

# **Diglossia and Codeswitching among Public Figures in the Arabian Gulf Region**

**by  
Arwa Alosaimi**

B.A, King Saud University, 2010

Thesis Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
Master of Arts

in the  
Department of Linguistics  
Faculty of Arts and Social Sciences

© Arwa Alosaimi 2019  
SIMON FRASER UNIVERSITY  
Fall 2019

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

# Approval

**Name:** Arwa Alosaimi

**Degree:** Master of Arts

**Title:** Diglossia and Codeswitching among Public Figures in the Arabian Gulf Region

**Examining Committee:**

**Chair:** Chung-hye Han  
Professor

**Panayiotis Pappas**  
Senior Supervisor  
Associate Professor

**Suzanne Hilgendorf**  
Supervisor  
Associate Professor

**Dean Mellow**  
Supervisor  
Associate Professor

**Catherine Léger**  
External Examiner  
Associate Professor  
Department of French  
University of Victoria

**Date Defended/Approved:** November 28, 2019

## Abstract

The Arab Gulf region is one of the standard examples of diglossia, where standard Arabic (SA), as a high variety, and colloquial Arabic (CA), as a low variety, co-exist (Ferguson, 1959; Kaye, 1970; Wardhaugh & Fuller, 2015). However, this linguistic pattern has been changing, with speakers increasingly using English due to globalization (Winford, 2002). In this paper, I examine the use of SA, CA, and English in the speech of two groups of public figures in the Arabian Gulf region: politicians and actors. It is hypothesized that the two groups will have distinct speech styles as they are constructing different personae (Eckert, 2004): a socially conservative one for politicians, and a more popular and cosmopolitan one for actors. It is further hypothesized that politicians will use standard Arabic more, because of its association with conservative values, and actors will use more colloquial Arabic and English for their popular association. The study uses naturalistic data that was obtained by examining a 30-minute interview, from public recordings on YouTube, for each of the 18 subjects: six politicians (three males, three females) and 12 actors (six males, six females). An analysis of 2,808 utterances yields results that support the hypothesis. More than half of the politicians' utterances are a mixture of SA and CA. Actors, on the other hand, adopt a different style in which CA is the dominant code, but they also use utterances that are a mixture of CA and English borrowings (12% of their speech). These findings can be seen as an indication of a linguistic change in the region: CA forms are replacing SA, not only in the speech of popular figures such as actors but also in that of more conservative speakers such as politicians. The appearance of English borrowings in the utterances of the actors is another indication that linguistic norms in the Arab Gulf region might be shifting towards the vernacular.

**Keywords:** variation; style; diglossia; codeswitching; gender; Gulf Arabic

To

the souls of my grandparents

my mother who models to me: strength, patience, and unending faith

my brother who has always been my pillar of support and encouragement

## **Acknowledgements**

First and foremost, I would like to express my sincere thanks and appreciation to Dr. Panayiotis Pappas, my senior supervisor, whose patience, moral support, understanding, and continuous encouragement made this study possible. He has been there providing his friendly guidance and expert advice throughout all stages of the thesis and the entire period of my studies in the Department of Linguistics at Simon Fraser University. I would also wish to express my gratitude to Dr. Suzanne Hilgendorf, my supervisor, for her valuable suggestions and comments, which have contributed greatly to the improvement of the thesis. My thanks are extended to the members of the committee, Dr. Chung-hye Han, Dr. Dean Mellow, and Dr. Catherine Léger, for the time and energy they spent reading my thesis and for their valuable comments. I am grateful to all of those with whom I have had the pleasure to work during my stay at SFU.

Nobody has been more important to me in the pursuit of this than the members of my family. I would like to thank my mother, Haya Alarifi, who supported me emotionally and financially. I am very grateful for my grandmother, Nourah Al-Khaledi, whose love and prayers were with me in whatever I pursued. Special thanks are due to my loving siblings, Lena and Assem, for their constant encouragement.

# Table of Contents

Approval.....	ii
Abstract.....	iii
Dedication.....	iv
Acknowledgements.....	v
Table of Contents.....	vi
List of Tables.....	viii
List of Figures.....	ix
List of Abbreviation.....	x
<b>Chapter 1. Introduction.....</b>	<b>1</b>
<b>Chapter 2. Literature Review.....</b>	<b>4</b>
2.1. Approaches to stylistic variation.....	4
2.1.1. The definition of style.....	4
2.1.2. Three waves of variation.....	6
2.1.2.1. The first wave: broad correlations.....	6
2.1.2.2. The second wave: local meanings.....	7
2.1.2.2.1. Ethnographic approaches.....	7
2.1.2.2.2. Bell's audience design.....	9
2.1.2.3. The third wave: speaker design.....	11
2.1.2.3.1. Practice.....	12
2.1.2.3.2. Indexicality.....	14
2.1.2.3.3. Stylistic practice.....	16
2.2. Diglossia in Arabic.....	17
2.2.1. Modern standard Arabic vs. colloquial Arabic in the Arabian Gulf region.....	21
2.2.2. Some linguistic characteristics of Gulf Arabic.....	24
2.2.3. Studies in diglossia.....	26
2.3. The role of English in the Arabian Gulf Region.....	27
2.4. Codeswitching.....	30
2.4.1. Unit of analysis in codeswitching.....	32
2.4.2. Codeswitching or borrowing? The terminology dilemma.....	33
2.4.3. Motivations for codeswitching.....	34
2.4.3.1. Codeswitching.....	34
2.4.3.2. Diglossic switching.....	35
2.4.4. Codeswitching and gender as a sociolinguistic variable.....	35
2.4.4.1. Codeswitching.....	35
2.4.4.2. Diglossic switching.....	36
<b>Chapter 3. Methodology.....</b>	<b>38</b>
3.1. Methodology in the previous literature.....	38
3.2. Subjects.....	41
3.3. Materials.....	42

3.4. Procedure .....	43
3.4.1. Coding .....	44
<b>Chapter 4. Results .....</b>	<b>49</b>
4.1. General description of the corpus .....	49
4.2. Structural aspects of codeswitching .....	52
4.2.1. Arabic/English codeswitching .....	53
4.2.1.1. General description of English lexical insertions .....	53
4.2.1.2. General description of English multi-lexical insertions .....	56
4.2.2. Colloquial Arabic/standard Arabic codeswitching .....	59
4.2.2.1. SA-dominant utterances .....	59
4.2.2.1.1. General description of CA lexical insertions .....	61
4.2.2.1.2. General description of CA multi-lexical insertions .....	65
4.2.2.1.3. General description of CA sentential insertions .....	69
4.2.2.2. CA-dominant utterances .....	70
4.2.2.2.1. General description of SA multi-lexical insertions .....	71
4.2.2.2.2. General description of SA sentential insertions .....	74
4.3. Sociolinguistics aspects of style variation .....	77
4.3.1. Style and occupation .....	77
4.3.2. Style and gender .....	78
4.3.3. Style and age .....	80
4.4. Summary of the results .....	81
<b>Chapter 5. Discussion and conclusion .....</b>	<b>83</b>
5.1. The structural characteristics of the overall pattern of language use .....	83
5.1.1. Diglossia .....	83
5.1.2. Codeswitching .....	84
5.2. Social characteristics and persona construction .....	86
5.2.1. Style and occupation .....	86
5.2.1.1. The conservative persona .....	86
5.2.1.2. The popular persona .....	88
5.2.2. Style and gender .....	88
5.2.2.1. Gender and the politician persona .....	89
5.2.2.2. Gender and the popular persona .....	90
5.2.3. Style and age .....	93
5.3. Conclusion .....	95
<b>References .....</b>	<b>99</b>

## List of Tables

Table 1.	Distribution of subjects by country.....	42
Table 2.	Types of utterances in the corpus .....	49
Table 3.	Lexical, multi-lexical, and sentential English elements.....	53
Table 4.	English single words in Arabic frames.....	54
Table 5.	EN multi-lexical elements in Arabic frames .....	57
Table 6.	Utterances according to the dominant language.....	59
Table 7.	Distribution of CA elements in SA-dominant utterances .....	60
Table 8.	CA single words in SA-dominant utterances .....	61
Table 9.	CA multi-lexical elements in SA-dominant utterances.....	65
Table 10.	CA sentential elements in SA-dominant utterances .....	69
Table 11.	Distribution of SA elements in CA-dominant utterances .....	71
Table 12.	SA multi-lexical elements in CA-dominant utterances.....	71
Table 13.	SA inter-sentential elements in CA-dominant utterances.....	75
Table 14.	Distribution of utterance types for politicians.....	78
Table 15.	Distribution of utterance types for actors.....	78
Table 16.	Distribution of utterance types for politicians by gender.....	79
Table 17.	Distribution of utterance types for actors by gender.....	80
Table 18.	Distribution of utterances types for actors by age .....	81



## List of Figures

Figure 1.	The genealogical tree of Arabic.....	22
Figure 2.	Distribution of monolingual SA, diglossic, and codeswitched utterances for politicians by gender .....	90
Figure 3.	Distribution of diglossic and codeswitched utterances for actors by gender .....	91
Figure 4.	Distribution of diglossic and codeswitched utterances for actors by age 95	

## List of Abbreviation

1PI	First-person plural pronoun
1Sig	First-person singular pronoun
2Sig	Second-person singular pronoun
3PI	Third-person plural pronoun
3Sig	Third-person singular pronoun
Adj	Adjective
ART	Arab Radio and Television
CA	Colloquial Arabic
Conj	Conjunction
CP	Complementizer phrase
Def	Definitive article
Dem	Demonstrative pronoun
EFL	English as a foreign language
EIAL	English as International Auxiliary Language
EL	Embedded language
EN	English
ENL	English as a native language
ESD	English as a second dialect
ESL	English as a second language
Fem	Feminine suffix
GA	Gulf Arabic
GCC	Gulf Cooperation Council
LBC	Lebanese Broadcasting Cooperation
ML	Matrix language
MLF	Matrix language-frame model
MSA	Modern Standard Arabic
Neg	Negation
Pass	Passive voice
Past	Past tense
PI	Plural suffix
PoS	Part of speech

Poss

Pres

SA

WE

Possessive

Present tense

Standard Arabic

World Englishes

# Chapter 1.

## Introduction

Arabic has attracted the attention of linguists not only for its diglossic nature, where speakers shift from a high variety (formal) to what they consider a low one (informal) or vice versa (Ferguson, 1959), but also for its widespread distribution in the Arab world. The Arab world is the region that spreads between the Atlantic Ocean and the Arabian Gulf. There are more than 20 Arabic-speaking countries in this region with a total population of more than 300 million people. Those countries are Saudi Arabia, United Arab Emirates, Bahrain, Kuwait, Oman, Qatar, Yemen, Iraq, Jordan, Syria, Lebanon, Palestine, Egypt, Sudan, Libya, Tunisia, Algeria, Morocco, Mauritania, Somalia, Comoros, and Djibouti (Habash, 2010). This wide use of Arabic by different speech communities in different geographical places results in an array of variation that has been a fertile area of research. In some Arabic-speaking countries, such as Morocco, that have experienced extensive language contact, borrowings and the influence of other languages have affected Arabic in several ways (Kachoub & Hilgendorf, 2019). This helps flourish the studies in codeswitching, which is the alternation between two or more languages while speaking (Myers-Scotton, 1993a). Other parts of the Arab world, such as the Arabian Gulf region, have not had contact with a language other than Arabic until the era of globalization and the revolution of technology (Winford, 2002). Globalization and technology open the gate for more interaction with the outside world in the Arabian Gulf region. In order to get connected with people around the world, English, as an international language, is the best choice for speakers there.

According to Kachru (1985), the spread and internationalization of English have led to the emergence of diverse varieties of this language. Kachru explains the historical spread of English through the World Englishes paradigm, which consists of three circles: the Inner Circle, the Outer Circle, and the Expanding Circle. The Inner Circle refers to any country that historically has used English as a primary language such as the United Kingdom and the United States. The Outer Circle includes all countries where English is not a native language, but is crucial for historical reasons and plays a pivotal role in the

countries' institutions such as India and the Philippines. Kachru considers the Gulf region countries to be part of the Expanding Circle, which encompasses all the countries that use English as a foreign language. English has not necessarily had a historical or governmental role in the countries of this circle. However, the situation has changed in the Gulf region, where most governments started some new curriculum reforms where English is favored as the medium of instruction (Ahmed, 2010). In Saudi Arabia, for example, English is officially mandated as the language of instruction in most majors in post-secondary education (Ebad, 2014). Such changes contribute to the increased use of English in the region. Those changes of English use in the Arabian Gulf region, as well as in other countries in the Expanding Circle, indicate a shift of the role of English from being a foreign language with limited usage to an additional language (Hilgendorf, 2018, p. 480). This has led me to the question of how speakers in the Gulf region use the three codes: standard Arabic, colloquial Arabic, and English.

I pursue this question as an issue of stylistic variation. There have been three main approaches to this concept. Labov (1972) attributes variation to how much attention the speaker gives to their speech. According to him, style can be either formal, when the speaker pays careful attention to their speech, or informal, when the attention decreases such as in daily communication. Bell (1984) finds many incidents of style variation that are not the result of how much attention is paid to speech, but of . This design basically indicates that style variation happens because “the speakers take most account of hearers in designing their talk” (p. 159). The third approach is Eckert's (2000), who challenges the two previous approaches as being too narrow. Eckert (2000) suggests a broader framework that links variation to social meaning. Eckert (2004) asserts that in order to understand style variation it is crucial “...to concern ourselves with the nature of the social meaning it carries and the mechanisms by which variation comes to be endowed with meaning” (p. 51). The main concept of Eckert's approach is that a speaker uses style in order to create a unique persona.

