapter 4. Unfortunately, the sheer number of necessary diphones (3350) would mean that recording each individual diphone in a single nonsense word or string of syllables is time-consuming. It is therefore necessary to record the diphones in a strategic manner, extracting multiple diphones per nonsense word or string of syllables.

For example, from the string /satum/, the diphones that can be extracted are:

silence – s

s — a

a — t

t — u

u — m

m — silence

Extracting every possible diphone from every recorded string of syllables would be a faster way than recording each diphone in a separate string of syllables. However, this speed comes at a price: the quality of the diphones will decrease, depending on their position in the recorded string of syllables.

The recorded strings of nonsense syllables take the following form: $C_1V_1C_2(C_3)V_2(C_4)$. Each recorded string of nonsense syllables must begin with a consonant, since onset-less syllables do not exist in Cairene Arabic. C_1 and C_2 are mandatory, while C3 and C4 are optional. The vocalic segments V_1 and V_2 may be either short vowels, long vowels, or diphthongs. From this schema, the possible diphones that can be extracted are:

silence – C_1

C_1 – V_1

V_1 – C_2

C_2 – C_3/ V_2

V_2 – C_4/silence

C_4 – silence

The complete list of nonsense strings of syllables to be recorded can be found in the appendix. A brief sample appears below:

(1) ?ihma

(2) ?ilX\a

(3) s`A:`s`i`s

(4) ?inka

(5) ?`i`d`ja

(6) ji?Ra

In total, there are 1090 strings of syllables to be recorded. This is rather lengthy, and contains many similar strings. Therefore, to ensure quality of recordings, these would ideally be made over multiple sessions, using a single speaker. The order of the strings to be recorded would be randomized, in order to prevent the same basic string of syllables appearing together (for example, all strings ending with /im/).

All the strings are disyllabic. The majority contain /i/ as the nucleus of the second syllable, due to the use of /i/ in Cairene Arabic as the default unaccented vowel. Additionally, the stress will fall on the long vowel, if there is one in the string. If not, the stress should fall on the first syllable. In the synthesized speech, use of an accented or unaccented prerecorded short vowel will not make a difference, as evidenced by Drullman et al. (1991) .

6. Extracting the Diphones

Following recording of the strings of syllables, the individual diphones need to be extracted from these strings. In *Festival*, the diphones do not need to be extracted into separate files. Instead, the recorded strings of syllables each need to be assigned a file name, from which the database shall automatically extract the diphone.

The diphones are then defined in relation to the file. Each diphone is described within the file with the start point, the midpoint, and the end point, all in milliseconds. The start and end points must be steady states, while the midpoint is the transition between the two phones that constitute the diphone. The diphone index would appearsimilar to the following:

?-i	S1	.343	.512	.671
i-h	S1	.672	.731	.892
X\-l	S2	.893	.978	1.012
n-k	S4	.672	.735	.995

The first and second diphone in the above diagram come from the same audio file, named S1. File S1 contains the string ?ihma. The first diphone begins at 343 ms, and ends at 671 ms. The second diphone, i-h, begins at 672 ms and ends at 892 ms. The third diphone in the diagram comes from the audio file named S2, which contains the string ?ilX\a, and begins at 893 ms and ends at 1.012 ms. The fourth diphone in the diagram comes from the audio file named S4, containing the string ?inka, and begins at 672 ms and ends at 995 ms.

These diphones are then used as the building blocks for the output. The diphone synthesis, in combination with waveform generation, make up the Digital Signal Processor which produces the output of synthesized speech. Thus the recorded diphones are concatenated to produce intelligible synthesized speech of Cairene Arabic.

7. Implementation of the Synthesizer

In order to produce Cairene Arabic synthesized speech with the methodology here, a variety of things would be necessary. This section predicts what would be required in order to actually carry out this project.

7.1 Phonological Transformation

The first step of the project involves the phonological transformation from Modern Standard Arabic orthography to Cairene Arabic in the International Phonetic Alphabet. This transformation involved both substitution/mapping rules and the framework of Optimality Theory. Therefore, an algorithm would have to be developed that could carry out both.

The algorithm that can carry out the framework of Optimality Theory is readily available. In Chapter 3, it was demonstrated that the program PRAAT could be implemented in such a way as to predict the output of the ranked constraints for resyllabification and syllable repair. Therefore, PRAAT can be implemented to predict the output of the ranked constraints for all the different steps in the transformation involving Optimality Theory: lexical syllabification and stress, emphasis spread, and phrasal stress.

Currently, the algorithm that would produce the mapping from Modern Standard Arabic orthography to the International Phonetic Alphabet, as well as make morphological substitutions, does not exist. Such an algorithm would need to be programmed by a person familiar with computer coding. At this moment, this is not feasible by myself, and another person would be needed.

With the help of a computer programmer, it would be thus possible to create an algorithm for the phonological transformation from Modern Standard Arabic orthography to fully-specified Cairene Arabic in the International Phonetic Alphabet.

7.2 Recording and Implementing Diphones in Festival

The second step of the project involves turning the output from the phonological transformation into synthesized speech. As discussed, this is to be achieved using the open source software *Festival*.

Before any work is to be done in *Festival*, it is pertinent that the diphones, in the strings of nonsense syllables, are recorded. This would need to be done by a native speaker of Cairene Arabic, who is also familiar with the International Phonetic Alphabet, thus able to produce the desired strings. Because the amount of material needed to be recorded is vast, it would be necessary to do the recordings over a number of sessions. The recordings would ideally be done in a sound-proof booth set up for professional audio recording.

Following the recordings, the diphones must be extracted from these audio files. This again is a rather lengthy process, one that involves determining the precise boundaries where one diphone ends and another begins. Additionally, there are many duplicate possible diphones in the recorded database of nonsense strings of syllables. Therefore, whomever is determining the beginning and ending points of the diphones must also determine which particular diphone is the ‘best’ of a pair of sounds. Consequently, this must be done by someone who is trained in both audio processing (to extract the diphones) as well as phonetics (to determine which particular diphone exemplifies the two phones). The diphones would then have to be organized via the system of *Festival*, so that they could be extracted when the software is run.

The next step would be to record the different prosodic contours. These different possible contours are specified in Chapter 3. The ideal way to do this would be to record the speaker three times: one declarative, one interrogative, and one exclamatory. The fundamental frequency could then be extracted from these different recordings, and used as the default, to be lengthened or shortened depending on the length of the phrase to be produced by the speech synthesizer.

After diphone extraction and generation of prosodic contours, the voice can be produced in *Festival*. Luckily, *Festival* provides a step-by-step description of how a voice can be created. Therefore, it is only necessary to implement different items as prescribed. Importantly, the letter-to-sound rules must be specified within *Festival* for Cairene Arabic, as Cairene Arabic (or any form of Arabic) is not supported by the main package of *Festival*. Again, this would be a rather lengthy process that would involve much programming. However, this could be achieved by anyone with knowledge of speech synthesis or programming.

The final step would be to check the speech that is synthesized. From this, the intelligibility of the speech must be assessed, as well as the naturalness. As a first attempt, it is not expected that the speech would sound perfectly natural, and perhaps not even be perfectly intelligible. Therefore, a perceptual experiment would be necessary. This experiment would require native speakers of Cairene Arabic as participants, who would then be asked to judge the naturalness and intelligibility of the synthesized speech, by the Diagnostic Rhyme test and the Modified Rhyme test, using both lexical items and phrases as test items. The results of the experiment would thus provide information on areas of sounds or specific sounds on which the speech synthesizer needs to improve. These improvements could then be researched and implemented, leading to a better speech synthesizer.

7.3 Time

Time constraints pose multiple problems when conducting such a large project. The project would require a good deal of time, ranging from the creation of the algorithms to achieve the phonological transformation to the time needed to record the nonsense strings of syllables. Quite possibly the project would take many months, especially when factoring in an end-experiment, to determine the success of the synthesizer. Because of time constraints, the project was not carried out in full here, and instead is only proposed as a possible way to synthesize Cairene Arabic speech from Modern Standard Arabic text, should one ever desire to do so.

7.4 Possible Problems

While in the ideal world, it would be possible to implement the current project perfectly on the first try, this is usually not the case in actuality. Therefore, forecasting possible problems provides a good way to look over the planning of the implementation, as well as the project as a whole.

The largest possible problem is that the phonological transformation to Cairene Arabic is faulty in some way. Being a student of Modern Standard Arabic and not a speaker of Cairene Arabic means that there is always the possibility that a flaw exists in the transformation, of which I am unaware. Were this the case, this should showcase itself in the proposed experiment following completion of an actual speech synthesizer, as the participants in the experiment would explain their dissatisfaction with the speech synthesizer, as well as what specifically these problems were. Using this information, the phonological transformation could thus be repaired to create legal Cairene Arabic phrases as output.

A related problem is the existing vocabulary difference between Modern Standard and Cairene Arabic. The algorithm as it stands accounts only for morphological substitutions, not lexical substitutions. In order to make the output a precise replica of Cairene Arabic speech, it would be necessary to take these lexical suppletions into account. This would need to be done via a large 'dictionary,'from which the algorithm could identify then rewrite any lexical items that differ between Modern Standard and Cairene Arabic not on the basis of phonology or morphology.

Another possible problem exists in the recording of the diphones. It could be that the nonsense strings of syllables are too repetitive, even to be done over multiple recording sessions. Additionally, it could be that the recordings simply do not produce the desired results in terms of diphones. Were this the case, it would be necessary to produce with a different list of strings to be recorded. These strings could be either nonsense strings or actually lexical items or phrases, which would require the help of a native speaker of Cairene Arabic to determine before recording.

There is also the problem that the generation of prosodic contours, via the PENTA model, would not work when implemented in the framework of *Festival*. Were this the case, the prosodic contours would have to be generated using the methods specified by *Festival* itself. An alternative would be to work with one familiar with programming, who could thus ensure that the PENTA model could be implemented via *Festival* to produce synthesized speech.

A final problem is that diphone synthesis as it is done in *Festival* does not produce intelligible or natural-sounding synthesized speech. This is especially true as speech technology advances and people become accustomed to high-quality, realistic-sounding synthesized speech. The open source software *Festival* has stated that it plans to remove the possibility of diphone synthesis in future versions. Therefore, it would be necessary to produce synthesized speech in a different method, which would require recording and extracting different numbers of sounds.

However, all of the possible problems can be dealt with. The solutions to these possible problems, if they arose with actual implementation of the project at hand, would lead to a better speech synthesizer of Cairene Arabic.

8. Conclusions

In this chapter, framework for a speech synthesizer via *Festival* was discussed. Importantly, the requirements for such a synthesizer were expounded upon.

As the project uses diphones as the unit of concatenation for speech synthesis, the different groups of diphones were explained. Prosody generation was also discussed, following the PENTA model as discussed in Chapter 4. Additionally, the full list of diphones was generated, as well as the nonsense strings of syllables to be recorded to generate these diphones is given. These nonsense strings would need to be recorded by a native speaker of Cairene Arabic. Following that, the diphones would need to be extracted from these audio files.

The actual implementation of the project would be time-consuming and require a certain skill set, as well as access to a native speaker of Cairene Arabic and a professional recording booth. The final speech synthesizer would need to be tested in terms of intelligibility and naturalness, through a perceptual experiment using native Cairene Arabic speakers as participants.

Some possible problems that could result from actual implementation of the project are discussed. These problems include possible issues with the phonological transformation, the PENTA model to generate prosody in *Festival*, and recording of the diphones. Additionally, the fact that diphones are somewhat outdated as a method of speech synthesis is discussed.

However, despite these possible problems, the groundwork laid in this chapter provides a good plan for the actual implementation and feasibility of the project.

6. Conclusion

This thesis demonstrated that it is indeed possible to derive Cairene Arabic phonetic transcriptions from Modern Standard Arabic. Beginning with a discussion of phonetics and morphology, Modern Standard and Cairene Arabic were compared. The phonology of the two forms of Arabic were discussed, including syllable and moraic theory, stress, and a few notable phonological processes. The conclusions from this discussion were used in the transformation from Modern Standard Arabic to Cairene Arabic.

The phonological transformation was completed via context-free and context-dependent rewrite rules and constraint-based Optimality Theory. The context-free rewrite rules concerned the phonetic transformation, from the phonemes of Modern Standard Arabic to those of Cairene Arabic. Context-dependent rewrite rules were used to transform the morphological aspects of Modern Arabic to reflect those of Cairene Arabic (namely pronouns and verbal inflections). Stratal Optimality Theory was used to account for the phonological processes in Cairene Arabic. There were two strata, each containing two sub-strata. The sub-strata on the lexical level were syllabification and emphasis spread, followed by lexical stress. The sub-strata on the post-lexical/phrasal level were resyllabification, syllable repair and emphasis spread, followed by phrasal stress. The ultimate output of this was Cairene Arabic, written in fully-specified IPA, which would be used as input for the speech synthesizer.