I will focus on the use of SA, CA and English in the speech of two groups of public figures in the Arabian Gulf region: politicians and actors. It is hypothesized that the two groups will have distinct speech styles as they are constructing different personae (Eckert, 2004): a socially conservative one for politicians, and a more popular and cosmopolitan one for actors. It is further hypothesized that politicians will use SA more, because of its association with conservative values, and actors will use more CA

and English for its popular association. In order to examine this hypothesis, the data is gathered by analyzing the speech in a 30-minute interview for each of 18 subjects who are divided into two main groups based on their occupations: six subjects and 12 actors. Gender and age are two grouping factors used for actors, whereas only gender is used to group politicians. The total length of the interviews was nine hours and they were obtained from the video-sharing website YouTube. The thesis will examine the following questions:

- What are the overall patterns of language use?
- What are the structural characteristics of these patterns?
- What are the social characteristics, and in particular are there differences in occupation, gender, and age associated to particular patterns?

By examining these questions, the thesis will try to determine how English fits in the diglossic linguistic environment in the Gulf region languages. It will also contribute to a greater understanding of English use in the Expanding Circle in the World Englishes paradigm. Moreover, it will help us to better understand style variation by giving the opportunity to those who are interested in this area to view style shifting in a different language and from different perspectives. The results of style variation in this study will be beneficial for cross-cultural research in sociolinguistics and will provide the field with new material that might improve our perception of variation.

The overall thesis objectives are two-fold: to examine the importance and the power of linguistic style in self-representation by tracking the use of SA, CA and English in the study's subjects, and to examine the rising phenomenon of English borrowings among Arabic speakers in this part of the Arab world that has resisted linguistic change for a long time.

The thesis is organized as follows. In chapter 2, I will illustrate and discuss the relevant literature on style variation, codeswitching, and diglossia. In chapter 3, I will discuss the different types of methods that have been adopted to collect and analyze data in previous research, and I will present the design of this study and how I collected and analyzed my data. Then, I will present the results related to the structural and sociolinguistic aspects of the data in chapter 4. Chapter 5 provides a discussion of the results, and a conclusion.

## Chapter 2.

### Literature Review

In this chapter, the necessary background information for some of the concepts that will be discussed in the rest of the thesis is provided. Section 2.1 reviews past and current approaches to analyzing linguistic style. Section 2.2 focuses on the concept of diglossia in Arabic and some linguistic characteristics of Gulf Arabic. Section 2.3 is about the role of English in the Arabian Gulf region. Finally, a review on the concept of codeswitching is provided in section 2.4.

#### 2.1. Approaches to stylistic variation

##### 2.1.1. The definition of style

Although style has been included as a factor in language variation from even the earliest studies of Labov (1966), there is no precise definition for this term in the literature, as it is defined differently according to the approach used to explore it. Coupland (2007) equates style with dialect, a term “used to refer to ways of speaking that are indexically linked to social groups, times and places...” (p. 2). Despite being general, Coupland’s definition would also fit for register and genre as two notions related to style. This *terminological maze*, as Moessner (2001) describes it, has led to discrepancies in variationist research. Therefore, the next paragraphs will shed some light on what the literature has said about the concepts of register, dialect, genre, and their relationship to style.

Coulthard (2014) acknowledges the fact that the concepts of both register and style might seem very close in meaning. However, he points out a key difference. He states that register is connected to certain situations and known by topic and “context-specific lexis”, whereas style is not linked to any situation and a speaker’s choice among styles has social meaning (Coulthard, 2014, p. 40). Holmes and Wilson (2017) also mention the unclear distinction between the two terms and how some sociolinguists do not even distinguish register from style (p. 280). Other researchers have tried to draw a line between the two notions by stating that register is more concerned with lexical

variation, whereas variation in style can be grammatical, syntactical, lexical, and phonological (Kortmann, 2005). In general, Mejdell (2006) indicates that the distinction is subject to the preferences of researchers who either treat the two concepts, register and style, synonymously or as completely different entities.

Irvine (2001) discusses this terminological dilemma and how many researchers focus on the terms that overlap with style, such as dialect and register. Irvine notices that some American linguists use style as a synonym for the British term register (Irvine, 2001, p. 27). As for the distinction between register and dialect, this lies in functions and not in forms. Speakers use a specific dialect to indicate their affiliation with certain social groups, whereas register is used to indicate characteristics of a current situation or social activity (Irvine, 2001, p. 27). Although Irvine explains the definitional differences of these terms, she encourages researchers to direct their efforts towards exploring the principles that govern the relationships and the distinguishing traits of varieties instead of focusing on taxonomic distinctions.

Genre is another problematic concept that overlaps not only with style but also with register. According to the *Encyclopedic Dictionary of Applied Linguistics* (Johnson and Johnson, 1998), “genres are types of spoken and written discourse recognized by a discourse community” (p. 140). Coupland (2007) believes that most of genre’s definitions correspond to that of social style (p. 15). Wardhaugh and Fuller (2015), however, believe that “dialect, style, register, and genre differences are largely independent” (p. 53). They state that “differences in dialect have to do with speakers and their regional or social identities, styles, registers, and genres have to do with different contexts of use” (p. 52).

An attempt to dissolve this terminological confusion regarding all three concepts is found in the work of Biber and Conrad (2019), where they explain the characteristics of each term. They consider register, genre, and style as three possible perspectives for the analysis of text varieties but not three types of varieties in themselves. In their explanation, register, as the focus of their book, has some similarities with style and genre. Both register and style perspectives are concerned with the analysis of linguistic features in a variety, however, style differs in the fact that “the use of these features is not functionally motivated by the situational context;” as is the case in register, but “rather, style features reflect aesthetic preferences, associated with particular authors or historical periods” (Biber & Conrad, 2019, p. 2). Genre and register, on the other hand,



are both connected to certain contexts, but the two notions are different in their linguistic analysis where the former focuses on “the conventional structures used to construct a complete text within the variety, for example, the conventional way in which a letter begins and ends” (Biber & Conrad, 2019, p. 2).

A recent attempt to construct an explicit definition of style is that of Theodoropoulou (2014). She suggests a broad definition that she claims covers three crucial analytical approaches to style: sociolinguistic variation, , and speaker design. According to her, style is “a varying and flexible semiotic repertoire through which individuals and groups craft and index their identities to the rest of the world, depending on communication circumstances” (Theodoropoulou, 2014, p. 7). This definition compresses what Eckert (2012) calls the three waves of variation in sociolinguistic studies, which will be explored, in the next sections.

### **2.1.2. Three waves of variation**

According to Eckert (2012), the study of variation in sociolinguistic research has developed in three waves. The first wave was characterized by general correlations between linguistic variables with social variables of class, age, gender, etc. The second wave used ethnographic methods to examine local categories that establish those social variables. The third wave treated variation as a social semiotic system that constitutes social meaning rather than reflecting it. I discuss each wave in more detail below.

#### **2.1.2.1. *The first wave: broad correlations***

Labov’s (1966) study on the social stratification of English in New York City is considered one of the first investigations on style variation. The data of this study were collected by two methods: interviewing 70 individuals, and anonymously observing people’s speech in public places in the Lower East Side of New York City. The primary focus is on obtaining data that represent people’s casual and natural speech. In order to achieve this goal, Labov (1966) developed a sociolinguistic interview that can be used to produce an array of speech styles that ranges from casual to formal. The sociolinguistic interview consists of four tasks: a reading passage, a word-list, a list of minimal pairs, and talking with the interviewee about a life experience, some of their beliefs, or their life in general. The first three tasks are designed to elicit formal, careful speech, whereas the last task is intended to trigger a more casual one. The various style shifts yielded

from this interview are attributed to how much attention is paid to speech. In other words, the style will be closer to the vernacular when the speaker is the least conscious of their speech. As consciousness to speech increases, the speaker then will use a style that is closer to the standard variety.

The findings of the New York survey support Labov's notion of the relationship between attention to speech and style variation. For instance, the variable (r) in postvocalic position has two variants: either pronounced as a constricted [r], or not pronounced. The results indicate that this variable is sensitive to the speaker's attention to speech. The constricted [r] is used more in minimal pairs when the speech is self-conscious. The frequency of this variant gradually drops as speech moves from the minimal pairs category to the casual category, where speech is less self-conscious. This trend is found in all of the other examined variables in Labov's study.

In this approach, the concept of standard has been associated with educated speakers from the upper middle class and it refers to any speech that does not have any regional or socially stigmatized features. The generalization that most studies share is that the social status of speakers is indicated by their linguistic varieties. They also see class stratification of language as a matter of linguistic prestige. Social meaning in this wave is directly linked to the socioeconomic hierarchy. Variables are viewed as social status markers and meaning is established on the basis of the binary opposition of the prestigious variety of language, i.e. formal speech, and the stigmatized vernacular.

### **2.1.2.2. *The second wave: local meanings***

#### **2.1.2.2.1. Ethnographic approaches**

Researchers in the second wave shift to ethnographic studies that pay particular attention to small communities for a quite long period of time. The main goal of these studies is to find out locally important social categories. According to Eckert (2009), those social categories might be local representations of the categories found in the survey studies of the first wave or they might be simply different (p. 7). Whether they are representations of the categories of the survey studies or not, they provide a better understanding of how variation is strongly linked with local meaning.

An example of an ethnographic study would be Labov's (1963) research on language variation in Martha's Vineyard island that is located south of Cape Cod in

Massachusetts. Labov conducted this study to investigate the linguistic variation in the pronunciation of two variables (ay) and (aw) and their social implications. It was noticed that local people of Martha's Vineyard tend to centralize the first elements of those diphthongs to become [əɪ], [əu] instead of the standard New England pronunciations [aɪ], [au]. Both diphthongs were examined in relation to age, ethnic group, occupation, and geographical location.

The results showed that centralization was frequent in the 31-45 age group. As for the geographical location, it was noticed that rural up-islanders centralized those diphthongs more than residents of the down island areas. In terms of occupation, fishermen were the lead category in these variants. The results of the ethnic groups, however, seemed puzzling in the sense that there is no explicit pattern that can account for the shift. This lack of clear correlations made Labov examine the social structure as well as the economic pressures that led to a linguistic change on the island.

The economy of Martha's Vineyard heavily depends on tourism, where nearly 42,000 tourists visit the island every summer. The dependence on tourism trade and the declining economy of the island, where the unemployment rate is very high and the average income is extremely low compared to other Massachusetts counties, put pressure on the local people. This pressure makes most of the Vineyarders perceive tourists as a threat to their independence. Fishers, who are descendants of English families and live in Chilmark, are the most resistant to tourists. They are the most economically stable group since their occupation preserves its value as it makes up a large part of the island's economy. This explains the high level of centralized diphthongs in Chilmark as well as in the upper island.

Although all ethnic groups showed high frequencies of centralization, they utilized this shift to achieve different aims. Labov pointed out that the main purpose behind centralized diphthongs among English people was to declare their resistance against tourists and to assert their possession of the island. Subsequent Portuguese and Indian immigrant groups, however, used it to prove their status as natives to the English group. The data reveal that social stratification is understood in local ways. This indicates that broader socioeconomic correlations are not the immediate result of education, income, and occupation, but they mirror local dynamics established deeply and firmly in practices and ideologies that are shaped by class.

This ethnographic approach revealed the need for a more nuanced way to study style, which was the objective of Bell (1984).

#### **2.1.2.2.2. Bell's audience design**

Bell (1977) studied how news is broadcasted differently on New Zealand's radio stations. He noticed that the speech style of the same broadcaster, who reads the same news, changes from one radio station to another. This variation cannot be attributed to the amount of attention paid to speech, which it is the traditional explanation followed in the Labovian approach (Labov, 1972), since the broadcaster applies the same strategy that is supposed to produce a careful speech, yet the style varies from one station to another.

According to Bell, there is a need for a framework that can rigorously account for this variation. It should explain the relation between style dimension, or "intraspeaker variation", and social dimension or "interspeaker variation". To find a base for such a framework, he proposes the term style axiom, which states that variation in the style dimension is derived from variation in the social dimension. In light of this axiom, Bell (1984) proposes audience design. The concept of this framework is that "the speakers take most account of hearers in designing their talk" (p. 159). Speakers primarily adjust their speech to their audience. It is important to mention that the effect of the intended hearer on the style of the speaker has long been a topic of interest in sociolinguistics. Hymes (1974) generally considers audience as an influencing factor among other factors such as setting, genre, and topic. In 1975, Giles and Powesland presented their influential accommodation theory, which brings audience more to the center of style variation research. The centrality of audience reaches its optimum in Bell's approach (1984), where it is treated as a main factor in stylistic variation, whereas other factors are viewed as secondary.

For Bell (1984), style shifting is either responsive, which can be studied through audience design, or initiative, which can be investigated using Referee Design. Starting with audience design, Bell divides the audience into four groups based on three principles: being addressed (directly spoken to), ratified (their presence is acknowledged), or known (known to be part of a speech context). The speaker is the center of the speech and called *first person*. The *second person*, or *the addressee*, is the listener that is addressed, ratified and known. Anyone who is not addressed by the

speaker but ratified and known is called *the auditor*. The third character in the audience is *the overhearer*- anyone who is neither addressed nor ratified but known to the speaker. The fourth component of Bell's audience is the *eavesdropper*, who does not meet any of the three principles, i.e., not known, ratified, or addressed. The last three categories, *auditors*, *overhearers*, and *eavesdroppers* are classified as the *third person*. In addition to audience variables, setting and topic are listed as examples of non-personal factors that are believed to play a lesser role in style variation. It is important to mention that both audience and non-audience factors are classified as responsive.