While an actual speech synthesizer was not developed, the framework for such an algorithm is provided. The proposed synthesizer accounts for a previous problem in Arabic speech synthesis: the lack of accuracy of a morpho-syntactic parser to provide diacritical marks (vowels) for the typically unvocalized Arabic text. Because the proposed synthesizer uses fully-specified IPA as input, a morpho-syntactic parser is not necessary.

The proposed synthesizer employs concatenative synthesis, with diphones as the unit of concatenation. The possible diphones for the language are given, as well as the nonsense strings of syllables that could be used to record the diphone base. Finally, a framework for how to carry out the synthesizer with the diphone database is given, using the open source software *Festival*. The framework also provides a way to account for the prosody of the language, using the PENTA model (Xu 2005).

From the study, it is clear that Cairene Arabic can, in fact, be derived from Modern Standard Arabic. Additionally, this derivation could be used to transform Modern Standard Arabic text into spoken, synthesized Cairene Arabic. The lack of colloquial Arabic in text-to-speech technology is a gap that the study intended to fill. Whether or not such a synthesizer could be carried out in a successful manner remains to be seen. Additionally, this project creates a new void: if it is possible to derive Cairene Arabic from Modern Standard Arabic, it should be possible to derive other dialects of Arabic from Modern Standard Arabic, and also derive Modern Standard Arabic from colloquial dialects of Arabic. Perhaps further research will demonstrate whether or not this is possible.

References

Aboul-Fetouh, Hilmi M. (1969).*A Morphological Study of Egyptian Colloquial Arabic.* The Netherlands.

Abu-Chacra, Faruk. (2007). *Arabic: An Essential Grammar.* Routledge, Oxon.

Abu Ghattas, N. and Abdel Nour, H. (2005). "Text-to-Speech Synthesis by Diphones for Modern Standard Arabic." An-Najah University J. Res Vol. 19.

Al-Haj, H., Hsiao, R., Lane, I. Black, A.W., Waibel, A. (2009). "Pronunciation Modeling for Dialectical Arabic Speech Recognition." *ASRU.*

Al-Jarrah, Rasheed. (2008). "Cairene Arabic Word Stress: A Constraint-Based Analysis." *Dirasat*, Human and Social Sciences, 25:3. 1-16.

Amayreh, Mousa M. (2003). "Completion of the Consonant Inventory of Arabic." *Journal of Speech, Language and Hearing Research* 46. 517-529.

Anttila, Arto. (1995). "Derivation variation from grammar: A study of Finnish genitives." Stanford University, 1-28.

Archangeli, Diana and Pulleyblank, Douglas. (1994). *Grounded Phonology*. The MIT Press:Cambridge, Massachusetts.

Assaf, M.M. (2005). "A Prototype of an Arabic Diphone Speech Synthesizer in Festival." Master's Thesis in Computational Linguistics, Uppsala Universitet.

Broselow, Ellen I. (1976). *The Phonology of Egyptian Arabic*. Amherst, Massachusetts.

Broselow, Ellen I. (1979). *The Phonology of Egyptian Arabic*. Ann Arbor, Michigan.

Chouireb, F. and Guerti, M. (2008). "Towards a high quality Arabic speech synthesis system based on neural networks and residual excited vocal tract model." SIViP 2. 73-87.

Davis, Stuart. (1995). "Emphasis Spread in Arabic and Grounded Phonology." *Linguistic Inquiry*26:3. 465-498.

Davis, Stuart. (2011). "Prosody in First Language Acquisition: Mora Structure in Cairene Arabic and Foot Structure in American English." Talk, Universiteit Utrecht.

Ditters, Wilhelmus Everhardus. (1992). *A Formal Approach to Arabic Syntax.* Luxor: Nijmegen.

Drullman, Rob and Collier, Rene. (1991). "On the combined use of accented and unaccented diphones in speech synthesis." *J. Acoustical Society of America* 90: 1766-1775.

El-Imam, Yousif A. 2004. "Phonetization of Arabic: rules and algorithms." *Computer Speech and Language* 18. 339-373.

El-Sadany, T.A. (2010). "An Arabic morphological system." *IBM Systems Journal.*

Elshafei, M. Al-Muhtaseb, H., and Al-Ghamdi, M. (2002). "Techniques for high quality Arabic speech synthesis." *Information Sciences* 140 (255-267).

Elshafei, M. Al-Muhtaseb, H., and Al-Ghamdi, M. (2006). "Machine Generation of Arabic Diacritical Marks." International Conference on Machine Learning; Models, Technologies and Applications.

Esling, John H. (1999). "The IPA Categories 'Pharyngeal' and 'Epiglottal': Laryngoscopic Observations of Pharyngeal Articulations and Larynx Height." *Language and Speech* 42:4. 349-372.

Ferguson, C.H. (1989). "Diglossia." *Word* 15, 325-340.

Ferguson, Charles A. (1956). "The Emphatic *l* in Arabic." *Language* 32:3. 446-452.

Festival Speech System. The Centre for Speech Technology Research, The University of Edinburgh. http://www.cstr.ed.ac.uk/projects/festival/

Goldwater, Sharon and Johnson, Mark. (2003). "Learning OT Constraint Rankings Using aMaximum Entropy Model." In Jennifer Spenader, Anders Eriksson, and Osten Dahl, editors,46 *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. StockholmUniversity, 111-120.

Habash, N. and Rambow, O. (2007). "Arabic Diacritization through Full Morphological Tagging." *Proceedings of NAACL HLT 2007, Companion Volume.* 53-56.

Harrell, Richard S. (1957). *The Phonology of Colloquial Egyptian Arabic.* New York: American Council of Learned Societies.

Harrell, Richard S., Tewfik, Laila Y., and Selim, George Dimitri. (1963). *Lessons in colloquial Egyptian Arabic.* Washington: Georgetown U.P.

Hassig, Hannah. (2010). "Emphasis Spread in Cairene Arabic: Grounded Phonology, Optimality Theory, and Free Variation." Universiteit Utrecht, unpublished.

Hassig, Hannah. (2010). "From Modern Standard Arabic to Egyptian Arabic: A Syntactic and Phonological Transformation." Utrecht Universiteit, unpublished.

Hellmuth, Sam. (2004). "Prosodic weight and phonological phrasing in Cairene Arabic." *CLS 40: The Main Session,* 97-111.

Hinds, Martin. (1986). *A dictionary of Egyptian Arabic: Arabic-English.* Beirut: Librairie du Liban.

Kiparsky, Paul. (2010). "Reduplication in Stratal OT." In Linda Uyechi and Lan Hee Wee (eds.) *Reality Exploration and Discovery: Pattern Interaction in Language and Life*.

Khoja, S. (2001). "APT: Arabic part-of-speech tagger." *Proceedings of the Student Workshop at NAACL-2001*.

Khoja, S., Garside, R. and Knowles, G. (2001). "A tagset for the morphosyntactic tagging of Arabic."

Laufer, Asher and Baer, Thomas. (1988). "The Emphatic and Pharyngeal Sounds in Hebrew and in Arabic." *Language and Speech* 31:2. 181-205.

Lehn, Walter. (1963). "Emphasis in Cairo Arabic." *Language* 39:1. 29-39.

McCarthy, John J. (1979). "On Stress and Syllabification." *Linguistic Inquiry* 10:3. 443-465.

McCarthy, John J. (1997). "Process-Specific Constraints in Optimality Theory." *Linguistic Inquiry,* 28:2. 231-251.

Norlin, Kjell. (1987). *A Phonetic Study of Emphasis and Vowels in Egyptian Arabic.* Lund: Department of Linguistics, Lund University.

Prince, Alan and Smolensky, Paul. (1993, 2002). "Optimality Theory: Constraint Interaction in Generative Grammar." ROA version: August, 2002.

SAMPA: Speech Assessment Methods Phonetic Alphabet. UCL Division of Psychology and Language Sciences. http://www.phon.ucl.ac.uk/home/sampa/

Shattuck-Hufnagel, Stefanie, Turk, Alice E. (1996). "A Prosody Tutorial for Investigators of Auditory Sentence Processing." *Journal of Psycholinguistic Research* 25: 2. 193-247.

Tomokiyo, L.M., Black, A.W., Lenzo, K.A. (2003). "Arabic in my Hand: Small-footprint Synthesis of Egyptian Arabic." *Eurospeech*, Geneva.

Vergyri, D. and Kirchhoff, K. (2004). "Automatic diacritization of Arabic for Acoustic Modeling in Speech Recognition." Workshop on Computational Approaches to Arabic Script-based Languages. University of Geneva. 66 – 73.

Watson, Janet C.E. (1999). "The Directionality of Emphasis Spread in Arabic." *Linguistic Inquiry*30:2. 289-300.

Watson, Janet C.E. (2002).*The Phonology and Morphology of Arabic.* Oxford University Press, London.

Xu, Yi. (2005), "Speech melody as articulatorily implemented communicative functions." *Speech Communication* 46: 220-251.

Yaghi, J. and Yagi, S.M. (2004). "Systematic Verb Stem Generation for Arabic." Workshop on Computational Approaches to Arabic Script-based Languages. University of Geneva. 23-30.

Youssef, Islam M. (2006). "Vocalic Feature Assimilation in Cairene Arabic and Buchan Scots: Contrast and Phonological Activity in the Parallel Structures Model of Feature Geometry." Master's Thesis, University of Tromsø.

Appendix

IPA to SAMPA

IPA	SAMPA
m	m
n	n
b	b
d	d
t	t
dˤ	d`
tˤ	t`
g	g
k	k
q	q
ʔ	?
v	v
f	f
z	z
s	s
zˤ	z`
sˤ	s`
ʃ	S
ʏ	R
x	x
ʕ	?`
ħ	X\
h	h
r	r
l	l
j	j
w	w
i	i
u	u
a	a
ɑˤ	A`
i:	i:
u:	u:
a:	a:
ɑ:ˤ	A:`

Diphone List

C – C diphones

C1	C2	C1	C2	C1	C2	C1	C2
m	m	Z	l	f	l`	R`	k
m	n	Z	j	f	j`	R`	q
m	b	Z	w	f	w`	R`	?
m	d	S	m	z	m`	R`	v
m	t	S	n	z	n`	R`	f
m	d`	S	b	z	b`	R`	z
m	t`	S	d	z	g`	R`	s
m	g	S	t	z	k`	R`	z`
m	k	S	d`	z	v`	R`	s`
m	q	S	t`	z	f`	R`	Z
m	?	S	g	z	Z`	R`	S
m	v	S	k	z	S`	R`	R
m	f	S	q	z	R`	R`	x
m	z	S	?	z	x`	R`	?`
m	s	S	v	z	X\`	R`	X\
m	z`	S	f	z	h`	R`	h
m	s`	S	z	z	r`	R`	r
m	Z	S	s	z	l`	R`	l
m	S	S	z`	z	j`	R`	j
m	R	S	s`	z	w`	R`	w
m	x	S	Z	s	m`	x`	m
m	?`	S	S	s	n`	x`	n
m	X\	S	R	s	b`	x`	b
m	h	S	x	s	g`	x`	d
m	r	S	?`	s	k`	x`	t
m	l	S	X\	s	v`	x`	d`
m	j	S	h	s	f`	x`	t`
m	w	S	r	s	Z`	x`	g
n	m	S	l	s	S`	x`	k
n	n	S	j	s	R`	x`	q
n	b	S	w	s	x`	x`	?
n	d	R	m	s	X\`	x`	v
n	t	R	n	s	h`	x`	f
n	d`	R	b	s	r`	x`	z
n	t`	R	d	s	l`	x`	s
n	g	R	t	s	j`	x`	z`
n	k	R	d`	s	w`	x`	s`
n	q	R	t`	z`	m`	x`	Z
n	?	R	g	z`	n`	x`	S
n	v	R	k	z`	b`	x`	R
n	f	R	q	z`	g`	x`	x
n	z	R	?	z`	k`	x`	?`

n	s	R	v	z`	v`	x`	X\
n	z`	R	f	z`	f`	x`	h
n	s`	R	z	z`	Z`	x`	r
n	Z	R	s	z`	S`	x`	l
n	S	R	z`	z`	R`	x`	j
n	R	R	s`	z`	x`	x`	w
n	x	R	Z	z`	X\`	X\`	m
n	?`	R	S	z`	h`	X\`	n
n	X\	R	R	z`	r`	X\`	b
n	h	R	x	z`	l`	X\`	d
n	r	R	?`	z`	j`	X\`	t
n	l	R	X\	z`	w`	X\`	d`
n	j	R	h	s`	m`	X\`	t`
n	w	R	r	s`	n`	X\`	g
b	m	R	l	s`	b`	X\`	k
b	n	R	j	s`	g`	X\`	q
b	b	R	w	s`	k`	X\`	?
b	d	x	m	s`	v`	X\`	v
b	t	x	n	s`	f`	X\`	f
b	d`	x	b	s`	Z`	X\`	z
b	t`	x	d	s`	S`	X\`	s
b	g	x	t	s`	R`	X\`	z`
b	k	x	d`	s`	x`	X\`	s`
b	q	x	t`	s`	X\`	X\`	Z
b	?	x	g	s`	h`	X\`	S
b	v	x	k	s`	r`	X\`	R
b	f	x	q	s`	l`	X\`	x
b	z	x	?	s`	j`	X\`	?`
b	s	x	v	s`	w`	X\`	X\
b	z`	x	f	Z	m`	X\`	h
b	s`	x	z	Z	n`	X\`	r
b	Z	x	s	Z	b`	X\`	l
b	S	x	z`	Z	g`	X\`	j
b	R	x	s`	Z	k`	X\`	w
b	x	x	Z	Z	v`	h`	m
b	?`	x	S	Z	f`	h`	n
b	X\	x	R	Z	Z`	h`	b
b	h	x	x	Z	S`	h`	d
b	r	x	?`	Z	R`	h`	t
b	l	x	X\	Z	x`	h`	d`
b	j	x	h	Z	X\`	h`	t`
b	w	x	r	Z	h`	h`	g
d	m	x	l	Z	r`	h`	k
d	n	x	j	Z	l`	h`	q

d	b	x	w	Z	j`	h`	?
d	d	?`	m	Z	w`	h`	v
d	t	?`	n	S	m`	h`	f
d	d`	?`	b	S	n`	h`	z
d	t`	?`	d	S	b`	h`	s
d	g	?`	t	S	g`	h`	z`
d	k	?`	d`	S	k`	h`	s`
d	q	?`	t`	S	v`	h`	Z
d	?	?`	g	S	f`	h`	S
d	v	?`	k	S	Z`	h`	R
d	f	?`	q	S	S`	h`	x
d	z	?`	?	S	R`	h`	?`
d	s	?`	v	S	x`	h`	X\
d	z`	?`	f	S	X\`	h`	h
d	s`	?`	z	S	h`	h`	r
d	Z	?`	s	S	r`	h`	l
d	S	?`	z`	S	l`	h`	j
d	R	?`	s`	S	j`	h`	w
d	x	?`	Z	S	w`	r`	m
d	?`	?`	S	R	m`	r`	n
d	X\	?`	R	R	n`	r`	b
d	h	?`	x	R	b`	r`	d
d	r	?`	?`	R	g`	r`	t
d	l	?`	X\	R	k`	r`	d`
d	j	?`	h	R	v`	r`	t`
d	w	?`	r	R	f`	r`	g
t	m	?`	l	R	Z`	r`	k
t	n	?`	j	R	S`	r`	q
t	b	?`	w	R	R`	r`	?
t	d	X\	m	R	x`	r`	v
t	t	X\	n	R	X\`	r`	f
t	d`	X\	b	R	h`	r`	z
t	t`	X\	d	R	r`	r`	s
t	g	X\	t	R	l`	r`	z`
t	k	X\	d`	R	j`	r`	s`
t	q	X\	t`	R	w`	r`	Z
t	?	X\	g	?`	m`	r`	S
t	v	X\	k	?`	n`	r`	R
t	f	X\	q	?`	b`	r`	x
t	z	X\	?	?`	g`	r`	?`
t	s	X\	v	?`	k`	r`	X\
t	z`	X\	f	?`	v`	r`	h
t	s`	X\	z	?`	f`	r`	r
t	Z	X\	s	?`	Z`	r`	l
t	S	X\	z`	?`	S`	r`	j
t	R	X\	s`	?`	R`	r`	w

t	x	X\	Z	?`	x`	l`	m
t	?`	X\	S	?`	X\`	l`	n
t	X\	X\	R	?`	h`	l`	b
t	h	X\	x	?`	r`	l`	d
t	r	X\	?`	?`	l`	l`	t
t	l	X\	X\	?`	j`	l`	d`
t	j	X\	h	?`	w`	l`	t`
t	w	X\	r	X\	m`	l`	g
d`	m	X\	l	X\	n`	l`	k
d`	n	X\	j	X\	b`	l`	q
d`	b	X\	w	X\	g`	l`	?
d`	d	h	m	X\	k`	l`	v
d`	t	h	n	X\	v`	l`	f
d`	d`	h	b	X\	f`	l`	z
d`	t`	h	d	X\	Z`	l`	s
d`	g	h	t	X\	S`	l`	z`
d`	k	h	d`	X\	R`	l`	s`
d`	q	h	t`	X\	x`	l`	Z
d`	?	h	g	X\	X\`	l`	S
d`	v	h	k	X\	h`	l`	R
d`	f	h	q	X\	r`	l`	x
d`	z	h	?	X\	l`	l`	?`
d`	s	h	v	X\	j`	l`	X\
d`	z`	h	f	X\	w`	l`	h
d`	s`	h	z	h	m`	l`	r
d`	Z	h	s	h	n`	l`	l
d`	S	h	z`	h	b`	l`	j
d`	R	h	s`	h	g`	l`	w
d`	x	h	Z	h	k`	j`	m
d`	?`	h	S	h	v`	j`	n
d`	X\	h	R	h	f`	j`	b
d`	h	h	x	h	Z`	j`	d
d`	r	h	?`	h	S`	j`	t
d`	l	h	X\	h	R`	j`	d`
d`	j	h	h	h	x`	j`	t`
d`	w	h	r	h	X\`	j`	g
t`	m	h	l	h	h`	j`	k
t`	n	h	j	h	r`	j`	q
t`	b	h	w	h	l`	j`	?
t`	d	r	m	h	j`	j`	v
t`	t	r	n	h	w`	j`	f
t`	d`	r	b	r	m`	j`	z
t`	t`	r	d	r	n`	j`	s
t`	g	r	t	r	b`	j`	z`
t`	k	r	d`	r	g`	j`	s`

t`	q	r	t`	r	k`	j`	Z
t`	?	r	g	r	v`	j`	S
t`	v	r	k	r	f`	j`	R
t`	f	r	q	r	Z`	j`	x
t`	z	r	?	r	S`	j`	?`
t`	s	r	v	r	R`	j`	X\
t`	z`	r	f	r	x`	j`	h
t`	s`	r	z	r	X\`	j`	r
t`	Z	r	s	r	h`	j`	l
t`	S	r	z`	r	r`	j`	j
t`	R	r	s`	r	l`	j`	w
t`	x	r	Z	r	j`	w`	m
t`	?`	r	S	r	w`	w`	n
t`	X\	r	R	l	m`	w`	b
t`	h	r	x	l	n`	w`	d
t`	r	r	?`	l	b`	w`	t
t`	l	r	X\	l	g`	w`	d`
t`	j	r	h	l	k`	w`	t`
t`	w	r	r	l	v`	w`	g
g	m	r	l	l	f`	w`	k
g	n	r	j	l	Z`	w`	q
g	b	r	w	l	S`	w`	?
g	d	l	m	l	R`	w`	v
g	t	l	n	l	x`	w`	f
g	d`	l	b	l	X\`	w`	z
g	t`	l	d	l	h`	w`	s
g	g	l	t	l	r`	w`	z`
g	k	l	d`	l	l`	w`	s`
g	q	l	t`	l	j`	w`	Z
g	?	l	g	l	w`	w`	S
g	v	l	k	j	m`	w`	R
g	f	l	q	j	n`	w`	x
g	z	l	?	j	b`	w`	?`
g	s	l	v	j	g`	w`	X\
g	z`	l	f	j	k`	w`	h
g	s`	l	z	j	v`	w`	r
g	Z	l	s	j	f`	w`	l
g	S	l	z`	j	Z`	w`	j
g	R	l	s`	j	S`	w`	w
g	x	l	Z	j	R`	m`	m`
g	?`	l	S	j	x`	m`	n`
g	X\	l	R	j	X\`	m`	b`
g	h	l	x	j	h`	m`	g`
g	r	l	?`	j	r`	m`	k`
g	l	l	X\	j	l`	m`	v`
g	j	l	h	j	j`	m`	f`

g	w	l	r	j	w`	m`	Z`
k	m	l	l	w	m`	m`	S`
k	n	l	j	w	n`	m`	R`
k	b	l	w	w	b`	m`	x`
k	d	j	m	w	g`	m`	X\`
k	t	j	n	w	k`	m`	h`
k	d`	j	b	w	v`	m`	r`
k	t`	j	d	w	f`	m`	l`
k	g	j	t	w	Z`	m`	j`
k	k	j	d`	w	S`	m`	w`
k	q	j	t`	w	R`	n`	m`
k	?	j	g	w	x`	n`	n`
k	v	j	k	w	X\`	n`	b`
k	f	j	q	w	h`	n`	g`
k	z	j	?	w	r`	n`	k`
k	s	j	v	w	l`	n`	v`
k	z`	j	f	w	j`	n`	f`
k	s`	j	z	w	w`	n`	Z`
k	Z	j	s	m`	m	n`	S`
k	S	j	z`	m`	n	n`	R`
k	R	j	s`	m`	b	n`	x`
k	x	j	Z	m`	d	n`	X\`
k	?`	j	S	m`	t	n`	h`
k	X\	j	R	m`	d`	n`	r`
k	h	j	x	m`	t`	n`	l`
k	r	j	?`	m`	g	n`	j`
k	l	j	X\	m`	k	n`	w`
k	j	j	h	m`	q	b`	m`
k	w	j	r	m`	?	b`	n`
q	m	j	l	m`	v	b`	b`
q	n	j	j	m`	f	b`	g`
q	b	j	w	m`	z	b`	k`
q	d	w	m	m`	s	b`	v`
q	t	w	n	m`	z`	b`	f`
q	d`	w	b	m`	s`	b`	Z`
q	t`	w	d	m`	Z	b`	S`
q	g	w	t	m`	S	b`	R`
q	k	w	d`	m`	R	b`	x`
q	q	w	t`	m`	x	b`	X\`
q	?	w	g	m`	?`	b`	h`
q	v	w	k	m`	X\	b`	r`
q	f	w	q	m`	h	b`	l`
q	z	w	?	m`	r	b`	j`
q	s	w	v	m`	l	b`	w`
q	z`	w	f	m`	j	g`	m`
q	s`	w	z	m`	w	g`	n`

q	Z	w	s	n`	m	g`	b`
q	S	w	z`	n`	n	g`	g`
q	R	w	s`	n`	b	g`	k`
q	x	w	Z	n`	d	g`	v`
q	?`	w	S	n`	t	g`	f`
q	X\	w	R	n`	d`	g`	Z`
q	h	w	x	n`	t`	g`	S`
q	r	w	?`	n`	g	g`	R`
q	l	w	X\	n`	k	g`	x`
q	j	w	h	n`	q	g`	X\`
q	w	w	r	n`	?	g`	h`
?	m	w	l	n`	v	g`	r`
?	n	w	j	n`	f	g`	l`
?	b	w	w	n`	z	g`	j`
?	d	m	m`	n`	s	g`	w`
?	t	m	n`	n`	z`	k`	m`
?	d`	m	b`	n`	s`	k`	n`
?	t`	m	g`	n`	Z	k`	b`
?	g	m	k`	n`	S	k`	g`
?	k	m	v`	n`	R	k`	k`
?	q	m	f`	n`	x	k`	v`
?	?	m	Z`	n`	?`	k`	f`
?	v	m	S`	n`	X\	k`	Z`
?	f	m	R`	n`	h	k`	S`
?	z	m	x`	n`	r	k`	R`
?	s	m	X\`	n`	l	k`	x`
?	z`	m	h`	n`	j	k`	X\`
?	s`	m	r`	n`	w	k`	h`
?	Z	m	l`	b`	m	k`	r`
?	S	m	j`	b`	n	k`	l`
?	R	m	w`	b`	b	k`	j`
?	x	n	m`	b`	d	k`	w`
?	?`	n	n`	b`	t	v`	m`
?	X\	n	b`	b`	d`	v`	n`
?	h	n	g`	b`	t`	v`	b`
?	r	n	k`	b`	g	v`	g`
?	l	n	v`	b`	k	v`	k`
?	j	n	f`	b`	q	v`	v`
?	w	n	Z`	b`	?	v`	f`
v	m	n	S`	b`	v	v`	Z`
v	n	n	R`	b`	f	v`	S`
v	b	n	x`	b`	z	v`	R`
v	d	n	X\`	b`	s	v`	x`
v	t	n	h`	b`	z`	v`	X\`
v	d`	n	r`	b`	s`	v`	h`
v	t`	n	l`	b`	Z	v`	r`