As for initiative style-shifting, Bell (1984) proposes that, in addition to responding to an immediate audience, speakers are considered to diverge from their present audience and to try to accommodate to a non-present audience called *referees* through referee design. In such situations, speakers can initiate a shift in style that is associated with their absent in-group or an out-group that they wish to identify. Bell (2001) revisits audience design to rework some principles so they can cover more areas of style shifting. Although audience design has been an influential model in studying linguistic variation, it is criticized for being reductionist in the sense that it relies mainly on the responsive dimension of style shifting and places less attention on the initiative dimension. To elaborate, the complex nature of a speaker's use of language to express identity is considered exceptional or occasional rather than essential. Bell agrees with this critique and says that, although initiative style shifting has never been neglected in Bell (1984), it is dealt with as a peripheral dimension. To overcome this issue, a more balanced view of the responsive and initiative dimensions is needed. Therefore, Bell suggests the integration of audience and referee designs in analyzing style variation. The reworked framework now views responsive and initiative dimensions to "be two complementary and coexistent dimensions of style, which operate simultaneously in all speech events" (Bell, 2001, p. 165).

In spite of the fact that Bell (2001) reworks audience design to give a more balanced view of responsive and initiative style shifting, he still believes that the initiative dimension is derived from the responsive dimension (Wolfram & Schilling, 2006). His reasoning on this derivative connection is that speakers are incapable of establishing social meaning for linguistic features. In fact, they create their own style by using linguistic features that already have social meaning. This point in particular has been challenged. Eckert (2004) challenges audience as underlying stylistic variation, arguing

that speakers, through their styles, can actually create new social meanings or alter meanings that already exist (p. 45).

Another limitation of audience design lies in the fact that it decides in advance the factors that affect speakers' choices of certain styles (Biber & Finegan, 1994). Bell (2001) responds to this point by saying that any framework with no "a priori categorization" cannot be falsified. He believes that one of the audience design's strengths is "its falsifiability" (p. 163). Schilling-Estes (2002) argues that the unidimensionality of Bell's framework will prevent researchers from capturing the complex nature of style variation in everyday communication. Schilling-Estes also criticizes audience design for its unclear framework that combines the new version of the communication accommodation theory (Giles, Coupland, & Coupland, 1991) with speaker design. She explains that audience design is similar to the accommodation theory in the sense that speakers are not restricted to linguistic features to show solidarity with their audience, but they also can use other "communicative strategies". As for bringing audience design closer to speaker design, Schilling-Estes (2002) states:

in recognizing that immediate conversational purpose is just as important as more permanent speaker characteristics in shaping speech style (as well as recognizing the importance of initiative style shift), they are moving the audience design framework in the direction of what we might call "speaker design" approaches, following Coupland (2000). (p. 388)

### **2.1.2.3.        *The third wave: speaker design***

The third wave comes to compensate for the previous absence of agency as an important element in studying variation (Coupland, 2000; Eckert, 2004; Zhang, 2005, 2008; Hernández-Campoy & Cutillas-Espinosa, 2012). In the process of styling, Coupland (2000) states that speakers should not be seen as passive or reactive but rather as having an active role where they utilize variation to project different versions of themselves. Therefore, researchers' attention should shift to the self, i.e. speakers, instead of texts. Moreover, stylistic variation should be perceived as individual variation because a speaker uses style to represent an identity that reduces or increases the social and cultural differences with another identity (the interlocutor's identity). Bucholtz (2015) believes that style should be recognized as a multimodal phenomenon and a dynamic system for identity construction.

Eckert (2012) argues that the first and second waves focused mainly on the structure of variation and failed to account for agency or how speakers are restricted by social structure in their everyday communication and what may be the potential relations that make that structure remain in place. The third wave does not devalue the pivotal role of structure but instead it encourages researchers to understand the relationship between structure and practice. In this relation, structure constrains practice which gives practice a very important role in producing and reproducing structure. Eckert believes that researchers will gain two crucial benefits when the focus is shifted to practice in variation studies. First, the meaning of variation will be the center of attention and that will enable us to understand, explore, and explain speaking but not speech, and styling but not style (Coupland, 2007). Second, realizing the meaning of variation will provide the needed tools to help study the real process of change. This wave has four key terms that most of the proposed theoretical frameworks share: *practice*, *indexicality*, and *stylistic practice*. Each of these key terms will be explained in the next sections.

#### **2.1.2.3.1. Practice**

Eckert (2009) uses her extended study on jocks and burnouts in a Detroit high school (Eckert, 1989) as one example of the third wave studies where the meaning of linguistic differences between those two groups of students in the suburbs of Detroit is achieved through a deep observation and examination of their everyday practices that shape who they are. The deep-rooted ideological concerns in speakers' everyday communication and experiences are indicators of the motivations behind the use of variation on both social and personal levels. In her analysis, Eckert employs Pierre Bourdieu's (1977) theory of practice as one means to interpret the data. According to this theory, the process of social reproduction has habitus in its center. The habitus is a set of behaviors, affects, and thoughts that a person acquires and improves during their life in a specific social position. The main role of this habitus is to act as a mediator between social structure and individual behavior and help shape, but not set, individual action.

In the light of practice theory, Eckert (1989) elaborately discusses the social background and practice of both jocks and burnouts to understand their distinctive linguistic patterns. It is proposed that, "language is part of a broader semiotic system that includes such things as clothing, territory, musical taste, activities, and stances" (Eckert,

2004, p. 47). The jocks come from middle-class families where they mostly depend on the care and services of their parents in their early age, even in their social and friendship relations through what is known as play dates. In contrast, burnouts are classified as part of the working class where neighborhood ties constitute a very important support system for kids and adults. Unlike the jocks, the burnouts usually blend in the neighborhood peer community at an early age. This helps the burnouts get to choose their friends before the age of schooling, whereas the jocks get the chance to make their friends in elementary school. Eckert (1989) goes further to the ideology of American school regarding friendship making, where students are provided with two resources: age-graded classes and extracurricular activities. Friendships in these environments are affected by competition and hierarchal relations, which perfectly suit the jocks but not the burnouts. The burnouts find school boring and cannot wait to leave it to enjoy the neighborhood networks, which extends later on to larger areas as they grow up.

This dynamic explains the behavior, beliefs, and values of both groups. In the school environment, working-class kids who want to be jocks frustratingly experience the lack of emotional support among the jocks. The jocks, on the other hand, who try to blend in the burnout community, find the tremendous emotional support among the burnouts to be awkwardly overwhelming. Eckert (1989) is interested in whom they talk to about their personal concerns and private problems. She finds that the jocks either bottle their problems up or reveal them to someone outside their school. As for the burnouts, they share their concerns with any of the burnouts. Eckert (2009) states that, “jocks develop habits of personal relations that are adaptive to a competitive corporate hierarchy, while burnouts develop habits that are adaptive to a solidary working-class milieu” (p. 15). Eckert refers to this as social reproduction, which is the center of Bourdieu’s (1977) theory of practice.

The distinction between jocks and burnouts is not only rooted locally but also in the socioeconomic landscape of the Detroit metropolis. This makes Eckert (1989) explore those groups on a sociogeographic level. To elaborate, she notices that the socioeconomic status gets increasingly perceived as we go farther from the city to the suburbs. The burnouts in schools that are adjacent to the city are considered more burn out than the burnouts farther in the suburbs and are highly respected by suburban burnouts. The jocks in the urban areas are seen more burn out when compared to the



jocks in the suburbs. Unlike the burnouts group, the situation is reversed with the jocks, in which the urban jocks are envious of the financial and political growth and stability of the suburban jocks. Grasping this distinction supports the notion that linguistic variation should be linked to social meaning proposed by Eckert (2004). From her data, the burnouts generally tend to use more urban variants than the jocks and it is social meaning that explains why urban burnouts and jocks use more urban variants than their suburban counterparts.

#### **2.1.2.3.2. Indexicality**

Besides the theory of practice, the third wave approach makes use of Silverstein's (1976, 2003) notion of indexical orders, which is a very crucial concept that explains the links between an individual's style and macrosociological factors in the process of stylistic practice. This notion helps explain the relations between linguistic features, styles, and social meaning. In indexicality, linguistic variables are not signs in themselves, but they are considered so once they are associated with social meaning. To elaborate, there is an initial stage where the speech of a certain population becomes salient and its linguistic features attract attention. In this stage, those linguistic features are called *first-order indexes*. After this speech is known as belonging to a specific group of people, its linguistic features can be recognized on their own even if they are extracted from their contexts. They basically become membership indexes of that specific population and can be used either by speakers from within this group to confirm their membership or by speakers outside the group to recall any related stereotypes. Those indexes are called *second-order indexes*. When those features are repeated, they create more effects and at some point they can be used for more indexical signs (Eckert, 2008, p. 463-465).

Eckert (2008) importantly mentions that this linear ordering of indexes does not mean that this process happens sequentially but rather can happen concurrently in many directions. Those indexes build a group of ideologically linked meanings that at any specific time establish an indexical field. By incorporating this notion in style investigations, researchers will deal with variables as having indexical fields that speakers use to make ideological changes.

Eckert (2012) criticizes the first wave interpretation of gender variation and how it is directly linked to the opposition between prestige and stigma. For example, the more standard linguistic pattern that females show is attributed to the idea that women need to give the image of being more educated in order to reflect high status. For Eckert, this does not seem to be accurate and it is built on speculations rather than on a grounded theory. She deals with linguistic variation among her female and male participants in the Detroit study by putting Silverstein's notion of indexicality into use. When one initially examines her data, they would be tempted to conclude that the burnout girls use more urban variants than the burnout boys because they need to reflect the image of being tough and rebellious more than reflecting their identification as burnouts. Therefore, one can say that urban variants such as negative concord index rebelliousness. The jock boys, on the other hand, show more urban variants in their speech compared to the jock girls. Why is the pattern reversed in the jocks' data? How could it be explained? Don't the jock girls need to prove their toughness as much as the burnout girls? Are the jock boys trying to express some sort of self-strength through urban variants? Again, it is only social meaning that can offer an explanation to these data.

In order to understand the indexical link between style and gender in Eckert's (1989) study, it is important to note that the two categories burnouts and jocks can be looked at as a continuum that places burnouts on one end and the jocks on the other. Not all students are classified as pure burnouts or pure jocks. There is a considerable number of students who find themselves in-between the polar categories. Eckert (1989) discusses the division among burnout girls, who are divided into two categories: the regular burnouts, and the burned-out burnouts. When looking at them from a social point of view, the regular burnouts build their identities following working class values such as friendship, loyalty, and school dissociation. The burned-out burnouts girls, on the other hand, are known for their rebellious behavior, and they separate themselves from the regular burnouts and view them with scorn. Knowing this fact helps explain the different correlations between the urban variants and social affiliation in those two clusters. The burned-out burnouts use much more urban variants compared to the regular burnouts. As for the jock boys who use more urban variants than the jock girls, Eckert found that the jock boys socialize more with the burnouts in the school's sport activities; therefore, they acquire some of their linguistic behavior. This example provides a piece of evidence that linking linguistic style to social meaning is the best way to interpret those stylistic

differences instead of the inadequate explanations such as the lack of monitoring or proving one's status. This example also shows the way Eckert handles gender differences in her data. She goes beyond superficially attributing linguistic variation among her subjects to gender difference; instead, she researches thoroughly the social affiliation and roles of each group to better understand their styles. In Eckert's (2009) words, "it makes it clear that gender differences in language do not index 'male' and 'female', but the qualities, activities, and stances that constitute a variety of ways of being 'male' and 'female' " (p. 19).

#### **2.1.2.3.3. Stylistic practice**

Style in the first wave is viewed from formality and informality perspectives, which are rejected by researchers of the third wave who believe that variation is used to create personal and social styles that are related to social types. They are also interested in the concept of stylistic practice as a bricolage procedure where people mix elements in creative ways to establish new meanings or alter old ones. Irvine (2001) states that styles are known for their distinctiveness, where each style can exclusively obtain its meaning through how it is connected with other styles (p. 22). Furthermore, Irvine stresses that those connections are ideologically mediated. Eckert (2012) agrees with Irvine's statements and adds that variationists should not focus only on the link between linguistic features and social categories, but they should try harder to interpret the social significance of those types of links.

Another important point Eckert (2012) stresses is the intentional stylistic choices which are ignored in the first and second waves, since the interest is entirely on unintentional stylistic choices. The third wave, however, finds those conscious stylistic choices appealing because they enable meaning construction through variation. Some linguistic styles are associated with stereotypes such as the Valley Girl style, which is associated with shallowness, materialism, and non-intellectuality (Bucholtz, Bermudez, Fung, and Edwards, 2007, p. 329). They emphasize that stereotypes, linked to some conscious stylistic choices, are not chosen randomly but they are chosen according to social hierarchies.

The most distinct aspect of Eckert's approach is its flexibility and depth in interpreting style variation. The idea of linking linguistic variation to social meaning provides researchers with more room to tackle and examine several instances that

previous approaches are unable to explain. Speaker design approach does not restrict variation to any pre-existing motivation until it goes under a deep investigation that puts speakers themselves on the spot. Style is considered as a means towards presenting one's image or persona to the world. Eckert (2004) proposes an approach that focuses on "speakers' linguistic performance as a continual construction of a persona (or personae), and variables as resources for this construction" (p. 41).

Eckert (2012) proposes a group of hypotheses that underline the theoretical foundation of the third wave. The first hypothesis states that variation conveys and voices the full scope of social concerns in a specific community. The second hypothesis insists that variation is not used to mirror, but to establish social meaning; therefore it is considered a power in social change. To reach this goal, researchers need to not investigate the process of variation generally, but as it is manifested in speech. The third hypothesis suggests that the meaning of single variables is not as distinctive as it is in the context of discourse and in the creation of styles. Therefore, variation studies have to focus on stylistic practice.

I now turn to a discussion of language usage in the Arab Gulf region in order to illuminate the stylistic resources that are available for speakers to use in the process of image projection.