v	g	n	j`	b`	S	v`	l`
v	k	n	w`	b`	R	v`	j`
v	q	b	m`	b`	x	v`	w`
v	?	b	n`	b`	?`	f`	m`
v	v	b	b`	b`	X\	f`	n`
v	f	b	g`	b`	h	f`	b`
v	z	b	k`	b`	r	f`	g`
v	s	b	v`	b`	l	f`	k`
v	z`	b	f`	b`	j	f`	v`
v	s`	b	Z`	b`	w	f`	f`
v	Z	b	S`	g`	m	f`	Z`
v	S	b	R`	g`	n	f`	S`
v	R	b	x`	g`	b	f`	R`
v	x	b	X\`	g`	d	f`	x`
v	?`	b	h`	g`	t	f`	X\`
v	X\	b	r`	g`	d`	f`	h`
v	h	b	l`	g`	t`	f`	r`
v	r	b	j`	g`	g	f`	l`
v	l	b	w`	g`	k	f`	j`
v	j	d	m`	g`	q	f`	w`
v	w	d	n`	g`	?	Z`	m`
f	m	d	b`	g`	v	Z`	n`
f	n	d	g`	g`	f	Z`	b`
f	b	d	k`	g`	z	Z`	g`
f	d	d	v`	g`	s	Z`	k`
f	t	d	f`	g`	z`	Z`	v`
f	d`	d	Z`	g`	s`	Z`	f`
f	t`	d	S`	g`	Z	Z`	Z`
f	g	d	R`	g`	S	Z`	S`
f	k	d	x`	g`	R	Z`	R`
f	q	d	X\`	g`	x	Z`	x`
f	?	d	h`	g`	?`	Z`	X\`
f	v	d	r`	g`	X\	Z`	h`
f	f	d	l`	g`	h	Z`	r`
f	z	d	j`	g`	r	Z`	l`
f	s	d	w`	g`	l	Z`	j`
f	z`	t	m`	g`	j	Z`	w`
f	s`	t	n`	g`	w	S`	m`
f	Z	t	b`	k`	m	S`	n`
f	S	t	g`	k`	n	S`	b`
f	R	t	k`	k`	b	S`	g`
f	x	t	v`	k`	d	S`	k`
f	?`	t	f`	k`	t	S`	v`
f	X\	t	Z`	k`	d`	S`	f`
f	h	t	S`	k`	t`	S`	Z`
f	r	t	R`	k`	g	S`	S`

f	l	t	x`	k`	k	S`	R`
f	j	t	X\`	k`	q	S`	x`
f	w	t	h`	k`	?	S`	X\`
z	m	t	r`	k`	v	S`	h`
z	n	t	l`	k`	f	S`	r`
z	b	t	j`	k`	z	S`	l`
z	d	t	w`	k`	s	S`	j`
z	t	d`	m`	k`	z`	S`	w`
z	d`	d`	n`	k`	s`	R`	m`
z	t`	d`	b`	k`	Z	R`	n`
z	g	d`	g`	k`	S	R`	b`
z	k	d`	k`	k`	R	R`	g`
z	q	d`	v`	k`	x	R`	k`
z	?	d`	f`	k`	?`	R`	v`
z	v	d`	Z`	k`	X\	R`	f`
z	f	d`	S`	k`	h	R`	Z`
z	z	d`	R`	k`	r	R`	S`
z	s	d`	x`	k`	l	R`	R`
z	z`	d`	X\`	k`	j	R`	x`
z	s`	d`	h`	k`	w	R`	X\`
z	Z	d`	r`	v`	m	R`	h`
z	S	d`	l`	v`	n	R`	r`
z	R	d`	j`	v`	b	R`	l`
z	x	d`	w`	v`	d	R`	j`
z	?`	t`	m`	v`	t	R`	w`
z	X\	t`	n`	v`	d`	x`	m`
z	h	t`	b`	v`	t`	x`	n`
z	r	t`	g`	v`	g	x`	b`
z	l	t`	k`	v`	k	x`	g`
z	j	t`	v`	v`	q	x`	k`
z	w	t`	f`	v`	?	x`	v`
s	m	t`	Z`	v`	v	x`	f`
s	n	t`	S`	v`	f	x`	Z`
s	b	t`	R`	v`	z	x`	S`
s	d	t`	x`	v`	s	x`	R`
s	t	t`	X\`	v`	z`	x`	x`
s	d`	t`	h`	v`	s`	x`	X\`
s	t`	t`	r`	v`	Z	x`	h`
s	g	t`	l`	v`	S	x`	r`
s	k	t`	j`	v`	R	x`	l`
s	q	t`	w`	v`	x	x`	j`
s	?	g	m`	v`	?`	x`	w`
s	v	g	n`	v`	X\	X\`	m`
s	f	g	b`	v`	h	X\`	n`
s	z	g	g`	v`	r	X\`	b`
s	s	g	k`	v`	l	X\`	g`

s	z`	g	v`	v`	j	X\`	k`
s	s`	g	f`	v`	w	X\`	v`
s	Z	g	Z`	f`	m	X\`	f`
s	S	g	S`	f`	n	X\`	Z`
s	R	g	R`	f`	b	X\`	S`
s	x	g	x`	f`	d	X\`	R`
s	?`	g	X\`	f`	t	X\`	x`
s	X\	g	h`	f`	d`	X\`	X\`
s	h	g	r`	f`	t`	X\`	h`
s	r	g	l`	f`	g	X\`	r`
s	l	g	j`	f`	k	X\`	l`
s	j	g	w`	f`	q	X\`	j`
s	w	k	m`	f`	?	X\`	w`
z`	m	k	n`	f`	v	h`	m`
z`	n	k	b`	f`	f	h`	n`
z`	b	k	g`	f`	z	h`	b`
z`	d	k	k`	f`	s	h`	g`
z`	t	k	v`	f`	z`	h`	k`
z`	d`	k	f`	f`	s`	h`	v`
z`	t`	k	Z`	f`	Z	h`	f`
z`	g	k	S`	f`	S	h`	Z`
z`	k	k	R`	f`	R	h`	S`
z`	q	k	x`	f`	x	h`	R`
z`	?	k	X\`	f`	?`	h`	x`
z`	v	k	h`	f`	X\	h`	X\`
z`	f	k	r`	f`	h	h`	h`
z`	z	k	l`	f`	r	h`	r`
z`	s	k	j`	f`	l	h`	l`
z`	z`	k	w`	f`	j	h`	j`
z`	s`	q	m`	f`	w	h`	w`
z`	Z	q	n`	Z`	m	r`	m`
z`	S	q	b`	Z`	n	r`	n`
z`	R	q	g`	Z`	b	r`	b`
z`	x	q	k`	Z`	d	r`	g`
z`	?`	q	v`	Z`	t	r`	k`
z`	X\	q	f`	Z`	d`	r`	v`
z`	h	q	Z`	Z`	t`	r`	f`
z`	r	q	S`	Z`	g	r`	Z`
z`	l	q	R`	Z`	k	r`	S`
z`	j	q	x`	Z`	q	r`	R`
z`	w	q	X\`	Z`	?	r`	x`
s`	m	q	h`	Z`	v	r`	X\`
s`	n	q	r`	Z`	f	r`	h`
s`	b	q	l`	Z`	z	r`	r`
s`	d	q	j`	Z`	s	r`	l`
s`	t	q	w`	Z`	z`	r`	j`

s`	d`	?	m`	Z`	s`	r`	w`
s`	t`	?	n`	Z`	Z	l`	m`
s`	g	?	b`	Z`	S	l`	n`
s`	k	?	g`	Z`	R	l`	b`
s`	q	?	k`	Z`	x	l`	g`
s`	?	?	v`	Z`	?`	l`	k`
s`	v	?	f`	Z`	X\	l`	v`
s`	f	?	Z`	Z`	h	l`	f`
s`	z	?	S`	Z`	r	l`	Z`
s`	s	?	R`	Z`	l	l`	S`
s`	z`	?	x`	Z`	j	l`	R`
s`	s`	?	X\`	Z`	w	l`	x`
s`	Z	?	h`	S`	m	l`	X\`
s`	S	?	r`	S`	n	l`	h`
s`	R	?	l`	S`	b	l`	r`
s`	x	?	j`	S`	d	l`	l`
s`	?`	?	w`	S`	t	l`	j`
s`	X\	v	m`	S`	d`	l`	w`
s`	h	v	n`	S`	t`	j`	m`
s`	r	v	b`	S`	g	j`	n`
s`	l	v	g`	S`	k	j`	b`
s`	j	v	k`	S`	q	j`	g`
s`	w	v	v`	S`	?	j`	k`
Z	m	v	f`	S`	v	j`	v`
Z	n	v	Z`	S`	f	j`	f`
Z	b	v	S`	S`	z	j`	Z`
Z	d	v	R`	S`	s	j`	S`
Z	t	v	x`	S`	z`	j`	R`
Z	d`	v	X\`	S`	s`	j`	x`
Z	t`	v	h`	S`	Z	j`	X\`
Z	g	v	r`	S`	S	j`	h`
Z	k	v	l`	S`	R	j`	r`
Z	q	v	j`	S`	x	j`	l`
Z	?	v	w`	S`	?`	j`	j`
Z	v	f	m`	S`	X\	j`	w`
Z	f	f	n`	S`	h	w`	m`
Z	z	f	b`	S`	r	w`	n`
Z	s	f	g`	S`	l	w`	b`
Z	z`	f	k`	S`	j	w`	g`
Z	s`	f	v`	S`	w	w`	k`
Z	Z	f	f`	R`	m	w`	v`
Z	S	f	Z`	R`	n	w`	f`
Z	R	f	S`	R`	b	w`	Z`
Z	x	f	R`	R`	d	w`	S`
Z	?`	f	x`	R`	t	w`	R`
Z	X\	f	X\`	R`	d`	w`	x`

Z	h	f	h`	R`	t`	w`	X\`
Z	r	f	r`	R`	g	w`	h`
w`	w`	w`	j`	w`	l`	w`	r`

C – V diphones

C	V	C	V	C	V	C	V
m	i	m	i:	m	aw	m`	i`
n	i	n	i:	n	aw	n`	i`
b	i	b	i:	b	aw	b`	i`
d	i	d	i:	d	aw	g`	i`
t	i	t	i:	t	aw	k`	i`
d`	i	d`	i:	g	aw	v`	i`
t`	i	t`	i:	k	aw	f`	i`
g	i	g	i:	q	aw	Z`	i`
k	i	k	i:	?	aw	S`	i`
q	i	q	i:	v	aw	R`	i`
?	i	?	i:	f	aw	x`	i`
v	i	v	i:	z	aw	X\`	i`
f	i	f	i:	s	aw	h`	i`
z	i	z	i:	Z	aw	r`	i`
s	i	s	i:	S	aw	l`	i`
z`	i	z`	i:	R	aw	j`	i`
s`	i	s`	i:	x	aw	w`	i`
Z	i	Z	i:	X\	aw	m`	u`
S	i	S	i:	h	aw	n`	u`
R	i	R	i:	r	aw	b`	u`
x	i	x	i:	l	aw	g`	u`
?`	i	?`	i:	j	aw	k`	u`
X\	i	X\	i:	w	aw	v`	u`
h	i	h	i:	d`	A`	f`	u`
r	i	r	i:	t`	A`	Z`	u`
l	i	l	i:	z`	A`	S`	u`
j	i	j	i:	s`	A`	R`	u`
w	i	w	i:	?`	A`	x`	u`
m	i	m	u:	d`	A:`	X\`	u`
n	i	n	u:	t`	A:`	h`	u`
b	i	b	u:	z`	A:`	r`	u`
d	i	d	u:	s`	A:`	l`	u`
t	i	t	u:	?`	A:`	j`	u`
g	i	g	u:	d`	i`	w`	u`
k	i	k	u:	t`	i`	m`	aj`

q	i	q	u:	z`	i`	n`	aj`
?	i	?	u:	s`	i`	b`	aj`
v	i	v	u:	?`	i`	g`	aj`
f	i	f	u:	d`	u`	k`	aj`
z	i	z	u:	t`	u`	v`	aj`
s	i	s	u:	z`	u`	f`	aj`
Z	i	Z	u:	s`	u`	Z`	aj`
S	i	S	u:	?`	u`	S`	aj`
R	i	R	u:	d`	aj`	R`	aj`
x	i	x	u:	t`	aj`	x`	aj`
X\	i	X\	u:	z`	aj`	X\`	aj`
h	i	h	u:	s`	aj`	h`	aj`
r	i	r	u:	?`	aj`	r`	aj`
l	i	l	u:	d`	aw`	l`	aj`
j	i	j	u:	t`	aw`	j`	aj`
w	i	w	u:	z`	aw`	w`	aj`
m	u	m	a:	s`	aw`	m`	aw`
n	u	n	a:	?`	aw`	n`	aw`
b	u	b	a:	d`	i:`	b`	aw`
d	u	d	a:	t`	i:`	g`	aw`
t	u	t	a:	z`	i:`	k`	aw`
g	u	g	a:	s`	i:`	v`	aw`
k	u	k	a:	?`	i:`	f`	aw`
q	u	q	a:	d`	u:`	Z`	aw`
?	u	?	a:	t`	u:`	S`	aw`
v	u	v	a:	z`	u:`	R`	aw`
f	u	f	a:	s`	u:`	x`	aw`
z	u	z	a:	?`	u:`	X\`	aw`
s	u	s	a:	m`	A`	h`	aw`
Z	u	Z	a:	n`	A`	r`	aw`
S	u	S	a:	b`	A`	l`	aw`
R	u	R	a:	g`	A`	j`	aw`
x	u	x	a:	k`	A`	w`	aw`
X\	u	X\	a:	v`	A`	m`	i:`
h	u	h	a:	f`	A`	n`	i:`
r	u	r	a:	Z`	A`	b`	i:`
l	u	l	a:	S`	A`	g`	i:`
j	u	j	a:	R`	A`	k`	i:`
w	u	w	a:	x`	A`	v`	i:`
m	a	m	aj	X\`	A`	f`	i:`
n	a	n	aj	h`	A`	Z`	i:`
b	a	b	aj	r`	A`	S`	i:`
d	a	d	aj	l`	A`	R`	i:`

t	a	t	aj	j`	A`	x`	i:`
g	a	g	aj	w`	A`	X\`	i:`
k	a	k	aj	m`	A:`	h`	i:`
q	a	q	aj	n`	A:`	r`	i:`
?	a	?	aj	b`	A:`	l`	i:`
v	a	v	aj	g`	A:`	j`	i:`
f	a	f	aj	k`	A:`	w`	i:`
z	a	z	aj	v`	A:`	m`	u:`
s	a	s	aj	f`	A:`	n`	u:`
Z	a	Z	aj	Z`	A:`	b`	u:`
S	a	S	aj	S`	A:`	g`	u:`
R	a	R	aj	R`	A:`	k`	u:`
x	a	x	aj	x`	A:`	v`	u:`
X\	a	X\	aj	X\`	A:`	f`	u:`
h	a	h	aj	h`	A:`	Z`	u:`
r	a	r	aj	r`	A:`	S`	u:`
l	a	l	aj	l`	A:`	R`	u:`
j	a	j	aj	j`	A:`	x`	u:`
w	a	w	aj	w`	A:`	X\`	u:`
r`	u:`	l`	u:`	j`	u:`	h`	u:`
w`	u:`						

V – C diphones

V	C	V	C	V	C	V	C
i	m	i:	h`	i:`	v`	i:`	m
i	n	i:	r`	i:`	f`	i:`	n
i	b	i:	l`	i:`	Z`	i:`	b
i	d	i:	j`	i:`	S`	i:`	d
i	t	i:	w`	i:`	R`	i:`	t
i	g	u	m`	i:`	x`	i:`	g
i	k	u	n`	i:`	X\`	i:`	k
i	q	u	b`	i:`	h`	i:`	q
i	?	u	g`	i:`	r`	i:`	?
i	v	u	k`	i:`	l`	i:`	v
i	f	u	v`	i:`	j`	i:`	f
i	z	u	f`	i:`	w`	i:`	z
i	s	u	Z`	u`	m`	i:`	s
i	Z	u	S`	u`	n`	i:`	Z
i	S	u	R`	u`	b`	i:`	S

i	R	u	x`	u`	g`	i:`	R
i	x	u	X\`	u`	k`	i:`	x
i	X\	u	h`	u`	v`	i:`	X\
i	h	u	r`	u`	f`	i:`	h
i	r	u	l`	u`	Z`	i:`	r
i	l	u	j`	u`	S`	i:`	l
i	j	u	w`	u`	R`	i:`	j
i	w	u:	m`	u`	x`	i:`	w
i:	m	u:	n`	u`	X\`	u`	m
i:	n	u:	b`	u`	h`	u`	n
i:	b	u:	g`	u`	r`	u`	b
i:	d	u:	k`	u`	l`	u`	d
i:	t	u:	v`	u`	j`	u`	t
i:	g	u:	f`	u`	w`	u`	g
i:	k	u:	Z`	u:`	m`	u`	k
i:	q	u:	S`	u:`	n`	u`	q
i:	?	u:	R`	u:`	b`	u`	?
i:	v	u:	x`	u:`	g`	u`	v
i:	f	u:	X\`	u:`	k`	u`	f
i:	z	u:	h`	u:`	v`	u`	z
i:	s	u:	r`	u:`	f`	u`	s
i:	Z	u:	l`	u:`	Z`	u`	Z
i:	S	u:	j`	u:`	S`	u`	S
i:	R	u:	w`	u:`	R`	u`	R
i:	x	a	m`	u:`	x`	u`	x
i:	X\	a	n`	u:`	X\`	u`	X\
i:	h	a	b`	u:`	h`	u`	h
i:	r	a	g`	u:`	r`	u`	r
i:	l	a	k`	u:`	l`	u`	l
i:	j	a	v`	u:`	j`	u`	j
i:	w	a	f`	u:`	w`	u`	w
u	m	a	Z`	aj`	m`	u:`	m
u	n	a	S`	aj`	n`	u:`	n
u	b	a	R`	aj`	b`	u:`	b
u	d	a	x`	aj`	g`	u:`	d
u	t	a	X\`	aj`	k`	u:`	t
u	g	a	h`	aj`	v`	u:`	g
u	k	a	r`	aj`	f`	u:`	k
u	q	a	l`	aj`	Z`	u:`	q
u	?	a	j`	aj`	S`	u:`	?
u	v	a	w`	aj`	R`	u:`	v
u	f	a:	m`	aj`	x`	u:`	f
u	z	a:	n`	aj`	X\`	u:`	z
u	s	a:	b`	aj`	h`	u:`	s
u	Z	a:	g`	aj`	r`	u:`	Z
u	S	a:	k`	aj`	l`	u:`	S

u	R	a:	v`	aj`	j`	u:`	R
u	x	a:	f`	aj`	w`	u:`	x
u	X\	a:	Z`	aw`	m`	u:`	X\
u	h	a:	S`	aw`	n`	u:`	h
u	r	a:	R`	aw`	b`	u:`	r
u	l	a:	x`	aw`	g`	u:`	l
u	j	a:	X\`	aw`	k`	u:`	j
u	w	a:	h`	aw`	v`	u:`	w
u:	m	a:	r`	aw`	f`	aj`	m
u:	n	a:	l`	aw`	Z`	aj`	n
u:	b	a:	j`	aw`	S`	aj`	b
u:	d	a:	w`	aw`	R`	aj`	d
u:	t	aj	m`	aw`	x`	aj`	t
u:	g	aj	n`	aw`	X\`	aj`	g
u:	k	aj	b`	aw`	h`	aj`	k
u:	q	aj	g`	aw`	r`	aj`	q
u:	?	aj	k`	aw`	l`	aj`	?
u:	v	aj	v`	aw`	j`	aj`	v
u:	f	aj	f`	aw`	w`	aj`	f
u:	z	aj	Z`	i`	d`	aj`	z
u:	s	aj	S`	i`	t`	aj`	s
u:	Z	aj	R`	i`	z`	aj`	Z
u:	S	aj	x`	i`	s`	aj`	S
u:	R	aj	X\`	i`	?`	aj`	R
u:	x	aj	h`	i:`	d`	aj`	x
u:	X\	aj	r`	i:`	t`	aj`	X\
u:	h	aj	l`	i:`	z`	aj`	h
u:	r	aj	j`	i:`	s`	aj`	r
u:	l	aj	w`	i:`	?`	aj`	l
u:	j	aw	m`	u`	d`	aj`	j
u:	w	aw	n`	u`	t`	aj`	w
a	m	aw	b`	u`	z`	aw`	m
a	n	aw	g`	u`	s`	aw`	n
a	b	aw	k`	u`	?`	aw`	b
a	d	aw	v`	u:`	d`	aw`	d
a	t	aw	f`	u:`	t`	aw`	t
a	g	aw	Z`	u:`	z`	aw`	g
a	k	aw	S`	u:`	s`	aw`	k
a	q	aw	R`	u:`	?`	aw`	q
a	?	aw	x`	aj`	d`	aw`	?
a	v	aw	X\`	aj`	t`	aw`	v
a	f	aw	h`	aj`	z`	aw`	f
a	z	aw	r`	aj`	s`	aw`	z
a	s	aw	l`	aj`	?`	aw`	s
a	Z	aw	j`	aw`	d`	aw`	Z
a	S	aw	w`	aw`	t`	aw`	S

a	R	i	d`	aw`	z`	aw`	R
a	x	i	t`	aw`	s`	aw`	x
a	X\	i	z`	aw`	?`	aw`	X\
a	h	i	s`	A`	m	aw`	h
a	r	i	?`	A`	n	aw`	r
a	l	i:	d`	A`	b	aw`	l
a	j	i:	t`	A`	d	aw`	j
a	w	i:	z`	A`	t	aw`	w
a:	m	i:	s`	A`	g	i`	m`
a:	n	i:	?`	A`	k	i`	n`
a:	b	u	d`	A`	q	i`	b`
a:	d	u	t`	A`	?	i`	g`
a:	t	u	z`	A`	v	i`	k`
a:	g	u	s`	A`	f	i`	v`
a:	k	u	?`	A`	z	i`	f`
a:	q	u:	d`	A`	s	i`	Z`
a:	?	u:	t`	A`	Z	i`	S`
a:	v	u:	z`	A`	S	i`	R`
a:	f	u:	s`	A`	R	i`	x`
a:	z	u:	?`	A`	x	i`	X\`
a:	s	a	d`	A`	X\	i`	h`
a:	Z	a	t`	A`	h	i`	r`
a:	S	a	z`	A`	r	i`	l`
a:	R	a	s`	A`	l	i`	j`
a:	x	a	?`	A`	j	i`	w`
a:	X\	a:	d`	A`	w	i:`	m`
a:	h	a:	t`	A:`	m	i:`	n`
a:	r	a:	z`	A:`	n	i:`	b`
a:	l	a:	s`	A:`	b	i:`	g`
a:	j	a:	?`	A:`	d	i:`	k`
a:	w	aj	d`	A:`	t	i:`	v`
aj	m	aj	t`	A:`	g	i:`	f`
aj	n	aj	z`	A:`	k	i:`	Z`
aj	b	aj	s`	A:`	q	i:`	S`
aj	d	aj	?`	A:`	?	i:`	R`
aj	t	aw	d`	A:`	v	i:`	x`
aj	g	aw	t`	A:`	f	i:`	X\`
aj	k	aw	z`	A:`	z	i:`	h`
aj	q	aw	s`	A:`	s	i:`	r`
aj	?	aw	?`	A:`	Z	i:`	l`
aj	v	A`	d`	A:`	S	i:`	j`
aj	f	A`	t`	A:`	R	i:`	w`
aj	z	A`	z`	A:`	x	u`	m`
aj	s	A`	s`	A:`	X\	u`	n`
aj	Z	A`	?`	A:`	h	u`	b`
aj	S	A:`	d`	A:`	r	u`	g`

aj	R	A:`	t`	A:`	l	u`	k`
aj	x	A:`	z`	A:`	j	u`	v`
aj	X\	A:`	s`	A:`	w	u`	f`
aj	h	A:`	?`	A`	m`	u`	Z`
aj	r	A`	m`	A`	n`	u`	S`
aj	l	A`	n`	A`	b`	u`	R`
aj	j	A`	b`	A`	g`	u`	x`
aj	w	A`	g`	A`	k`	u`	X\`
aw	m	A`	k`	A`	v`	u`	h`
aw	n	A`	v`	A`	f`	u`	r`
aw	b	A`	f`	A`	Z`	u`	l`
aw	d	A`	Z`	A`	S`	u`	j`
aw	t	A`	S`	A`	R`	u`	w`
aw	g	A`	R`	A`	x`	u:`	m`
aw	k	A`	x`	A`	X\`	u:`	n`
aw	q	A`	X\`	A`	h`	u:`	b`
aw	?	A`	h`	A`	r`	u:`	g`
aw	v	A`	r`	A`	l`	u:`	k`
aw	f	A`	l`	A`	j`	u:`	v`
aw	z	A`	j`	A`	w`	u:`	f`
aw	s	A`	w`	A:`	m`	u:`	Z`
aw	Z	A:`	m`	A:`	n`	u:`	S`
aw	S	A:`	n`	A:`	b`	u:`	R`
aw	R	A:`	b`	A:`	g`	u:`	x`
aw	x	A:`	g`	A:`	k`	u:`	X\`
aw	X\	A:`	k`	A:`	v`	u:`	h`
aw	h	A:`	v`	A:`	f`	u:`	r`
aw	r	A:`	f`	A:`	Z`	u:`	l`
aw	l	A:`	Z`	A:`	S`	u:`	j`
aw	j	A:`	S`	A:`	R`	u:`	w`
aw	w	A:`	R`	A:`	x`	aj`	m`
i	m`	A:`	x`	A:`	X\`	aj`	n`
i	n`	A:`	X\`	A:`	h`	aj`	b`
i	b`	A:`	h`	A:`	r`	aj`	g`
i	g`	A:`	r`	A:`	l`	aj`	k`
i	k`	A:`	l`	A:`	j`	aj`	v`
i	v`	A:`	j`	A:`	w`	aj`	f`
i	f`	A:`	w`	i`	m	aj`	Z`
i	Z`	i`	m`	i`	n	aj`	S`
i	S`	i`	n`	i`	b	aj`	R`
i	R`	i`	b`	i`	d	aj`	x`
i	x`	i`	g`	i`	t	aj`	X\`
i	X\`	i`	k`	i`	g	aj`	h`
i	h`	i`	v`	i`	k	aj`	r`
i	r`	i`	f`	i`	q	aj`	l`
i	l`	i`	Z`	i`	?	aj`	j`