## **2.2. Diglossia in Arabic**

Diglossia as a term was first found in the work of Marçais (1930) to explain incidences where spoken Arabic dialects and Classical Arabic are combined and used as a means of communication in the Arabic world. Classical Arabic is "the language of pre-Islamic poetry and of the Qura'n" (Abdul-Raof, 2013, p. 15). Diglossia was then reintroduced to the realm of linguistics by Ferguson (1959) to generally describe a linguistic behavior where two varieties of a certain language co-exist in a speech community. Arabic, Modern Greek, Swiss German, and Haitian Creole are classified by Ferguson as diglossic cases where each language has two varieties: one is perceived as a superposed version and called the high variety [H], and the other is a dialectal version called the low variety [L]. Regarding the functions of those varieties, H is usually preferred in formal situations, whereas L is seen as the one and only appropriate

medium of informal communication. The following list shows the occasions where it is expected to use either H or L, as indicated by Ferguson (1959, p. 329).

<b>H</b>	<b>L</b>
Sermon in church and mosque	Instructions to servants, waiters, workmen and clerks
Political speech	Conversation with family, friends, and colleagues
University lecture	Radio soap opera
News broadcast	Caption on political cartoon
Newspaper editorial, news story, caption on picture	Folk literature

Since this paper mainly focuses on the Arabic language, the presented literature will tackle diglossia in the Arab world only. Wardhaugh and Fuller (2015) classify Classical Arabic, or Al-Fusha, as the H variety and the other regional colloquial dialects, called Al-ammyiah, as L varieties. This classification cannot be discussed without considering prestige. For instance, Classical Arabic is believed to be superior by nearly all of the Arabs and this attitude is driven from the fact that this H variety is the language of the holy scripture of Islam, the Qur'an. Another reason for the prestige of Classical Arabic is that most of the admired literature was written in this variety. Besides prestige, diglossia can be also defined through an array of criteria that accurately draw a clear line between the two varieties of Arabic. Those criteria are acquisition, standardization, grammar, lexicon, and phonology (Ferguson, 1959; Kaye, 1970; Saville-Troike, 1982; Hudson-Edwards, 1984).

In terms of acquisition, although children might come into contact with Classical Arabic in their early age by different means such as TV programs, they do not practically acquire it until they get into school, where it is taught to them as an L2. Even after spending years studying the grammar and vocabulary of Classical Arabic, not all speakers will be able to use Classical Arabic proficiently. This is due to the fact that this variety is restricted to formal situations that not all speakers are put in. In contrast with Classical Arabic, colloquial Arabic is the variety mainly used in the household and on a daily-basis in interactions. Being exposed to colloquial Arabic enables children to acquire this variety in "the normal way of learning one's mother tongue" as stated by Ferguson (1959, p. 313).

Another characteristic that describes the diglossic nature of Arabic is the idea of standardization. Scholars notice that the H variety, i.e., Classical Arabic, is well established when it comes to the study of its grammatical structure, pronunciation,

vocabulary, and orthography. This systematic approach in dealing with Classical Arabic makes it an established norm that permits very limited variation. In contrast, colloquial Arabic lacks this feature of standardization and variation in all levels of linguistic structure occurs due to several factors such as regional differences (Ferguson, 1959). For example, each country in the Arabic-speaking world has a distinctive L that is considered the standard L and is seen as an indicator of the speaker's origin like Egyptian Arabic in Egypt, Moroccan Arabic in Morocco, and Gulf Arabic, which is the variety spoken by most of the speakers in the Arabian Gulf countries.

As for grammar, scholars recognize that colloquial Arabic has a simpler grammatical system as compared to the H variety Classical Arabic (Ferguson, 1959; Kaye, 1970). This is exemplified by noun cases and duality in adjectives and verbs that exist in Classical Arabic but not in colloquial varieties. Nouns in Classical Arabic have three cases that are shown by marks on word endings. The word *house*, for instance, is /baytu/ in the nominal case, /bayta/ in the accusative case, and /bayti/ in the genitive case. Those marks are completely absent in colloquial Arabic, and the equivalent of *house* in Cairo Arabic, for example, is one word /beet/. The following examples are taken from Kaye (1970, p. 384) on the syntactic and morphological differences between Classical Arabic as the H variety and Cairo Arabic as the L variety:

**Classical Arabic**

- (1) /jaaʔa lwaladaani/
- (2) /aibaytaani lkabiiran/
- (3) /kitaabu lmaliki/

The two boys came.  
The two big houses.  
The king's book.

**Cairo Arabic**

- /gum ilwaladeen/
- /ilbiteen ilkubaar/
- /ilkitaab bitaʕ lmalik/

In (1), the verb /jaaʔa / in Classical Arabic agrees with the subject /lwaladaani/ in duality. This agreement is not present in the Cairo Arabic example, where the verb /gum/ is plural and the subject /ilwaladeen/ is dual. As for (2), the Classical Arabic example shows agreement in duality between the noun /aibaytaani/ and the adjective /lkabiiran/. Again, the agreement is not needed in the Cairo Arabic sentence, where the noun /ilbiteen/ is dual, but the adjective is plural /ilkubaar/. Example (3) shows the difference between the two varieties in expressing possession. In Cairo Arabic, the particle / bitaʕ/ corresponds to *of* in English and is used to illustrate possession. This particle has equivalents in other L varieties of Arabic, such as /maal/ in Baghdadi Arabic and /tabaʕ/ in Palestinian Arabic. Classical Arabic, however, does not have this particle at all. Syntactically speaking, nouns in (1), (2), and (3) are strictly marked with specific cases in

Classical Arabic sentences, while those declensions are not required in the Cairo Arabic sentences.

Lexical distinction is a feature that Ferguson (1959) believes to be a strong indicator of a diglossic situation. Vocabulary is shared between H and L varieties to some extent, but there are words and expressions that are found in one variety but not in the other. For example, H has some technical words that have no equivalents in L simply because those words are only used in H domains and never discussed in L domains and vice versa. However, what are called *lexical doublets* also exist, in which two words, one in H and its equivalents in L, are deployed to indicate the same concept or object. Ferguson (1959) states that:

A striking feature of diglossia is the existence of many paired items, one H one L, referring to fairly common concepts frequently used in both H and L, where the range of meaning of the two items is roughly the same, and the use of one or the other immediately stamps the utterance or written sequence as H or L. (p. 343)

An example of those paired items is found in Suleiman's (1985) study of diglossia in Jordanian Arabic, where he lists the lexical doublets observed in the data of students of Yarmouk University, a public university in northern Jordan. For instance, /uriidu/, /aðhaba/, and /likay/ in Classical Arabic correspond with /biddi/, /aruuħ/, and /minjan/ in Jordanian Arabic, which mean 'I want', 'to go', and 'in order to', respectively (Suleiman, 1985, p. 43).

The last characteristic of diglossia is phonology, which Ferguson claims to be 'moderately different' between Classical and colloquial Arabic. He asserts that "the sound systems of H and L constitute a single phonological structure of which the L phonology is the basic system and the divergent features of H phonology are either a subsystem or a parasystem" (Ferguson, 1959, p. 335). Kaye (1970) disagrees with Ferguson's assertion that says that H and L varieties in Arabic have a single phonological system. He, however, attributes this phonological gap between the two varieties of Arabic to the idea that Classical Arabic has an ill-defined phonological system compared to the well-defined system in colloquial Arabic. To elaborate, Kaye explains that the single word /raaħ/ 'he went' in Syrian Arabic has many SA versions, which range from most to least acceptable /ðahaba/, /ðahab/, /dahab/. Salih (2015) also expresses an opinion against Ferguson and asserts that the phonological variation is

actually significant after examining data from Iraqi Arabic. Salih states that there are at least six phonological aspects that are found only in Iraqi Arabic. Those aspects are three consonants and three vowels. An example concerning the consonant aspects is the initial consonant clusters in colloquial Iraqi Arabic, such as the initial cluster in /ʃlo:n/ and /ʃbīk/ which does not exist in the Classical Arabic equivalents /māḏā/ and /mābik/. Those mean ‘how’ and ‘what is wrong’, respectively. An example concerning the vowel differences will be the four extra vowels found in Iraqi Arabic: /ɔ:/, /aii/, /aia/, and /ea/ (Salih, 2015, p. 110-114).

### **2.2.1. Modern standard Arabic vs. colloquial Arabic in the Arabian Gulf region**

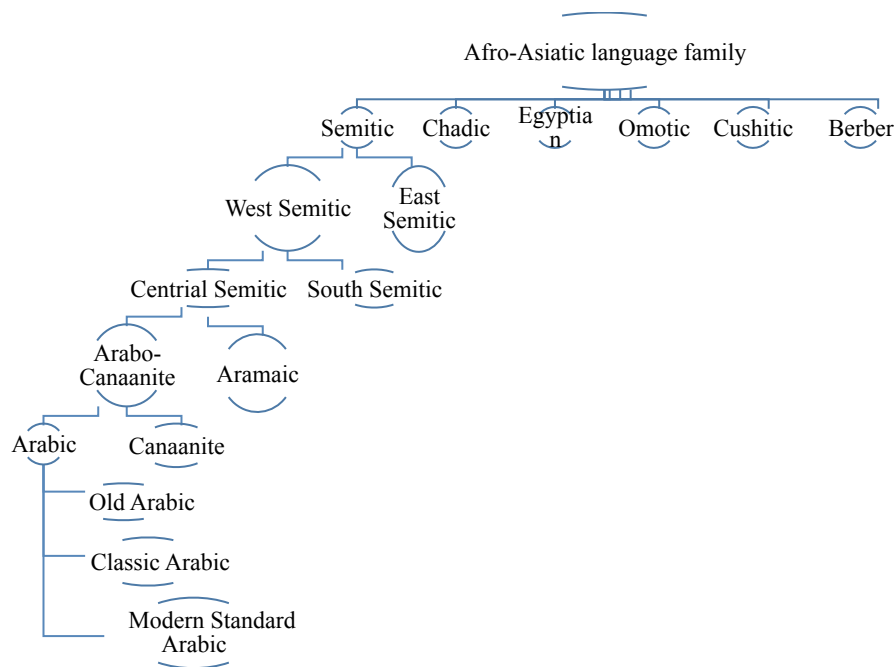
In order to better understand diglossia in a specific region, the varieties and their roles should be taken into consideration. Non-Arabic speakers have the impression that Arabic is a single monolithic language that is spoken by all Arabs in the Arab-speaking countries. To some extent, this assumption is true, as there is a common Arabic that is used and understood by most educated people in the Arab world. This common or formal Arabic is called Modern Standard Arabic (MSA), hereafter MSA, and it is the main language that is taught in school, used in the media, and exclusively adopted in officially written documents in more than 20 countries in the Arab world (Badawi, 2006; Bassiouney, 2009). What is misleading in the previous assumption is the notion that MSA is the only Arabic used for both formal and everyday communication. The unknown fact to most non-Arabic speakers is that MSA is rarely, if ever, used in informal linguistic environments, such as family or friends (Adouane & Johansson, 2016). Colloquial Arabic dialects are used instead. Habash (2010) best described those dialects as “the true native language forms” that “are generally restricted in use for informal daily communication” (p. 1).

This other form of Arabic that is called colloquial and is claimed by some researchers to be the native language of the majority of speakers in the Arab world raises an interesting question: how did Arabic varieties develop? To answer this question, linguists have followed different approaches by classifying those varieties genealogically, chronologically, geographically, sociologically, and linguistically (Cuvalay-Haak, 1997). Genealogically speaking, Arabic is part of the Central Semitic group that is descended from the Afro-Asiatic family of world languages (Katzner, 2002).



On a more narrow scope, the Arabic language family consists of all the varieties that are descended from Proto-Arabic (Figure 1). Shvitiel (1991) classified the descendants of Proto-Arabic chronologically by marking three main periods. According to his classification, Old Arabic was dominant up to the beginning of the seventh century, Classic Arabic was used from CE 600-1800, and Modern Arabic started from CE 1800 and is used to the current day (p. 1441). Versteegh (2014) points out that Old Arabic was spoken in the Arabian Peninsula, Classic Arabic was spoken in the Arabian Peninsula and gradually extended to the regions outside its borders as Islam continued to spread from Spain to India, and Modern Arabic represents the language spoken in the Arab countries today.

Based on this chronological order of Arabic, the term Classical Arabic will be substituted with the term Modern Standard Arabic, MSA. Bassiouney (2009) discussed the difference between the two terms, in which Classical Arabic refers to “the religious language of the Qur’an.... or ...older classical texts”, whereas MSA is the High variety “used in public speech” (p. 12). For Ferguson (1959), MSA is the outcome of native speakers attempting to speak Classical Arabic. Bateson (2003) argues that MSA is believed to stem from Classical Arabic and to have simple syntax, different lexicon, and distinct style.



**Figure 1. The genealogical tree of Arabic**

The genealogical tree in Figure 3 shows only the formal varieties of Arabic chronologically. This does not completely answer the question about how Arabic varieties developed. Another specific question would be: how did the Arabic dialects develop in particular? Cuvalay-Haak (1997) states that in the seventh century Arabic underwent massive linguistic changes due to the rapid spread of this language outside the borders of the Arabian Peninsula during the expansion of the Islamic Empire. He believed that “these linguistic changes were either accelerated or provoked by the numerous language contact situations in the new empire” (Cuvalay-Haak, 1997, p. 6). Zaidan and Callison-Burch (2014) divided Arabic dialects into five dialect groups according to geographical location: Moroccan, Egyptian, Gulf, Levantine, and Iraqi. Since the scope of this study is restricted to the Arabian Gulf region, the focus here is on Gulf Arabic. The term Gulf Arabic covers all colloquial varieties that are mutually intelligible and spoken in the Gulf’s six states: Kuwait, Qatar, Bahrain, Oman, United Arab Emirates, and Saudi Arabia (Habash, 2010; Khalifa, Habash, Abdulrahim, & Hassan, 2016). Some scholars believe that Gulf Arabic is used only in the eastern part of the Arabian Peninsula, which includes the previously mentioned countries and only the eastern region of Saudi Arabia (Holes, 1990). No one denies that those spoken dialects have differences, but “the fact that these differences exist, ... should not be allowed to obscure the fact of their essential unity as a dialect group” (Johnstone, 1967, p. 18). In spite of the morphological and phonological differences, it is argued that Gulf Arabic is the closest variety to MSA because several grammatical rules and much lexicon are shared (Holes, 1984; Versteegh, 2014). To clearly explain the relationship between MSA and Gulf Arabic, Hary (1996) proposed a continuum with MSA at one end and colloquial Arabic at the other end. Although the two ends represent either pure MSA or pure CA, these idealized varieties do not actually exist. What exists is MSA mixed with some colloquial lexicon or vice versa (Hary, 1996, p. 72).

It has been noticed that the literature on Gulf Arabic is not as rich as that of other dialects of the Arab world. Most of the studies that have been conducted have been about the grammar or basics of Gulf Arabic (Qafisheh, 1975, 1977; Holes, 1984, 1990). Holes (2004, 2011) shed some light on the history of language change and dialects formation in the Gulf region. One of the most important factors that influenced this region linguistically was the migration that took place 15 centuries ago in the eastern part of Arabia. There were two kinds of migration. One movement was from the central and

northern parts of Arabia, as the Bedouin tribes moved east for a better life. The second kind of migration to the Gulf region was driven by commerce and was from other countries, such as Persia and India, by sea to the area. Those movements resulted in the integration of the Bedouin's dialect brought from central and northern Arabia with the languages of the overseas migrants. This linguistic change also influenced the dialect of the people who were already settled in the area. Those changes that have been introduced to the local dialects yield more colloquial forms that eventually constituted what is known today as Gulf Arabic (Holes, 2011, p. 131-132).

### 2.2.2. Some linguistic characteristics of Gulf Arabic

The crucial question is how can MSA be distinguished from Gulf Arabic (GA)? A very recent corpus study showed that vocabulary, semantics, and syntactic structures are the most salient indicators that enable researchers to distinguish MSA from GA (Adouane & Johansson, 2016). Starting with vocabulary, there are two defining characteristics of GA's lexicon: "the use of Arabized English (English written in Arabic script) and the use of special vocabulary" (p. 2711). This can be clearly seen in the following examples taken from Adouane and Johansson (2016, p. 2711):

هذا المطعم إدمان يا ويلي ع الطعم الشاورما حقهيم تبع تكشن مرة لذيدة و سبايسي روعة  
و ساندويتشاتهم خرافية خصوصا بحطون تشيز كثير فرش عن جد برافو عليهم! يمبيي و هذي عشرة

(1)

haða	il-matʕam	idman	ya	weili	ʕala
this	Def-restuarant	addiction	oh	wow	on
i-tʕaʕam	i-ʕawirma	ħaħum	tabaʕ	<b>chicken</b>	marah
Def-taste	Def-shawatma	their	for	chicken	very
laðiðah	wi	<b>spicy</b>	rawʕah	wa	<b>sadwitch-at-hum</b>
delicious	and	spicy	wonderful	and	sandwitch-PI-their
xurafiah	biħitʕ-un	<b>cheese</b>	kti:r	<b>fresh</b>	ʕan
unbelievable	put-3PI	cheese	lots	fresh	as
dzid	<b>bravo</b>	ʕaleihum	<b>yummy</b>	wa	haði
fact	bravo	on-them	yummy	and	this
ʕaʕrah					
ten					

'This restaurant is an addiction. The taste of its **chicken** shawarma is very delicious and **spicy**. Its **sandwiches** are unbelievable; especially they put lots of **fresh cheese**. Truly, **bravo yummy** and these are ten stars for you.'

(2)

مطعم حلو مره و كتيير أروح له ابي اروح له الحين بس اخر زيارة له ارتفع ضغطي ما ادري عشان في زحمة  
أو عشان الناس باجازة و منشفحين على الأكل و المطاعم كان زحمة و خرسة مو مثل أي يوم

matʕam	ħilu	marah	w	kti:r	aru:ħ
restaurant	good	very	and	frequently	go-1Sig
li-h	<b>abi</b>	aru:ħ	li-h	<b>alħi:n</b>	bas
to-it	want-1Sig	go-1Sig	to-it	now	but
axir	ziyarah	li-h	irtifaʕ	dʕaytʕ-i	ma
last	visit	to-it	Increase-Past	blood pressure-my	Neg
adri	<b>ʕaʕan</b>	fi	zaħmah	aw	<b>ʕaʕan</b>
know-1Sig	because	there	crowd	or	because
i-nas	bi-dʕazah	w	<b>minʕafħi:n</b>	ʕala	al-akil
Def-people	on-vacation	and	voracious	on	Def-food
w	al- matʕaʕim	kan	zaħmah	w	<b>xriʕah</b>
and	Def-restaurant	was	crowded	and	busy
mu	miθil	ay	yum		
Neg	like	any	day		

'The restaurant is **very** nice and I visit it frequently. **Now I want** to go there but at the last visit I got high blood pressure, I'm not sure if it was **because** of the long queuing or **because** people are on vacation and are **craving for food** and restaurants. It was crowded and **busy** not as any other day.'

In example (1), the Arabized words are bolded in both the example and the translation. Example (2) shows the "special vocabulary" or in other words the dialectal vocabulary. Semantically speaking, those dialectal words are divided into two main groups based on whether they are in the MSA dictionary or not. Regarding the dialectal words that exist in the MSA dictionary, some of them have either a "conflicting part of speech (PoS) or totally a different meaning" or both. For instance, the word [abi] in example (2) is used in MSA and GA, but its part of speech and meaning are completely different in each variety. It is a noun in MSA and means [father], but a verb in dialectal Arabic that means [I want]. Another interesting case regarding MSA and GA is the "conflicting vocabulary where the same word form has exactly the same part of speech between MSA and Gulf Arabic but with the opposite meaning." (Adouane & Johansson, 2016, p. 2711). The words [rahi:b/ terrible] and [xati:r/ dangerous] in the following example (3) have the same part of speech in MSA and Gulf Arabic, in which they are both adjectives. However, the dialectal versions of those words are opposite in meaning

to their MSA's counterparts. The two words convey a positive connotation in Gulf Arabic but a negative one in MSA.

(3)

الشاورما حفته رهيبه و طعمها خطير

i-ʃawirma	ħagitiħ	<b>rahi:bah</b>	wa	ʔaʕam-aha	<b>xatʔ:r</b>
Def-shawarma	his	terrible	and	taste-its	dangerous

'His shawarma is terrible and its taste is dangerous.'

(Adouane & Johansson, 2016, p. 2712)

In addition to vocabulary and semantics, Adouane and Johansson (2016) discussed the syntactic differences between MSA and Gulf Arabic. They stated that MSA permits a relatively free word order. This can be exemplified by topicalization cases where changes occur to the order of the sentence elements in order to express emphasis. In Gulf Arabic, "this free word order is even loose since there is no standard structure." (p. 2712).

Besides word order, negation in Gulf Arabic is different than in MSA. In MSA, there are two ways to negate. One is to use a negative word such as /akrah/, which means 'I hate'. The other way is to use some particles in order to reverse a positive expression. The tense of the verb in MSA determines which particle to use. For instance, /la/ is used if the verb is in the present tense, /lam/ for the past, and /lan/ for the future. In Gulf Arabic, the word's part of speech determines which negation particle to use. In addition to the negation particles used in MSA, GA has the extra particles [mu, muħub, miħ, muħ, mu] that are exclusively used to negate prepositional phrases, pronouns, adjectives, adverbs, and nouns (Qafisheh, 1977, p. 238-242).

### 2.2.3. Studies in diglossia

Variationists have examined diglossia in several regions of the Arabic-speaking communities (Harrel & Blanc, 1960; Al-Toma, 1969; Suleiman, 1985; Bentahila, 1991; Dahir, 1998; Al-Saidat, 1999; Tamimi and l'lawi, 2006; Al-Saidat & Al-Momani, 2010; Stadlbauer, 2010; Al-Sobh, Abu-Melhim & Bani-Hani, 2015; Qudah, 2018). It is important

to mention that a considerable amount of literature has been published on four colloquial varieties of Arabic: Jordanian, Egyptian, Syrian, and Moroccan Arabic. Al-Saidat and Al-Momani (2010) examine the grammatical differences between MSA and Jordanian Arabic, in particular, the use of future markers, and find that the markers used in Jordanian Arabic differ from those in MSA. Moreover, those forms, in contrast to MSA, are inflected to express grammatical categories such as person, number, and gender.

Some scholars find mass media a fertile source for data that enables them to describe diglossia (Alshamrani, 2008; Bassiouney, 2009; Albirini, 2011; Ababtain, 2013; Dashti & Dashti, 2015). Alshamrani (2008), for instance, focuses on the language used in three TV stations: Aljazeera, the Lebanese Broadcasting Cooperation (LBC), and Arab Radio and Television (ART). He finds that two varieties of Arabic were used on different frequencies by both the broadcasters and the guests in the programs televised on those stations. The study indicates that the type of program on all three channels affects the nature of diglossic switching. Political, religious, and documentary programs show a high-frequency use of MSA. However, the overall use of MSA drastically decreases in songs, series, and news reports on movie stars. Interestingly, Aljazeera has the highest percentage of using MSA (83%), while the overall use of MSA in LBC and ART are (17%) and (15%) respectively.

Another recent qualitative study was conducted on a limited population, where the language of Kuwaiti presenters in nine Kuwaiti TV channels undergoes a thorough investigation (Dashti & Dashti, 2015). The results reveal that there is a strong link between the program's type and the used variety, which is entirely consistent with Alshamrani's findings (2008). Again MSA is spoken in the news, political programs, and religious discussions. In contrast, CA dominates in sports and fashion programs. What is new in the results of Dashti and Dashti (2015) is the emergence of English as a variety from another language in the data where CA is highly used. The next section will discuss the literature concerning the role of English in the Arabian Gulf region and some examples of recent research on English borrowings in the Arabic-speaking communities.

### **2.3. The role of English in the Arabian Gulf Region**

The classification of English in speech communities has been a heated topic in the literature. Görlach classifies English based on "the functional range and the norms of

correctness” as a foreign language (EFL), a second language (ESL), a native language (ENL), or a second dialect (ESD) (1998, p. 19). Smith (1976/2015) questions such conceptualizations and believes that those terms do not accurately capture all English users and uses. He proposes the additional concept of English as an international auxiliary language (EIAL). According to Smith, English is an international language that speakers use to communicate with each other worldwide. It is also an auxiliary language that speakers of the same country use, besides their L1, for internal communication (Smith, 1976/2015, p. 159). Smith insists that English should not be considered as an exclusive language to native speech communities, but as an inclusive language that belongs to every speech community that uses it.

The notion of English as a language that “belongs to the world” (Smith, 1976/2015, p. 160) was introduced first in Kachru’s (1966) work on Indian English, where he noticed that there are specific characteristics of the variety of English used in India. Kachru (1985) then introduced the World Englishes paradigm (WE), which consists of three circles of English use: the Inner Circle, the Outer Circle, and the Expanding Circle. Those circles represent the historical stages of English spread and are based on three principles: “the types of spread, the patterns of acquisition and the functional domains in which English is used across cultures and languages” (Kachru, 1985, p. 12). The Inner Circle is composed of mostly monolingual speech communities where English is used as a primary language such as the United Kingdom, the United States, Canada, Australia, and New Zealand. The Outer Circle includes all speech communities that were colonized under British imperialism and now use English as an additional language. English has an important role in terms of communication and language policies in countries such as India, Nigeria, and Zambia. The last circle is the Expanding Circle that covers speech communities where English historically was a foreign language but more recently has been gaining uses and spreading to an increasing number of domains. English reaches speech communities of this circle by contacts and interactions across national boundaries. China, and Indonesia are considered part of the Expanding Circle.

As discussed in Hilgendorf (2015, 2018), the World Englishes paradigm and its underlining concept of plurality of language have been criticized for being marginalizing and segregating. Hilgendorf (2018) states that critics:

seemingly argue that acknowledging such different macro-level English-using communities lends itself to drawing comparisons. Dynamics and tensions between communities then can become more prominent, which in turn can heighten tacit, if not explicit, hierarchies. Such hierarchies can lead to marginalization, bringing advantage and privilege to some communities at the expense of fair and equal treatment for all populations. Speech communities with greater power, prestige, and/or status may gain attention and recognition, assuming a more central position in the constellation of English-using communities. (p. 476).

Hilgendorf (2015) responds to the criticism by laying out the fundamental concepts in sociolinguistics for the plurality of language in general and the WE in particular. Hilgendorf argues that English is a tool that speakers use to communicate and express their thoughts in a meaningful way within their speech communities in the Outer and Expanding Circles; therefore, it is essential to focus on speakers in our research. Besides focusing on speakers, research on language must be carried out in its social context, which is the orientation of Labov's (1963, 1966, 1972) studies on linguistic variation. The concept of language as a social phenomenon that needs to be studied in its social context reveals its collaborative dimension, where language is viewed as a code shared and used among a group of speakers. Speakers on a collective level play a pivotal role in the dynamics of language use where they sustain, reproduce, and change codes (Hilgendorf, 2015, p. 57-58).