i	j`	i`	S`	i`	v	aj`	w`
i	w`	i`	R`	i`	f	aw`	m`
i:	m`	i`	x`	i`	z	aw`	n`
i:	n`	i`	X\`	i`	s	aw`	b`
i:	b`	i`	h`	i`	Z	aw`	g`
i:	g`	i`	r`	i`	S	aw`	k`
i:	k`	i`	l`	i`	R	aw`	v`
i:	v`	i`	j`	i`	x	aw`	f`
i:	f`	i`	w`	i`	X\	aw`	Z`
i:	Z`	i:`	m`	i`	h	aw`	S`
i:	S`	i:`	n`	i`	r	aw`	R`
i:	R`	i:`	b`	i`	l	aw`	x`
i:	x`	i:`	g`	i`	j	aw`	X\`
i:	X\`	i:`	k`	i`	w	aw`	h`
aw`	l`	aw`	j`	aw`	w`	aw`	r`

Silence – C diphones

0	**C**	**0**	**C**
0	m	0	X\
0	n	0	h
0	b	0	r
0	d	0	l
0	t	0	j
0	d`	0	w
0	t`	0	m`
0	g	0	n`
0	k	0	b`
0	q	0	g`
0	?	0	k`
0	v	0	v`
0	f	0	f`
0	z	0	Z`
0	s	0	S`
0	z`	0	R`
0	s`	0	x`
0	Z	0	X\`
0	S	0	h`
0	R	0	r`
0	x	0	l`
0	?`	0	j`
		0	w`

C – Silence diphones

C	**0**	**C**	**0**
m	0	X\	0
n	0	h	0
b	0	r	0
d	0	l	0
t	0	j	0
d`	0	w	0
t`	0	m`	0
g	0	n`	0
k	0	b`	0
q	0	g`	0
?	0	k`	0
v	0	v`	0
f	0	f`	0
z	0	Z`	0
s	0	S`	0
z`	0	R`	0
s`	0	x`	0
Z	0	X\`	0
S	0	h`	0
R	0	r`	0
x	0	l`	0
?`	0	j`	0
		w`	0

V – Silence diphones

V	**0**
i	0
a	0
A`	0

Recorded Strings of Syllables

Total: 1090

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	m	a	m		i	m	
	m	a:	m		i	m	
	m	i:	m		i	m	
	m	u:	m		i	m	
	m	u	m		i	m	
	n	a	n		i	n	
	n	a:	n		i	n	
	n	i:	n		i	n	
	n	u:	n		i	n	
	n	u	n		i	n	
	b	a	b		i	b	
	b	a:	b		i	b	
	b	i:	b		i	b	
	b	u:	b		i	b	
	b	u	b		i	b	
	d	a	d		i	d	
	d	a:	d		i	d	
	d	i:	d		i	d	
	d	u:	d		i	d	
	d	u	d		i	d	
	t	a	t		i	t	
	t	a:	t		i	t	
	t	i:	t		i	t	
	t	u:	t		i	t	
	t	u	t		i	t	
	g	a	g		i	g	
	g	a:	g		i	g	
	g	i:	g		i	g	
	g	u:	g		i	g	
	g	u	g		i	g	
	k	a	k		i	k	
	k	a:	k		i	k	
	k	i:	k		i	k	
	k	u:	k		i	k	
	k	u	k		i	k	
	q	a	q		i	q	
	q	a:	q		i	q	
	q	i:	q		i	q	
	q	u:	q		i	q	
	q	u	q		i	q	
	?	a	?		i	?	
	?	a:	?		i	?	

silence	**C1**	**V1**	**C2**	**(C3)**	**V2**	**(C4)**	**silence**
	?	u:	?		i	?	
	?	u	?		i	?	
	v	a	v		i	v	
	v	a:	v		i	v	
	v	i:	v		i	v	
	v	u:	v		i	v	
	v	u	v		i	v	
	f	a	f		i	f	
	f	a:	f		i	f	
	f	i:	f		i	f	
	f	u:	f		i	f	
	f	u	f		i	f	
	z	a	z		i	z	
	z	a:	z		i	z	
	z	i:	z		i	z	
	z	u:	z		i	z	
	z	u	z		i	z	
	s	a	s		i	s	
	s	a:	s		i	s	
	s	i:	s		i	s	
	s	u:	s		i	s	
	s	u	s		i	s	
	Z	a	Z		i	Z	
	Z	a:	Z		i	Z	
	Z	i:	Z		i	Z	
	Z	u:	Z		i	Z	
	Z	u	Z		i	Z	
	S	a	S		i	S	
	S	a:	S		i	S	
	S	i:	S		i	S	
	S	u:	S		i	S	
	S	u	S		i	S	
	R	a	R		i	R	
	R	a:	R		i	R	
	R	i:	R		i	R	
	R	u:	R		i	R	
	R	u	R		i	R	
	x	a	x		i	x	
	x	a:	x		i	x	
	x	i:	x		i	x	
	x	u:	x		i	x	
	x	u	x		i	x	
	X\	a	X\		i	X\	
	X\	a:	X\		i	X\	
	X\	i:	X\		i	X\	

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	X\	u	X\		i	X\	
	h	a	h		i	h	
	h	a:	h		i	h	
	h	i:	h		i	h	
	h	u:	h		i	h	
	h	u	h		i	h	
	r	a	r		i	r	
	r	a:	r		i	r	
	r	i:	r		i	r	
	r	u:	r		i	r	
	r	u	r		i	r	
	l	a	l		i	l	
	l	a:	l		i	l	
	l	i:	l		i	l	
	l	u:	l		i	l	
	l	u	l		i	l	
	j	a	j		i	j	
	j	a:	j		i	j	
	j	i:	j		i	j	
	j	u:	j		i	j	
	j	u	j		i	j	
	w	a	w		i	w	
	w	a:	w		i	w	
	w	i:	w		i	w	
	w	u:	w		i	w	
	w	u	w		i	w	
	d`	i:`	d`		i`	d`	
	d`	u:`	d`		i`	d`	
	d`	A`	d`		i`	d`	
	d`	A:`	d`		i`	d`	
	t`	i:`	t`		i`	t`	
	t`	u:`	t`		i`	t`	
	t`	A`	t`		i`	t`	
	t`	A:`	t`		i`	t`	
	s`	i:`	s`		i`	s`	
	s`	u:`	s`		i`	s`	
	s`	A`	s`		i`	s`	
	s`	A:`	s`		i`	s`	
	z`	i:`	z`		i`	z`	
	z`	u:`	z`		i`	z`	
	z`	A`	z`		i`	z`	
	z`	A:`	z`		i`	z`	
	m`	i:`	d`		i`	m`	
	m`	u:`	d`		i`	m`	
	m`	A`	d`		i`	m`	

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	d`	i:`	m`		i`	m`	
	n`	i:`	d`		i`	n`	
	n`	u:`	d`		i`	n`	
	n`	A`	d`		i`	n`	
	n`	A:`	d`		i`	n`	
	d`	i:`	n`		i`	n`	
	b`	i:`	d`		i`	b`	
	b`	u:`	d`		i`	b`	
	b`	A`	d`		i`	b`	
	b`	A:`	d`		i`	b`	
	d`	i:`	b`		i`	b`	
	g`	i:`	d`		i`	g`	
	g`	u:`	d`		i`	g`	
	g`	A`	d`		i`	g`	
	g`	A:`	d`		i`	g`	
	d`	i:`	g`		i`	g`	
	k`	i:`	d`		i`	k`	
	k`	u:`	d`		i`	k`	
	k`	A`	d`		i`	k`	
	k`	A:`	d`		i`	k`	
	d`	i:`	k`		i`	k`	
	?`	i:`	d`		i`	?`	
	?`	u:`	d`		i`	?`	
	?`	A`	d`		i`	?`	
	?`	A:`	d`		i`	?`	
	d`	i:`	?`		i`	?`	
	v`	i:`	d`		i`	v`	
	v`	u:`	d`		i`	v`	
	v`	A`	d`		i`	v`	
	v`	A:`	d`		i`	v`	
	d`	i:`	v`		i`	v`	
	f`	i:`	d`		i`	f`	
	f`	u:`	d`		i`	f`	
	f`	A`	d`		i`	f`	
	f`	A:`	d`		i`	f`	
	d`	i:`	f`		i`	f`	
	Z`	i:`	d`		i`	Z`	
	Z`	u:`	d`		i`	Z`	
	Z`	A`	d`		i`	Z`	
	Z`	A:`	d`		i`	Z`	
	d`	i:`	Z`		i`	Z`	
	S`	i:`	d`		i`	S`	
	S`	u:`	d`		i`	S`	
	S`	A`	d`		i`	S`	
	S`	A:`	d`		i`	S`	