When the collective-level interaction results into some common features that become over time indicators of a certain group of speakers, it is referred to as speech community. Hilgendorf (2015) states that speech community is another key concept for plurality of language and the WE paradigm, and sheds some light on studies that point out the characteristics of this concept (p. 58-59). First, since language is a social phenomenon, any speech community should be treated as a social entity instead of a linguistic one (Hymes, 1974, p. 47). Second, speech communities must be viewed in terms of plurality because an individual of a certain community can be also an individual of many other communities at the same time (Bolinger, 1975, p. 333 ). Third, speech communities can be geographically independent and their individuals are not restricted to either one variety of a language or even its standard form (Firth, 1959, p. 208) . By listing the sociolinguistic fundamental concepts for language plurality and the WE paradigm, Hilgendorf (2015) shows that in order to understand English varieties, they should be studied in their contexts and their speakers should be the focus of variation studies. Hilgendorf (2018) asserts that, "the social reality of plurality in language use is



irrespective of how much relative power and influence a speech community may have” (p. 478).

According to Kachru (1985), the countries of the Arabian Gulf region are part of the Expanding Circle. Since his study, several changes have happened to the role of English in the countries of this Circle due to mobility, globalization, and the effect of online social networking that opens doors to international communication (Winford, 2002, p. 30). For example, English in the United Arab Emirates is now used as an acrolectal lingua franca as a result of the high rate of immigration to the country between 1990- 2010 (Boyle, 2012). Another important change that is believed to influence the role of English in the Gulf Cooperation Council (GCC) countries is the new movement towards using English as the medium of instruction in most of the countries' higher educational institutions (Ahmed, 2010). Regarding the status of English, it is stated that “Kuwaitis view English as a prestigious language .... They use it as a sign of social glamour. Overall, English is becoming an increasingly important language in the Arab world...” (Dashti, 2015, p. 33). Fussell (2011) believes that English has spread across the Gulf region over a short period of time and the increased use of English has led to the emergence of a new variety, Gulf English. Those rapid and continuous changes of English uses in the Expanding Circle indicates a shift for this language from being a language of limited purposes to an additional language (Hilgendorf, 2018, p. 480).

The change of English's role in the region and the evolution of online social communication have given rise to some linguistic behaviors such as codeswitching. The next section will present a review of recent literature on the topic of codeswitching.

## **2.4. Codeswitching**

Some researchers, such as (Ferguson, 1959), restrict diglossia to mixing between two varieties of the same language, but others include varieties from even different languages (Fishman, 1967). In an attempt to differentiate between Ferguson's diglossia and Fishman's diglossia, Heath (1989) proposes the term 'diglossic switching' in addition to the existing term 'codeswitching'. The former term covers Ferguson's concept of diglossia, while the latter covers the part related to switching between languages in Fishman's definition. There is no consensus either on the definition or the terminology of codeswitching in the literature. Bassiouney (2009) believes that the

definitions offered by Gumperz (1982) and Myers-Scotton (1993a) are more flexible and accurate to some extent in describing this verbal behavior. Gumperz states that “Conversational codeswitching can be defined as the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems” (Gumperz, 1982, p. 59). Myers-Scotton (1993a) defines codeswitching as the “use of two or more languages in the same conversation, usually within the same conversational turn, or even within the same sentence of that turn” (p. 47). Both Gumperz and Myers-Scotton see codeswitching as the situation where a speaker would choose a language over another, a variety over another, or a style over another.

Scholars such as Poplack (1980) and Myers-Scotton (1989) distinguish between two main types of codeswitching. One is inter-sentential, which occurs outside or at the boundaries of sentences or clauses. The other is intra-sentential, where the switch happens between words on the sentence or clause level. The intra-sentential codeswitching has two subcategories: tag switching, and intra-word switching. Tag switching is to switch either a word or a tag phrase or both from a language to another. Intra-word switching, on the other hand, happens at morpheme boundaries. The following examples are taken from Ismail (2015, p. 5) for more clarification:

#### (4) Inter-sentential codeswitching

raʒʕat-i:	min	ʔəmri:kah	ka:nat	marah	mutʕibah
Return-my	from	America	was	very	tiring

The weather conditions delayed flights for like almost twenty-four hours.

‘My return from America was very tiring. **The weather conditions delayed flights for like almost twenty-four hours.**’

#### (5) Intra-sentential codeswitching

qadam-t	ʕala	<b>one month</b>	iʒ a:zah	fi	sʕeif
submit-l	on	one month	vacation	in	summer

wə	ŋʕa:	allah	jagbəl-u:n-ha:
and	willing	God	accept-they-it-Fem

‘I applied for a **one month** vacation in summer and God willing they will accept it.’

### **2.4.1. Unit of analysis in codeswitching**

There are several theoretical frameworks that have been postulated to predict where codeswitching might happen and to explain any grammatical constraints that regulate this process. Some of these approaches are constraint-based, such as the Two Constraints theory by Poplack (1981), the Government Principle theory by Di Sciullo, Muysken, and Singh (1986), and the Matrix Language-Frame model (MLF) by Myers-Scotton (1993a). Other approaches are not constraint-based, such as the Constraint-free approach by MacSwan (1999). The most influential framework in the field of codeswitching studies is Myers-Scotton's (1993a) Matrix Language-Frame model (MLF). This model is based on the distinction between the matrix language (ML) and the embedded language (EL). The ML is mainly in charge of supplying the morphosyntactic structure of the sentence or clause. The EL, on the other hand, provides content morphemes to be embedded into the ML.

Another important aspect of MLF model is that it uses the CP (complementizer phrase) as the basic unit of analysis, rather than the sentence because "even within a sentence, the grammars may not be in contact" (Myers-Scotton, 2002, p. 55). Other researchers, however, argue that the sentence is the best unit of analysis (Jacobson, 2001; Bentahila & Davies, 1998). Jacobson's reasoning is that CP would be suitable as the analytic unit when "code-switching operates according to two mechanisms, one in which L1 dominates L2, and the other where L2 dominates L1", but in some cases a sentence should be considered when "three mechanisms emerge, that is (a) L1 dominates L2, (b) L2 dominates L1 and (c) neither dominates the other." (Jacobson, 2001, p. 71).

In this study, the unit of analysis is the utterance, because the diglossic situation of Arabic together with the increased usage of English gives rise to the possibility of both intra-sentential, as well as inter-sentential switching between the three codes. Even though I am not adhering to the MLF model, there are (as I will show in the results) some patterns of switching that raise interesting questions about its proposals.

#### **2.4.2. Codeswitching or borrowing? The terminology dilemma**

The terminological distinction between codeswitching and borrowing seems to be problematic for many scholars in this field. Views in the literature are confusing, as some consider the two terms to be different from each other, while other linguists assert that borrowing is a synonym for codeswitching. Haugen (1956) proposes the idea of a continuum where codeswitching represents the maximal distinction and borrowing represents the minimal one. According to him, borrowing is “the regular use of material from one language in another so that there is no longer either switch or overlapping except in a historical sense”, whereas codeswitching happens when “a bilingual introduces a completely unassimilated word from another language into his speech” (p. 40).

Lipski (2005), in his summary of switch/borrowing typology, believes that the borrowing process is always conscious and deliberate, while codeswitching can be conscious in certain situations and unconscious in others. Bentahila and Davies (1983) propose two criteria to differentiate between codeswitching and borrowing. The first criterion restricts codeswitching to bilingual speakers, whereas borrowing could be found in the speech of both bilinguals and monolinguals. The second criterion posits that in borrowing, unlike in codeswitching, lexical items undergo phonological and morphological changes to adapt to the matrix language. For Myers-Scotton (1993a, 2002), borrowing and codeswitching are similar phenomena and both display an interaction between the matrix language and the embedded language. The only distinction that she notices is that borrowed words are usually highly frequent.

Obiamalu and Mbagwu (2007) argue that previous research and data show that constituents in codeswitching could also be subjected to phonological and morphological modification. They suggest that codeswitching is divided into three types: borrowing, quasi-borrowing, and true codeswitching. Borrowing happens when words from another language are inserted, after morphological and phonological adjustments, into the host language. In this type, the main reason for borrowing is the lack of the equivalents of the borrowed vocabulary. As for the quasi-borrowing, the equivalent of the lexical item is available in the host language and is used by monolingual and bilingual speakers. It is important to mention that linguistic assimilation of this type might or might not happen.

The last type is the true codeswitching where speakers also borrow terms even when a word already exists in their language. Only bilinguals adapt this linguistic behavior.

### **2.4.3. Motivations for codeswitching**

#### **2.4.3.1. Codeswitching**

In addition to studies on the structural aspects of codeswitching, researchers are interested in knowing and discussing the reasons that might lead any speaker to switch from one language to another (Myers-Scotton, 1993b; Gardner-Chloros, 2009; Alrowais, 2012; Ben Nafa, 2013; Hamouda, 2015; Kniaz, 2017). One reason could be the desire to reflect a specific social status. Auer (2002) believes that “codeswitching carries a hidden prestige which is made explicit by attitudes” (p. 221). Moreover, showing competence in another language is perceived as a sign of being well-educated, which in turn helps speakers to distinguish themselves from others and build their own social status (Suleiman, 1999). Aside from prestige, shifting to another language can be an expression of solidarity between people either from the same or different ethnic groups. Holmes and Wilson (2017) points out that “a speaker may. . . switch to another language as a signal of group membership and shared ethnicity within an addressee” (p. 35).

Some studies have shown that codeswitching has an impact on persuasion. Nerghes (2011) claims that codeswitching drives attention to the individual’s speech and makes it more persuasive to a target audience. She argues that, “codeswitching is an effective strategy that leads to systematic processing of information especially when associated with strong arguments” (Nerghes, 2011, p. 4). In addition to the previous reasons, it is believed that “people may switch code within a speech event to discuss a particular topic” (p. 38). This is shown in the results of Atawneh’s study (2007), where women switch to English when talking about topics relating to health, body, or food. In some societies, talking about certain topics in the native language might not be preferred because the speakers do not want to deal with the pressure of going against social norms. Leung (2006) states, “Taboo words and topics are often code-switched so as to avoid to be expressed in the speaker’s native language” (p. 14). An example on this point would be the South Korean rappers who switch to English in order to use taboo words that are found in the American hip-hop culture but not acceptable in their society (Lee, 2007).

### **2.4.3.2. Diglossic switching**

Although diglossia is considered a type of codeswitching by some researchers, it is preferred to separate the motivations of this linguistic behavior from that in classic codeswitching for the sake of clarity. It is not surprising that the two processes share many motivations as indicated by previous research. For instance, Abu-Melhim (1991) claims that Arabic speakers shift between varieties to gain prestige as noticed in his study where Jordanian speakers might shift from Jordanian Arabic to Egyptian Arabic because the latter is considered a more prestigious variety. Studies on political speech show that politicians use diglossic practices in their speech community in order to construct a rhetorical strategy so they can be more persuasive and be able to reach different classes in their audience (Holes, 1993; Al-Wer, 1997; Mazraani, 1997).

Albirini (2011) examines codeswitching between SA and dialectal Arabic in video and audio recordings in different domains such as religion, politics, and sport. He explains that diglossic switching is deployed to reinforce the idea of Arab nationalism in which “speakers seem to emphasize their pan-Arab or Muslim identity to invoke their relationship with other Arabs and Muslims” (Albirini, 2011, p. 546). Another reason for diglossic switching is to create a friendly and warm atmosphere between the speaker and the audience as assumed by Al-Batal (2002) in his study on the linguistic style of one of the Lebanese channels, Lebanese Broadcasting Corporation International (LBCI), where broadcasters tend to insert some colloquial markers while reading the news in SA. This is not the classic definition of diglossia, but this happens in diglossic societies where the effect of diglossic situation manifests itself in the speech of broadcasters.

## **2.4.4. Codeswitching and gender as a sociolinguistic variable**

### **2.4.4.1. Codeswitching**

The current literature on codeswitching between English and Arabic abounds with examples of studies conducted to search for any possible correlation between codeswitching and gender (Sadiqi, 1995; Atawneh, 2007; Sayahi, 2011; Al Batoush, 2014; Ismail, 2015; Al Kaddour & Kaddoura, 2019). Several Arabic linguists conclude that gender plays a pivotal role in Arabic speakers’ codeswitching performance across many different geographical places in the Arab world. In Morocco, Sadiqi (1995) finds that Moroccan women shift between French and Moroccan Arabic more than men do in

order to gain high status or prestige in their society. The same conclusion is found in two studies that are carried out on university students in Palestine and Jordan. The results of the distributed questionnaires show notable gender differences in the use of English loans in female students' speech, in which they tend to use English borrowed words more than their male colleagues (Atawneh, 2007; Al Batoush, 2014).

According to a study on university students in Saudi Arabia, gender also correlates with codeswitching in which "CS into English would seem to be an emblematic part of Saudi female speech style since women overall resorted to substantially more CS into English than male speakers" (Ismail, 2015, p. 107). In contrast to all the previous studies, Sayahi (2011) shows no difference either in patterns or frequency of codeswitching between males and females in Tunisia. Instead of gender, education is proven to be a significant factor, with Tunisians, who are high school graduates, code-switch less than those who have university degrees.

#### **2.4.4.2. Diglossic switching**

Many studies concluded that women use more standard pronunciation than men in Western communities (Chambers, 1995). Sociolinguists notice that the Western pattern of standard and non-standard use of the language is reversed in Middle Eastern communities. Abd-el-Jawad (1981) examined the linguistic variation in spoken Arabic in Amman, Jordan. He found that gender played a major role in the use of this variable's three variants: the standard variant [q], the non-standard variant [g], and the urban variant [ʔ]. The results indicated that women in Amman use the non-standard variant more than men. This finding is similar to the findings of previous studies on the same variable but on different populations such as variation in Cairo (Schmidt, 1974) and in a group of international people from Syria, Egypt, Jordan, Lebanon, and Palestine (Salam, 1980). The opposite pattern of standard use in Arabic to that in English leads to controversial beliefs and explanations among sociolinguists. Bakir (1986), for example, attributes this reversed pattern to the impact of gender-based factors in the Middle Eastern world where women do not have equal chances to be exposed to the standard form of the language as men do.

Abd-el-Jawad (1987) disagrees with this view and discards the interpretation of those results as a reversed image of language use in Western societies. He argues that the Western pattern cannot be completely copied and applied to other languages without

considering the different social and cultural backgrounds. He claims that, although the urban variant [ʔ] is a non-standard form, it is viewed in Amman as prestigious. Therefore, women in his study tended to use the non-standard but prestigious variants in their speech, as compared with men. Some researchers do not only reject applying language patterns to another culture, but they also believe that it is impossible to treat one of the opposite genders in a certain culture as one entity. For example, Sadiqi (2003a) argues that for someone to have a good understanding of women in Morocco, they should not look at women as one entity. Moroccan women and their roles in society differ according to geographical, economic, religious, social, and political factors.

In this chapter I provided justification for the objectives of this thesis by covering the relevant literature on style, diglossia in Arabic speaking countries and codeswitching. On the basis of this review, I have decided that the appropriate research questions for this study are:

- What are the overall patterns of language use?
- What are the structural characteristics of these patterns?
- What are the social characteristics, and in particular are there differences in occupation, gender, and age associated to particular patterns?

In the next chapter I will discuss the details of the methodology that I employed to answer these questions.



## **Chapter 3.**

### **Methodology**

#### **3.1. Methodology in the previous literature**

As any topic in sociolinguistics, codeswitching is studied by using qualitative designs, quantitative designs, or mixed designs of both qualitative and quantitative methods. According to Friginal and Hardy (2014), qualitative approaches that have been used in sociolinguistic studies involve “ethnographies (such as field studies and field interviews), case studies, phenomenology, narrative research, grounded theory, historical research, critical social research, and ethical inquiry” (p. 9). Quantitative research, on the other hand, is “the systematic data collection and analysis involving the application of mathematical and statistical techniques in further describing the nature of linguistic variation within a particular research setting” (Friginal & Hardy, 2014, p. 13).

Deciding on which design to use from the above designs is not a mere personal preference for scholars. In fact, it is largely governed by the research question which works as a design indicator (Mallinson, Childs, & Herk, 2013; Holmes & Hazen, 2014; Albirini, 2016). Mallinson et al. (2013) point out three cases, each of which strictly requires a particular approach to be followed. First, if the research question is mainly about a language phenomenon and how it is shown in the linguistic system, then a quantitative approach is needed to track this change. Second, when the research question is about language attitudes or identity, a qualitative approach is more applicable to capture and analyze the topic on a macro level. The last case is when researchers go beyond the above two questions to study “how the linguistic system constrains language change as well as how speakers create new meanings by using the potential of language to be variable.” (Mallinson et al., 2013, p. 11). In this situation, a mixed-method approach is needed to appropriately answer those questions and to provide deeper insights.

In addition to research design, choosing the right type of data is of a great importance in any study. Failing to do so can jeopardize the reliability of the study, which in turn can lead to misleading conclusions (Wolfson, 1986, p. 689). Albirini (2016) lists

four kinds of data that sociolinguists use in their research: naturally produced data, elicited speech, retrospective and introspective data, and written texts. Out of those four forms, naturally produced data are preferred because they exclusively resemble everyday language use (Labov, 1972). This category includes all speeches that are not produced to serve as research data, such as university lectures, political speech, religious sermons, news interviews and reports, and daily communication. The only disadvantage of this form of data is that it sometimes does not provide the researcher with enough tokens to examine specific linguistic behaviors. Since the internet has been considered a rich source of information, the internet mediated approach enables researchers to gather non-reactive data, i.e., naturally produced speech, by non-reactive methods, i.e., “the undisclosed observation and analysis of documents and archives” (Hewson, 2008, p. 5). In contrast with natural speech, elicited speech, which is characterized as reactive data, is speech obtained from speakers who are informed about being participants in a certain study and about the tasks that they are expected to perform. The drawback of elicited data is that it is less natural as compared to naturally produced data. Similar to elicited data, introspective and retrospective data are considered less natural as well. These data are collected through many reactive methods such as questionnaires, face-to-face or group-based interviews, open-ended questions, self-reflections, etc. Although both elicited data and introspective and retrospective data lack the quality of being 100% natural, they are embraced by most scholars as a way to obtain a sufficient number of tokens for any investigated phenomenon (Albirini, 2016).

Besides choosing suitable instruments for data collection, sociolinguists have many methods to choose from in order to analyze their collected data. Depending on the project's focus, the data can be analyzed phonologically, phonetically, morphosyntactically, or lexically. There are also other approaches that have been used to analyze data such as corpus linguistics, conversation analysis, critical discourse analysis, etc. (Holmes & Hazen, 2014). Out of those methods, corpus linguistics has been recently adopted as “a method of analysis that can be used for almost any form of linguistic inquiry. Researchers examine large amounts of naturally occurring language data, using computer software to help them to identify unexpected patterns or confirm their hypotheses.” (Holmes & Hazen, 2014, p. 107). Halim and Maros (2014) used this method of analysis by creating their own corpus of Facebook written interactions for five

Malay-English speakers in order to discover the possible functions of their codeswitching.

Another method of analysis in codeswitching research is conversation analysis, which is a set of methods that “investigate and identify the general social interactional practices and competences shared by members of a speech community, competences which enable them to interact meaningfully with one another, and which (largely) do not depend on their particular speaker identities.” (Holmes & Hazen, 2014, p. 231). One of the recent studies that used conversation analysis was conducted by Balamoti (2010) to know the reasons behind codeswitching among Greek students in Edinburgh, Scotland.

In addition to corpus linguistics and conversation analysis, discourse analysis has been also favored as a method of analysis that “provides one way of exploring the relationship between language and society at the micro level of social interaction.” (Holmes & Hazen, 2014, p. 178). Mukenge and Chimbarange (2012) used discourse analysis to examine the use of codeswitching in the film *Yellow Card* as a strategy to avoid topics and terms that are seen as taboos in public interactions. The researchers believe that this method of analysis “allows the researchers to analyze the words and phrases used by the characters in the chosen films within their specific contexts” which in turn helps understand codeswitching. (Mukenge & Chimbarange, 2012, p. 582).

To have a close look at different research designs in codeswitching research, three recent studies will be discussed in this section: Asali, 2011; Christoffersen, 2014; and Ismail, 2015. Ismail (2015) conducted a quantitative study to examine both the frequency and kind of codeswitching in Saudis’ mixed-sex interactions. She collected her data by recording the speech of six Saudi couples who were all deemed to be “communicative Arabic-English bilinguals” (p. 102).

Asali (2011) carried out a qualitative study in the United States of America to investigate the different attitudes of Arab Americans towards codeswitching between Arabic and English. In addition, the study aimed at knowing when and why both codeswitching to English and codeswitching to Arabic took place. The researcher gathered her data through three instruments: questionnaires, interviews, and personal observation. The participants were 200 Arabs from three states: New York, Illinois, and Michigan. The demographic characteristics of her sample involved age, occupation,

gender, level of education, level of both English and Arabic proficiency, years of residence, place of residence, and original nationality.

Christoffersen (2014) used a mixed-method approach to compare codeswitching between two different groups of kindergarten children: 2 L1 bilinguals, who learned two languages as their first languages before reaching three years old, and L2 learners, who learned an additional language at the age of three or older. The study participants were thirty students. Six of them were Spanish/English 2 L1 bilinguals and the rest were L2 learners of Spanish. The researcher used two methods to gather data: observation and recording 12 hours of the participants' interaction. The data then were analyzed using both quantitative methods, where frequencies of codeswitching were quantified, and qualitative methods, where instances of codeswitching were grammatically and conversationally investigated.

As mentioned above, Mallinson et al. (2013) provide clear instructions on how to choose a research design. Since the objective of this study is to understand the general pattern of code usage first, and then explore the role that the codes can play in the construction of a certain persona, a quantitative design is the most appropriate. In the following sections, I discuss the detail of the study's methodology.

### **3.2. Subjects**

This study's design is comprised of 18 speakers divided into two main groups. The politicians' group contained six subjects (3 males, 3 females), and the actors' group comprised 12 speakers (6 males, 6 females). The actors are also classified into two groups by age. Six are born before 1963 (old), and six after 1984 (young). The level of education varied between the politicians and the actors. All of the politicians hold postgraduate degrees such as master's degrees or Ph.D. degrees. Unlike the politicians, the actors' educational background varies between high school diplomas and bachelor's degrees. As for English proficiency, the politicians have good command of English that can be seen through their interviews in non-Arabic TV channels where English is the language of communication. Moreover, their careers in politics require them to be fluent in English since they meet with their counterparts of different countries and they act as official spokespersons for major announcements. In addition, some of the politicians obtained their degrees abroad in universities where English is the language of

instruction. Some of the actors, on the other hand, did not have a high level proficiency in English. One of the actors made it clear that he does not understand English very well by telling his story when he worked with a non-Arab producer and how the communication was difficult due to language barriers. One of the female actors had a quick interview with a British presenter and it was obvious that the actor struggled to understand and to deliver speech in English. Since this research is conducted to investigate the linguistic variation in Gulf Arabic, the sample includes subjects from different countries of the Arabian Gulf region: Saudi Arabia, Kuwait, United Arab Emirates, Qatar, Bahrain, and Oman (Table 1).

**Table 1. Distribution of subjects by country**

Country	Politicians	Actors	Total
Saudi Arabia	1	2	3
United Arab Emirates	1	2	3
Bahrain	1	2	3
Kuwait	1	4	5
Qatar	1	1	2
Oman	1	1	2
<b>Total</b>	6	12	18

### 3.3. Materials

The data was obtained by collecting a pre-recorded video interview for each subject. The main source of those interviews was YouTube, which is a very well known video-sharing website. It is important to draw the readers' attention to the fact that these interviews were recorded and previously broadcasted on TV channels. Only the first 30 minutes of every interview was analyzed. The total length of the interviews was nine hours. They were all recorded relatively recently, between the years 2010 and 2015. Regarding the content's topics, the politicians' interviews were about political issues and affairs that were important for the Arabian Gulf region states on both national and international levels. Standard Arabic is expected in such formal settings and this type of interviews is of a great interest to the politics-oriented audience. Actors, on the other hand, discussed various topics about their accomplishments, recent works, and

struggles in their careers. All of the actors' interviews focused on their acting career and their concerns and hopes in this field. This type of programs attracts audiences, from different age groups and educational backgrounds in the six Gulf States, who are interested in knowing the latest updates of their actors. Colloquial Arabic is the variety that is used by the audience of such programs.

I chose this Internet mediated approach of interviews, mainly for reasons for access to data, but also because the genre of TV interview is well-established and certainly the politicians and actors that are being interviewed are very familiar with it. Thus, even though this is not a "natural" conversation in the sense of Labov (1972), it is a stylized genre that allows us to observe how linguistic resources can be used in the construction of a certain persona.

### **3.4. Procedure**

The data underwent three stages of encoding: transcription, segmentation, and coding. Starting with transcription, Edwards (2008) states that recordings are not of great importance without having them transcribed. This is because transcripts "... provide a distillation of the fleeting events of an interaction, frozen in time, freed from extraneous detail, and expressed in categories of interest to the researcher" (Edward, 2008, p. 321). Therefore, the first 30 minutes of every interview was carefully transcribed into Arabic except sentences or words that were pronounced in English. The interviewers' speech was not transcribed. Only the speech of the 18 subjects was transcribed since their language is the focus of this study.

The second stage of encoding is segmentation. I paid attention to what I call *primary segmentation*, which is every natural break that took place when the interviewer either interrupted or asked questions to the interviewee. This type of segmentation generated chunks of speech that were separated into utterances, which are the basic units of analysis in this study. As mentioned in the previous sections, the analytical unit is problematic and an area of contention by scholars in the field. Some believe CP "projection of complementizer" is the most appropriate unit to rely on in speech segmentation (Myers-Scotton, 2002), while other linguists advocate for the sentence (Bentahila & Davies, 1998; Jacobson, 2001). In the current study, the utterance was chosen to be the main unit of analysis. According to Crookes and Rulon (1985), an

utterance is “a stream of speech with at least one of the following characteristics: under one intonation contour, bounded by pauses, and constituting a single semantic unit.” (p. 9). Regarding the length of pauses, Scollon (1974) stated that a pause can indicate an utterance boundary if it is between 1 to 8 seconds. This approach in segmentation is believed to capture both diglossia and codeswitching more than sentence or CP approaches. This is because language production is a mental process that involves three levels: conceptualization, formulation, and articulation (Levelt, 1999). Relying only on the syntactic structure and neglecting other important aspects such as intonation and pauses in speech segmentation might not lead to accurate analysis, especially when working on data that has two linguistic behaviors: codeswitching and diglossia. What makes the utterance approach distinct from other approaches is the fact that intonation plays a major role in dividing the data, which in turn helps in indicating the “immediate constituents structure” of the language (Lea, 1973).

### **3.4.1. Coding**

After the data were segmented into utterances, every utterance was put under a certain category. Seven categories were postulated in order to track both diglossia and codeswitching in the subjects’ speech. Those categories are listed as follows:

- 1) **CA** (Colloquial Arabic utterances)
- 2) **SA** (Standard Arabic utterances)
- 3) **EN** (English utterances)
- 4) **SA/CA** (Standard and colloquial Arabic mixed utterances)
- 5) **CA/EN** (Colloquial Arabic and English mixed utterances)
- 6) **SA/EN** (Standard Arabic and English mixed utterances)
- 7) **SA/CA/EN** (Standard Arabic, colloquial Arabic, and English mixed utterances)

The first category was “CA” which was the label of colloquial Arabic utterances. Classifying utterances as colloquial was based on the studies of Qafisheh (1977), Holes

(1990), and Adouane and Johansson (2016) on Gulf Arabic. The following utterance is an example of “CA”:

(1)

Lah	ihni	fiḥ	iḥtmam	o	ihni
Neg	here	there	attention	and	there
ma	fiḥ	iḥtmam			
Neg	there	attention			

‘No, here there is an attention and here there is not.’