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	x`	i:`	d`		i`	x`	
	x`	u:`	d`		i`	x`	
	x`	A`	d`		i`	x`	
	x`	A:`	d`		i`	x`	
	d`	i:`	x`		i`	x`	
	X\`	i:`	d`		i`	X\`	
	X\`	u:`	d`		i`	X\`	
	X\`	A`	d`		i`	X\`	
	X\`	A:`	d`		i`	X\`	
	d`	i:`	X\`		i`	X\`	
	h`	i:`	d`		i`	h`	
	h`	u:`	d`		i`	h`	
	h`	A`	d`		i`	h`	
	h`	A:`	d`		i`	h`	
	d`	i:`	h`		i`	h`	
	r`	i:`	d`		i`	r`	
	r`	u:`	d`		i`	r`	
	r`	A`	d`		i`	r`	
	r`	A:`	d`		i`	r`	
	d`	i:`	r`		i`	r`	
	l`	i:`	d`		i`	l`	
	l`	u:`	d`		i`	l`	
	l`	A`	d`		i`	l`	
	l`	A:`	d`		i`	l`	
	d`	i:`	l`		i`	l`	
	j`	i:`	d`		i`	j`	
	j`	u:`	d`		i`	j`	
	j`	A`	d`		i`	j`	
	j`	A:`	d`		i`	j`	
	d`	i:`	j`		i`	j`	
	w`	i:`	d`		i`	w`	
	w`	u:`	d`		i`	w`	
	w`	A`	d`		i`	w`	
	w`	A:`	d`		i`	w`	
	d`	i:`	w`		i`	w`	
	?	i	m	m	a		
	?	i	m	n	a		
	?	i	m	b	a		
	?	i	m	d	a		
	?	i	m	t	a		
	?	i	m	d`	A`		
	?	i	m	t`	A`		
	?	i	m	g	a		
	?	i	m	k	a		
	?	i	m	q	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	m	v	a		
	?	i	m	f	a		
	?	i	m	z	a		
	?	i	m	s	a		
	?	i	m	z`	A`		
	?	i	m	s`	A`		
	?	i	m	Z	a		
	?	i	m	S	a		
	?	i	m	R	a		
	?	i	m	x	a		
	?	i	m	?`	A`		
	?	i	m	X\	a		
	?	i	m	h	a		
	?	i	m	r	a		
	?	i	m	l	a		
	?	i	m	j	a		
	?	i	m	w	a		
	?	i	n	m	a		
	?	i	n	n	a		
	?	i	n	b	a		
	?	i	n	d	a		
	?	i	n	t	a		
	?	i	n	d`	A`		
	?	i	n	t`	A`		
	?	i	n	g	a		
	?	i	n	k	a		
	?	i	n	q	a		
	?	i	n	?	a		
	?	i	n	v	a		
	?	i	n	f	a		
	?	i	n	z	a		
	?	i	n	s	a		
	?	i	n	z`	A`		
	?	i	n	s`	A`		
	?	i	n	Z	a		
	?	i	n	S	a		
	?	i	n	R	a		
	?	i	n	x	a		
	?	i	n	?`	A`		
	?	i	n	X\	a		
	?	i	n	h	a		
	?	i	n	r	a		
	?	i	n	l	a		
	?	i	n	j	a		
	?	i	n	w	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	b	n	a		
	?	i	b	b	a		
	?	i	b	d	a		
	?	i	b	t	a		
	?	i	b	d`	A`		
	?	i	b	t`	A`		
	?	i	b	g	a		
	?	i	b	k	a		
	?	i	b	q	a		
	?	i	b	?	a		
	?	i	b	v	a		
	?	i	b	f	a		
	?	i	b	z	a		
	?	i	b	s	a		
	?	i	b	z`	A`		
	?	i	b	s`	A`		
	?	i	b	Z	a		
	?	i	b	S	a		
	?	i	b	R	a		
	?	i	b	x	a		
	?	i	b	?`	A`		
	?	i	b	X\	a		
	?	i	b	h	a		
	?	i	b	r	a		
	?	i	b	l	a		
	?	i	b	j	a		
	?	i	b	w	a		
	?	i	d	m	a		
	?	i	d	n	a		
	?	i	d	b	a		
	?	i	d	d	a		
	?	i	d	t	a		
	?	i	d	d`	A`		
	?	i	d	t`	A`		
	?	i	d	g	a		
	?	i	d	k	a		
	?	i	d	q	a		
	?	i	d	?	a		
	?	i	d	v	a		
	?	i	d	f	a		
	?	i	d	z	a		
	?	i	d	s	a		
	?	i	d	z`	A`		
	?	i	d	s`	A`		
	?	i	d	Z	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	d	R	a		
	?	i	d	x	a		
	?	i	d	?`	A`		
	?	i	d	X\	a		
	?	i	d	h	a		
	?	i	d	r	a		
	?	i	d	l	a		
	?	i	d	j	a		
	?	i	d	w	a		
	?	i	t	m	a		
	?	i	t	n	a		
	?	i	t	b	a		
	?	i	t	d	a		
	?	i	t	t	a		
	?	i	t	d`	A`		
	?	i	t	t`	A`		
	?	i	t	g	a		
	?	i	t	k	a		
	?	i	t	q	a		
	?	i	t	?	a		
	?	i	t	v	a		
	?	i	t	f	a		
	?	i	t	z	a		
	?	i	t	s	a		
	?	i	t	z`	A`		
	?	i	t	s`	A`		
	?	i	t	Z	a		
	?	i	t	S	a		
	?	i	t	R	a		
	?	i	t	x	a		
	?	i	t	?`	A`		
	?	i	t	X\	a		
	?	i	t	h	a		
	?	i	t	r	a		
	?	i	t	l	a		
	?	i	t	j	a		
	?	i	t	w	a		
	?`	i`	d`	m	a		
	?`	i`	d`	n	a		
	?`	i`	d`	b	a		
	?`	i`	d`	d	a		
	?`	i`	d`	t	a		
	?`	i`	d`	d`	A`		
	?`	i`	d`	t`	A`		
	?`	i`	d`	g	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?`	i`	d`	q	a		
	?`	i`	d`	?	a		
	?`	i`	d`	v	a		
	?`	i`	d`	f	a		
	?`	i`	d`	z	a		
	?`	i`	d`	s	a		
	?`	i`	d`	z`	A`		
	?`	i`	d`	s`	A`		
	?`	i`	d`	Z	a		
	?`	i`	d`	S	a		
	?`	i`	d`	R	a		
	?`	i`	d`	x	a		
	?`	i`	d`	?`	A`		
	?`	i`	d`	X\	a		
	?`	i`	d`	h	a		
	?`	i`	d`	r	a		
	?`	i`	d`	l	a		
	?`	i`	d`	j	a		
	?`	i`	d`	w	a		
	?`	i`	t`	m	a		
	?`	i`	t`	n	a		
	?`	i`	t`	b	a		
	?`	i`	t`	d	a		
	?`	i`	t`	t	a		
	?`	i`	t`	d`	A`		
	?`	i`	t`	t`	A`		
	?`	i`	t`	g	a		
	?`	i`	t`	k	a		
	?`	i`	t`	q	a		
	?`	i`	t`	?	a		
	?`	i`	t`	v	a		
	?`	i`	t`	f	a		
	?`	i`	t`	z	a		
	?`	i`	t`	s	a		
	?`	i`	t`	z`	A`		
	?`	i`	t`	s`	A`		
	?`	i`	t`	Z	a		
	?`	i`	t`	S	a		
	?`	i`	t`	R	a		
	?`	i`	t`	x	a		
	?`	i`	t`	?`	A`		
	?`	i`	t`	X\	a		
	?`	i`	t`	h	a		
	?`	i`	t`	r	a		
	?`	i`	t`	l	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?`	i`	t`	w	a		
	?	i	g	m	a		
	?	i	g	n	a		
	?	i	g	b	a		
	?	i	g	d	a		
	?	i	g	t	a		
	?	i	g	d`	A`		
	?	i	g	t`	A`		
	?	i	g	g	a		
	?	i	g	k	a		
	?	i	g	q	a		
	?	i	g	?	a		
	?	i	g	v	a		
	?	i	g	f	a		
	?	i	g	z	a		
	?	i	g	s	a		
	?	i	g	z`	A`		
	?	i	g	s`	A`		
	?	i	g	Z	a		
	?	i	g	S	a		
	?	i	g	R	a		
	?	i	g	x	a		
	?	i	g	?`	A`		
	?	i	g	X\	a		
	?	i	g	h	a		
	?	i	g	r	a		
	?	i	g	l	a		
	?	i	g	j	a		
	?	i	g	w	a		
	?	i	k	m	a		
	?	i	k	n	a		
	?	i	k	b	a		
	?	i	k	d	a		
	?	i	k	t	a		
	?	i	k	d`	A`		
	?	i	k	t`	A`		
	?	i	k	g	a		
	?	i	k	k	a		
	?	i	k	q	a		
	?	i	k	?	a		
	?	i	k	v	a		
	?	i	k	f	a		
	?	i	k	z	a		
	?	i	k	s	a		
	?	i	k	z`	A`		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	k	Z	a		
	?	i	k	S	a		
	?	i	k	R	a		
	?	i	k	x	a		
	?	i	k	?`	A`		
	?	i	k	X\	a		
	?	i	k	h	a		
	?	i	k	r	a		
	?	i	k	l	a		
	?	i	k	j	a		
	?	i	k	w	a		
	?	i	q	m	a		
	?	i	q	n	a		
	?	i	q	b	a		
	?	i	q	d	a		
	?	i	q	t	a		
	?	i	q	d`	A`		
	?	i	q	t`	A`		
	?	i	q	g	a		
	?	i	q	k	a		
	?	i	q	q	a		
	?	i	q	?	a		
	?	i	q	v	a		
	?	i	q	f	a		
	?	i	q	z	a		
	?	i	q	s	a		
	?	i	q	z`	A`		
	?	i	q	s`	A`		
	?	i	q	Z	a		
	?	i	q	S	a		
	?	i	q	R	a		
	?	i	q	x	a		
	?	i	q	?`	A`		
	?	i	q	X\	a		
	?	i	q	h	a		
	?	i	q	r	a		
	?	i	q	l	a		
	?	i	q	j	a		
	?	i	q	w	a		
	j	i	?	m	a		
	j	i	?	n	a		
	j	i	?	b	a		
	j	i	?	d	a		
	j	i	?	t	a		
	j	i	?	d`	A`		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	j	i	?	g	a		
	j	i	?	k	a		
	j	i	?	q	a		
	j	i	?	?	a		
	j	i	?	v	a		
	j	i	?	f	a		
	j	i	?	z	a		
	j	i	?	s	a		
	j	i	?	z`	A`		
	j	i	?	s`	A`		
	j	i	?	Z	a		
	j	i	?	S	a		
	j	i	?	R	a		
	j	i	?	x	a		
	j	i	?	?`	A`		
	j	i	?	X\	a		
	j	i	?	h	a		
	j	i	?	r	a		
	j	i	?	l	a		
	j	i	?	j	a		
	j	i	?	w	a		
	?	i	v	m	a		
	?	i	v	n	a		
	?	i	v	b	a		
	?	i	v	d	a		
	?	i	v	t	a		
	?	i	v	d`	A`		
	?	i	v	t`	A`		
	?	i	v	g	a		
	?	i	v	k	a		
	?	i	v	q	a		
	?	i	v	?	a		
	?	i	v	v	a		
	?	i	v	f	a		
	?	i	v	z	a		
	?	i	v	s	a		
	?	i	v	z`	A`		
	?	i	v	s`	A`		
	?	i	v	Z	a		
	?	i	v	S	a		
	?	i	v	R	a		
	?	i	v	x	a		
	?	i	v	?`	A`		
	?	i	v	X\	a		
	?	i	v	h	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	v	l	a		
	?	i	v	j	a		
	?	i	v	w	a		
	?	i	f	m	a		
	?	i	f	n	a		
	?	i	f	b	a		
	?	i	f	d	a		
	?	i	f	t	a		
	?	i	f	d`	A`		
	?	i	f	t`	A`		
	?	i	f	g	a		
	?	i	f	k	a		
	?	i	f	q	a		
	?	i	f	?	a		
	?	i	f	v	a		
	?	i	f	f	a		
	?	i	f	z	a		
	?	i	f	s	a		
	?	i	f	z`	A`		
	?	i	f	s`	A`		
	?	i	f	Z	a		
	?	i	f	S	a		
	?	i	f	R	a		
	?	i	f	x	a		
	?	i	f	?`	A`		
	?	i	f	X\	a		
	?	i	f	h	a		
	?	i	f	r	a		
	?	i	f	l	a		
	?	i	f	j	a		
	?	i	f	w	a		
	?	i	z	m	a		
	?	i	z	n	a		
	?	i	z	b	a		
	?	i	z	d	a		
	?	i	z	t	a		
	?	i	z	d`	A`		
	?	i	z	t`	A`		
	?	i	z	g	a		
	?	i	z	k	a		
	?	i	z	q	a		
	?	i	z	?	a		
	?	i	z	v	a		
	?	i	z	f	a		
	?	i	z	z	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	z	z`	A`		
	?	i	z	s`	A`		
	?	i	z	Z	a		
	?	i	z	S	a		
	?	i	z	R	a		
	?	i	z	x	a		
	?	i	z	?`	A`		
	?	i	z	X\	a		
	?	i	z	h	a		
	?	i	z	r	a		
	?	i	z	l	a		
	?	i	z	j	a		
	?	i	z	w	a		
	?	i	s	m	a		
	?	i	s	n	a		
	?	i	s	b	a		
	?	i	s	d	a		
	?	i	s	t	a		
	?	i	s	d`	A`		
	?	i	s	t`	A`		
	?	i	s	g	a		
	?	i	s	k	a		
	?	i	s	q	a		
	?	i	s	?	a		
	?	i	s	v	a		
	?	i	s	f	a		
	?	i	s	z	a		
	?	i	s	s	a		
	?	i	s	z`	A`		
	?	i	s	s`	A`		
	?	i	s	Z	a		
	?	i	s	S	a		
	?	i	s	R	a		
	?	i	s	x	a		
	?	i	s	?`	A`		
	?	i	s	X\	a		
	?	i	s	h	a		
	?	i	s	r	a		
	?	i	s	l	a		
	?	i	s	j	a		
	?	i	s	w	a		
	?	i`	z`	m	a		
	?	i`	z`	n	a		
	?	i`	z`	b	a		
	?	i`	z`	d	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i`	z`	d`	A`		
	?	i`	z`	t`	A`		
	?	i`	z`	g	a		
	?	i`	z`	k	a		
	?	i`	z`	q	a		
	?	i`	z`	?	a		
	?	i`	z`	v	a		
	?	i`	z`	f	a		
	?	i`	z`	z	a		
	?	i`	z`	s	a		
	?	i`	z`	z`	A`		
	?	i`	z`	s`	A`		
	?	i`	z`	Z	a		
	?	i`	z`	S	a		
	?	i`	z`	R	a		
	?	i`	z`	x	a		
	?	i`	z`	?`	A`		
	?	i`	z`	X\	a		
	?	i`	z`	h	a		
	?	i`	z`	r	a		
	?	i`	z`	l	a		
	?	i`	z`	j	a		
	?	i`	z`	w	a		
	?	i`	s`	m	a		
	?	i`	s`	n	a		
	?	i`	s`	b	a		
	?	i`	s`	d	a		
	?	i`	s`	t	a		
	?	i`	s`	d`	A`		
	?	i`	s`	t`	A`		
	?	i`	s`	g	a		
	?	i`	s`	k	a		
	?	i`	s`	q	a		
	?	i`	s`	?	a		
	?	i`	s`	v	a		
	?	i`	s`	f	a		
	?	i`	s`	z	a		
	?	i`	s`	s	a		
	?	i`	s`	z`	A`		
	?	i`	s`	s`	A`		
	?	i`	s`	Z	a		
	?	i`	s`	S	a		
	?	i`	s`	R	a		
	?	i`	s`	x	a		
	?	i`	s`	?`	A`		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i`	s`	h	a		
	?	i`	s`	r	a		
	?	i`	s`	l	a		
	?	i`	s`	j	a		
	?	i`	s`	w	a		
	?	i	Z	m	a		
	?	i	Z	n	a		
	?	i	Z	b	a		
	?	i	Z	d	a		
	?	i	Z	t	a		
	?	i	Z	d`	A`		
	?	i	Z	t`	A`		
	?	i	Z	g	a		
	?	i	Z	k	a		
	?	i	Z	q	a		
	?	i	Z	?	a		
	?	i	Z	v	a		
	?	i	Z	f	a		
	?	i	Z	z	a		
	?	i	Z	s	a		
	?	i	Z	z`	A`		
	?	i	Z	s`	A`		
	?	i	Z	Z	a		
	?	i	Z	S	a		
	?	i	Z	R	a		
	?	i	Z	x	a		
	?	i	Z	?`	A`		
	?	i	Z	X\	a		
	?	i	Z	h	a		
	?	i	Z	r	a		
	?	i	Z	l	a		
	?	i	Z	j	a		
	?	i	Z	w	a		
	?	i	S	m	a		
	?	i	S	n	a		
	?	i	S	b	a		
	?	i	S	d	a		
	?	i	S	t	a		
	?	i	S	d`	A`		
	?	i	S	t`	A`		
	?	i	S	g	a		
	?	i	S	k	a		
	?	i	S	q	a		
	?	i	S	?	a		
	?	i	S	v	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	S	z	a		
	?	i	S	s	a		
	?	i	S	z`	A`		
	?	i	S	s`	A`		
	?	i	S	Z	a		
	?	i	S	S	a		
	?	i	S	R	a		
	?	i	S	x	a		
	?	i	S	?`	A`		
	?	i	S	X\	a		
	?	i	S	h	a		
	?	i	S	r	a		
	?	i	S	l	a		
	?	i	S	j	a		
	?	i	S	w	a		
	?	i	R	m	a		
	?	i	R	n	a		
	?	i	R	b	a		
	?	i	R	d	a		
	?	i	R	t	a		
	?	i	R	d`	A`		
	?	i	R	t`	A`		
	?	i	R	g	a		
	?	i	R	k	a		
	?	i	R	q	a		
	?	i	R	?	a		
	?	i	R	v	a		
	?	i	R	f	a		
	?	i	R	z	a		
	?	i	R	s	a		
	?	i	R	z`	A`		
	?	i	R	s`	A`		
	?	i	R	Z	a		
	?	i	R	S	a		
	?	i	R	R	a		
	?	i	R	x	a		
	?	i	R	?`	A`		
	?	i	R	X\	a		
	?	i	R	h	a		
	?	i	R	r	a		
	?	i	R	l	a		
	?	i	R	j	a		
	?	i	R	w	a		
	?	i	x	m	a		
	?	i	x	n	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	x	d	a		
	?	i	x	t	a		
	?	i	x	d`	A`		
	?	i	x	t`	A`		
	?	i	x	g	a		
	?	i	x	k	a		
	?	i	x	q	a		
	?	i	x	?	a		
	?	i	x	v	a		
	?	i	x	f	a		
	?	i	x	z	a		
	?	i	x	s	a		
	?	i	x	z`	A`		
	?	i	x	s`	A`		
	?	i	x	Z	a		
	?	i	x	S	a		
	?	i	x	R	a		
	?	i	x	x	a		
	?	i	x	?`	A`		
	?	i	x	X\	a		
	?	i	x	h	a		
	?	i	x	r	a		
	?	i	x	l	a		
	?	i	x	j	a		
	?	i	x	w	a		
	?	i`	?`	m	a		
	?	i`	?`	n	a		
	?	i`	?`	b	a		
	?	i`	?`	d	a		
	?	i`	?`	t	a		
	?	i`	?`	d`	A`		
	?	i`	?`	t`	A`		
	?	i`	?`	g	a		
	?	i`	?`	k	a		
	?	i`	?`	q	a		
	?	i`	?`	?	a		
	?	i`	?`	v	a		
	?	i`	?`	f	a		
	?	i`	?`	z	a		
	?	i`	?`	s	a		
	?	i`	?`	z`	A`		
	?	i`	?`	s`	A`		
	?	i`	?`	Z	a		
	?	i`	?`	S	a		
	?	i`	?`	R	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i`	?`	?`	A`		
	?	i`	?`	X\	a		
	?	i`	?`	h	a		
	?	i`	?`	r	a		
	?	i`	?`	l	a		
	?	i`	?`	j	a		
	?	i`	?`	w	a		
	?	i	X\	m	a		
	?	i	X\	n	a		
	?	i	X\	b	a		
	?	i	X\	d	a		
	?	i	X\	t	a		
	?	i	X\	d`	A`		
	?	i	X\	t`	A`		
	?	i	X\	g	a		
	?	i	X\	k	a		
	?	i	X\	q	a		
	?	i	X\	?	a		
	?	i	X\	v	a		
	?	i	X\	f	a		
	?	i	X\	z	a		
	?	i	X\	s	a		
	?	i	X\	z`	A`		
	?	i	X\	s`	A`		
	?	i	X\	Z	a		
	?	i	X\	S	a		
	?	i	X\	R	a		
	?	i	X\	x	a		
	?	i	X\	?`	A`		
	?	i	X\	X\	a		
	?	i	X\	h	a		
	?	i	X\	r	a		
	?	i	X\	l	a		
	?	i	X\	j	a		
	?	i	X\	w	a		
	?	i	h	m	a		
	?	i	h	n	a		
	?	i	h	b	a		
	?	i	h	d	a		
	?	i	h	t	a		
	?	i	h	d`	A`		
	?	i	h	t`	A`		
	?	i	h	g	a		
	?	i	h	k	a		
	?	i	h	q	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	h	v	a		
	?	i	h	f	a		
	?	i	h	z	a		
	?	i	h	s	a		
	?	i	h	z`	A`		
	?	i	h	s`	A`		
	?	i	h	Z	a		
	?	i	h	S	a		
	?	i	h	R	a		
	?	i	h	x	a		
	?	i	h	?`	A`		
	?	i	h	X\	a		
	?	i	h	h	a		
	?	i	h	r	a		
	?	i	h	l	a		
	?	i	h	j	a		
	?	i	h	w	a		
	?	i	r	m	a		
	?	i	r	n	a		
	?	i	r	b	a		
	?	i	r	d	a		
	?	i	r	t	a		
	?	i	r	d`	A`		
	?	i	r	t`	A`		
	?	i	r	g	a		
	?	i	r	k	a		
	?	i	r	q	a		
	?	i	r	?	a		
	?	i	r	v	a		
	?	i	r	f	a		
	?	i	r	z	a		
	?	i	r	s	a		
	?	i	r	z`	A`		
	?	i	r	s`	A`		
	?	i	r	Z	a		
	?	i	r	S	a		
	?	i	r	R	a		
	?	i	r	x	a		
	?	i	r	?`	A`		
	?	i	r	X\	a		
	?	i	r	h	a		
	?	i	r	r	a		
	?	i	r	l	a		
	?	i	r	j	a		
	?	i	r	w	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	l	n	a		
	?	i	l	b	a		
	?	i	l	d	a		
	?	i	l	t	a		
	?	i	l	d`	A`		
	?	i	l	t`	A`		
	?	i	l	g	a		
	?	i	l	k	a		
	?	i	l	q	a		
	?	i	l	?	a		
	?	i	l	v	a		
	?	i	l	f	a		
	?	i	l	z	a		
	?	i	l	s	a		
	?	i	l	z`	A`		
	?	i	l	s`	A`		
	?	i	l	Z	a		
	?	i	l	S	a		
	?	i	l	R	a		
	?	i	l	x	a		
	?	i	l	?`	A`		
	?	i	l	X\	a		
	?	i	l	h	a		
	?	i	l	r	a		
	?	i	l	l	a		
	?	i	l	j	a		
	?	i	l	w	a		
	?	i	j	m	a		
	?	i	j	n	a		
	?	i	j	b	a		
	?	i	j	d	a		
	?	i	j	t	a		
	?	i	j	d`	A`		
	?	i	j	t`	A`		
	?	i	j	g	a		
	?	i	j	k	a		
	?	i	j	q	a		
	?	i	j	?	a		
	?	i	j	v	a		
	?	i	j	f	a		
	?	i	j	z	a		
	?	i	j	s	a		
	?	i	j	z`	A`		
	?	i	j	s`	A`		
	?	i	j	Z	a		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	?	i	j	R	a		
	?	i	j	x	a		
	?	i	j	?`	A`		
	?	i	j	X\	a		
	?	i	j	h	a		
	?	i	j	r	a		
	?	i	j	l	a		
	?	i	j	j	a		
	?	i	j	w	a		
	?	i	w	m	a		
	?	i	w	n	a		
	?	i	w	b	a		
	?	i	w	d	a		
	?	i	w	t	a		
	?	i	w	d`	A`		
	?	i	w	t`	A`		
	?	i	w	g	a		
	?	i	w	k	a		
	?	i	w	q	a		
	?	i	w	?	a		
	?	i	w	v	a		
	?	i	w	f	a		
	?	i	w	z	a		
	?	i	w	s	a		
	?	i	w	z`	A`		
	?	i	w	s`	A`		
	?	i	w	Z	a		
	?	i	w	S	a		
	?	i	w	R	a		
	?	i	w	x	a		
	?	i	w	?`	A`		
	?	i	w	X\	a		
	?	i	w	h	a		
	?	i	w	r	a		
	?	i	w	l	a		
	?	i	w	j	a		
	?	i	w	w	a		
	m	aj		m	i		
	n	aj		n	u		
	b	aj		b	i		
	d	aj		d	u		
	t	aj		t	i		
	d`	aj`		d`	u`		
	t`	aj`		t`	i`		
	g	aj		g	u		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	q	aj		q	u		
	?	aj		?	i		
	v	aj		v	u		
	f	aj		f	i		
	z	aj		z	u		
	s	aj		s	i		
	z`	aj`		z`	u`		
	s`	aj`		s`	i`		
	Z	aj		Z	u		
	S	aj		S	i		
	R	aj		R	u		
	x	aj		x	i		
	?`	aj`		?`	u		
	X\	aj		X\	i		
	h	aj		h	u		
	r	aj		r	i		
	l	aj		l	u		
	j	aj		j	i		
	w	aj		w	u		
	m	aw		m	i		
	n	aw		n	u		
	b	aw		b	i		
	d	aw		d	u		
	t	aw		t	i		
	d`	aw`		d`	u		
	t`	aw`		t`	i		
	g	aw		g	u		
	k	aw		k	i		
	q	aw		q	u		
	?	aw		?	i		
	v	aw		v	u		
	f	aw		f	i		
	z	aw		z	u		
	s	aw		s	i		
	z`	aw`		z`	u`		
	s`	aw`		s`	i`		
	Z	aw		Z	u		
	S	aw		S	i		
	R	aw		R	u		
	x	aw		x	i		
	?`	aw`		?`	u		
	X\	aw		X\	i		
	h	aw		h	u		
	r	aw		r	i		
	l	aw		l	u		

silence	C1	V1	C2	(C3)	V2	(C4)	silence
	w	aw		w	u		
	m`	aj`		m`	i`		
	n`	aj`		n`	u`		
	b`	aj`		b`	i`		
	g`	aj`		g`	u`		
	k`	aj`		k`	i`		
	v`	aj`		v`	u`		
	f`	aj`		f`	i`		
	Z`	aj`		Z`	u`		
	S`	aj`		S`	i`		
	R`	aj`		R`	u`		
	x`	aj`		x`	i`		
	X\`	aj`		X\`	u`		
	h`	aj`		h`	i`		
	r`	aj`		r`	u`		
	l`	aj`		l`	i`		
	j`	aj`		j`	u`		
	w`	aj`		w`	i`		
	m`	aw`		m`	u`		
	n`	aw`		n`	i`		
	b`	aw`		b`	u`		
	g`	aw`		g`	i`		
	k`	aw`		k`	u`		
	v`	aw`		v`	i`		
	f`	aw`		f`	u`		
	Z`	aw`		Z`	i`		
	S`	aw`		S`	u`		
	R`	aw`		R`	i`		
	x`	aw`		x`	u`		
	X\`	aw`		X\`	i`		
	h`	aw`		h`	u`		
	r`	aw`		r`	i`		
	l`	aw`		l`	u`		
	j`	aw`		j`	i`		
	w`	aw`		w`	u`		