In Example (1), the whole utterance is colloquial Arabic on almost all linguistic levels: syntactic, phonological, and lexical. The SA counterpart would be a sentence like the following:

(2)

la	huna	juḍʒad	iḥtimam	huna	wa
Neg	here	there	attention	there	and
juḍʒad	iḥtimam				
there	attention				

‘No, here there is an attention, and here there is no attention.’

We can see that the colloquial /ihni/ was used in example (1) for the standard /huna/ in (2). Also, the colloquial /fiḥ/ in (1) was used for the standard /juḍʒad/ in (2). The negation also is typical of colloquial speech by using /ma/ in this context. There is also the word /iḥtimam/, which was pronounced colloquially as /iḥtmam/ where the vowel /i/ after /t/ was unpronounced.

The second category was “SA” which included utterances in only SA as illustrated in example (3).

(3)

al-etifaqiah	tuṣṭabar	tamdi:d	aw	tafʿi:l	li-l- etifaqiah
Def-agreement	consider-Pass	extension	or	activation	for-Def-agreement
a-sabiqah					
Def-last					

‘The agreement is considered an extension or an activation for the last agreement.’



The third category was “EN” which included any monolingual English utterance. Moving to the fourth category, any utterance that was a mixture of standard and colloquial Arabic was given the label “SA/CA”. Colloquial Arabic words were underlined as shown in utterance (4) where /ħina/ was used instead of the standard form /naħnu/. There is also the colloquial form /ħimrna/ for the standard /qatʕ/. The double underlined words are actually SA but pronounced colloquially. For instance, the word /libas-na/, which means *we wore*, was pronounced colloquially where the vowel of the short open syllable was raised. The SA’s pronunciation of this word is /labis-na/.

(4)

itʕlaqan absolutely	<u>ħina</u> we	<u>ħimrna</u> have	ma never	<u>libas-na</u> wear-1Pl	ħađihi this
a-nađʕarah Def-glasses	a-tʕaʕifeih Def-sectarian				

‘We absolutely have never worn those sectarian glasses.’

It was noticed that some CA utterances had English borrowings. Those utterances were put under the fifth category “CA/EN” as in example (5) where the English word *fame* was Arabized by prefixing it with the definite article /al-/.

(5)

ħiʕub-ha after-that	inxatʕab-na got engaged-1Pl	wa and	ħiʕub ma after	inxatʕab-na got engaged-1Pl	ħam Continue-Past
waqif stand up	maħia with-1Sig	wa and	lakin but	al- fame Def- <b>fame</b>	ħei thing
doodah-ih distracted-him	wa and	ma Neg	alum-ih blame-him	amantan honestly	

‘After that we got engaged and after getting engaged, he stood up for me but fame is something that distracted him and I honestly do not blame him.’

The use of English borrowings was also spotted in SA utterances. Those SA utterances mixed with English were put under the sixth category “SA/EN”. In example (6), the English word *agenda* was modified or Arabized by applying the Arabic rules of grammar and morphology. For instance, the English word *agenda* was treated as a feminine word by adding /-h/ to it. *Agenda* was also made definite by prefixing it with the Arabic definite article /al-/ to become /al-agenda-h/.

(6)

fʻabʻʻan of course	fi in	<b>al-agenda-h</b> Def-agenda-fem	al-awaliah Def-priority	hia is	moħarabat fighting
al-faqir Def-poverty	liʻna because	al-faqir Def-poverty	biħad it	ħatih self	huwa is
ʻmil factor	asasi main	fi in	Kaθi:r a lot	min of	a-douwal Def-countires
l-maʻakil for-problem	oxra other	miθil like	a-sʻaħah Def-health	wa and	a-taʻli:m Def-education
wa and	a-taʻmi:n Def-security	al-ʻiħaʻi: Def-food-adj			

'In the agenda, the priority, of course, is fighting poverty because poverty itself is the main factor for other problems such as health, education, and food security in a lot of countries.'

The seventh category was the one combining SA, CA, and EN "SA/CA/EN". In (7), the borrowed word is bolded and the colloquial ones are underlined. The Standard forms for the undelined colloquial words /illi/, /maħtʻutʻah/, /ras-na/ are /allati/, /wudʻiʻʻat/, and /ruʻusina/ respectively. As for the borrowed word in this example *imposing*, it preserved the original English pronunciation /ɪmpəʊzɪŋ/. However, it was prefixed with the definite article /l-/.

(7)

faliesat Neg	a-demoqratiħah Def-democracy	ma what	nusami:ha 1PI call	<b>bi-l-imposing</b> of-Def-imposing	yaʻni means
<u>illi</u> that	<u>maħtʻutʻah</u> put-Pass	ʻala on	<u>ras-na</u> head-Poss-1PI	jadʻzib have	an to
numaθilha 1PI-represent					

'Democracy is not what we call imposing, which we are forced to represent.'

It is typical for studies such so the present study to have a second rater to test the coherence of the coding scheme. This was not possible because of the inability to find someone who had the appropriate knowledge of Gulf Arabic and skills in linguistics to qualify for a second rater. What we did instead was rating 50% of the data in many sessions, where Dr. Pappas and I discussed and went over each of the utterances. Based on the results of the two raters who coded this subset of data, the criteria for each category was confirmed; thus, it was concluded that adequate reliability and validity had been achieved in the process of the coding.

To conclude this chapter, a quantitative design is being implemented to investigate the nature of code mixing among the subjects of this study who use language as a tool to project certain types of personae. The next chapter will present the results of the collected data.

## Chapter 4.

### Results

This chapter focuses on the description of the findings of this study. There are three main sections in this chapter. The first section 4.1 starts with a general description of the corpus. Section 4.2 offers a structural description of the morphosyntactic aspects of two types of codeswitching: English/Arabic 4.2.1, standard Arabic/colloquial Arabic mixing 4.2.2. The third section 4.3 presents the results that show the effect of occupation, gender, and age on the style of this study's subjects. Section 4.4 is a summary of the results.

#### 4.1. General description of the corpus

Codeswitching can be analyzed on both sociolinguistic and structural levels. In the following sections, this phenomenon will be investigated from a structural perspective where the phonological, morphological, and syntactic aspects of the data are going to be the focus of the analysis.

**Table 2. Types of utterances in the corpus**

Type of utterances	Number	Percentage
Colloquial Arabic (CA)	1669	59
Standard Arabic (SA)	314	11
English (EN)	15	1
CA/SA	477	17
CA/EN	273	10
SA/EN	11	0
CA/SA/EN	49	2
Total	2808	100

There are 2808 utterances in total in this corpus. Table 2 shows that the majority of the subjects' speech is in monolingual CA (59%). There are also 314 monolingual utterances in standard Arabic (SA, 11%), and only 15 utterances in English (EN, 1%).

The most common type of mixing is that between CA and SA with 477 utterances (17%). Next we find CA and EN mixed utterances (10%), 49 utterances (2%) that have all three codes, and just 11 utterances with SA and EN.

An example of the monolingual CA is given in the utterance in (1):

(1)

ʃahad	wa:ɪd	itwa:si:ni:	fi	h-a-salfah	haði
Def-agreement	a lot	comfort-1Sig	in	Dem-Def-matter	this

'Shahad comforts me a lot regarding this matter.'

Example (2) shows a pure SA utterance:

(2)

al-etifaqiah	tuʃtabar	tamdi:d	aw	tafʕi:l	li-l- etifaqiah
Def-agreement	consider-Pass	extension	or	activation	for-Def-agreement
a-sabiqah					
Def-last					

'The agreement is considered an extension or an activation for the last agreement.'

As for the English utterances, they are mainly short responses that can be words, phrases, or short sentences, such as in (3):

(3)

'Yes, exactly.'

As mentioned above, mixed utterances of CA and SA is the second common type of utterances in the corpus. Example (4) shows a mixed utterance of CA and SA where the bolded words are CA words. The underlined word /libas-na/ is actually SA but pronounced colloquially. This verb means *we wore* and it was pronounced colloquially, that is, the vowel of the short open syllable was raised. The SA pronunciation of this word is /labis-na/. Classifying utterances as colloquial was based on the studies of Qafisheh (1977), Holes (1990), and Adouane and Johansson (2016) on Gulf Arabic. The rest of the bolded words are colloquial lexical items. The first plural pronoun /hina/ is the colloquial form for the SA equivalent /naħnu/. The colloquial word /ʕimrna/ is used here as a negative intensifier with /ma libas-na/ to create the double negative construction /ʕimrna ma libas-na/. The double underlined words are actually SA but pronounced

colloquially. For instance, the word /libas-na/, which means we wore, was pronounced colloquially where the vowel of the short open syllable was raised. The SA's pronunciation of this word is /labis-na/.

(4)

it'laqan absolutely	<b>hina</b> we	<b>Simrna</b> have	ma never	<b>libas-na</b> wear-1PI	haðihi this
a-nað'arah Def-glasses	a-t'aʔifeih Def-sectarian				

'We absolutely have never worn those sectarian glasses.'

The utterance in example (5) is an illustration on the bilingual CA and EN utterances in the corpus. We can see that the utterance below contains one English noun:

(5)

ʕiqub-ha after-that	inxat'ab-na got engaged-1PI	wa and	ʕiqub ma after	inxat'ab-na got engaged-1PI	tam Continue-Past
waqif stand up	maʕia with-1Sig	wa and	lakin but	al- fame Def- <b>fame</b>	ʕei thing
doodah-ih distracted-him	wa and	ma Neg	alum-ih blame-him	amantan honestly	

'After that we got engaged and after getting engaged, he stood up for me but fame is something that distracted him and I honestly do not blame him.'

The corpus has some monolingual SA utterances with EN switches. In example (6), the English noun *agenda* is inserted in a SA utterance:

(6)

tʿabʿan of course	fi in	al- <b>agenda</b> -h Def-agenda-fem	al-awalah Def-priority	hia is	moʿharabat fighting
al-faqir Def-poverty	liʿna because	al-faqir Def-poverty	biḥad it	ḍatih self	huwa is
ʿmil factor	asasi main	fi in	Kaḥi:r a lot	min of	a-douwal Def-countires
l-maʿakil for-problem	oxra other	miḥil like	a-sʿaḥah Def-health	wa and	a-taʿli:m Def-education
wa and	a-taʿmi:n Def-security	al-ʿiḍaʿi: Def-food-adj			

'In the agenda, the priority, of course, is fighting poverty because poverty itself is the main factor for other problems such as health, education, and food security in a lot of countries.'

The last example (7) is an utterance that encompasses all three codes: SA, CA, and EN. There is an English word in this utterance *imposing*. This English insertion preserves its original pronunciation and prefix with the Arabic definite article –/l. The underlined clause is CA. The SA equivalent for /illi maḥtʿutʿah ʿala ras-na/ is /allati wudʿiʿat ʿala ruʿusuna/. To make it clearer, the SA form for the CA demonstrative pronoun /illi/ is /alliti/. The SA form for the colloquial passive verb /maḥtʿutʿah/ is /wudʿiʿat/. The object of the preposition /ras-na/ is used in the colloquial form where the speaker referred to herself as well as her people using this singular colloquially pronounced form instead of the SA plural form /ruʿusina/.

(7)

faliesat Neg	a-demoqratiʿiah Def-democracy	ma what	nusami:ha 1PI call	bi-l- <b>imposing</b> of-Def-imposing	yaʿni means
<b>illi</b> that	<b>maḥtʿutʿah</b> put-Pass	<b>ʿala</b> on	<b>ras-na</b> head-Poss-1PI	jadʿib have	an to
numaḥilha 1PI-represent					

'Democracy is not what we call imposing, which we are forced to represent.'

## 4.2. Structural aspects of codeswitching

This section will be devoted to describing the structural aspects of the two main types of codeswitching that are found in the data: Arabic/English (333 tokens, %12) and

SA/CA (477 tokens, 17%). It mainly focuses on the morphosyntactic characteristics of both types of codeswitching. The framework of the present study is based on the Matrix Language-Frame model (MLF) introduced by Myers-Scotton (1993a; 2002).

#### 4.2.1. Arabic/English codeswitching

As it is shown in Table 1, English is found in CA utterances, SA utterances, or mixed utterances of SA/CA. The data reveals that English/Arabic utterances occur on the lexical, the phrasal, and the sentential levels. This section focuses on the general description of Arabic/English codeswitching. Table 3 shows the number of lexical and phrasal bilingual utterances in the entire corpus.

**Table 3. Lexical, multi-lexical, and sentential English elements**

EN elements	Number	Percentage
Lexical instances	337	80
Multi-lexical instances	43	10
Sentential instances	41	10
Total	419	100

As shown in Table 2, 12% of 2808 utterances have at least either a lexical, a multi-lexical or sentential English element. The total number of English elements 419 exceeds the total number of utterances with English elements 333 due to the fact that some utterances contain more than one English element. Table 3 shows that English lexical elements form 80% with 337 insertions. English multi-lexical and sentential elements constitute 20% of a total of 419 English elements. The present study will only analyze lexical and multi-lexical insertions.

##### 4.2.1.1. General description of English lexical insertions

In this study, most English lexical elements found in the Arabic utterances are nouns, adjectives, or discourse markers. Table 4 provides an overview of the number of English single words embedded in Arabic frames.



**Table 4. English single words in Arabic frames**

Type	Number	%
Nouns	258	77
Adjectives	43	13
Discourse markers	19	6
Adverbs	8	2
Verbs	5	1
Response particles	4	1
Total	337	100

This table indicates that nouns are the most favored lexical elements among the subjects. 77% of a total of 337 English lexical elements are nouns. Adjectives come after nouns in terms of frequency where they form 13% of lexical English borrowings. Then, discourse markers constitute 6% of the lexical elements in the corpus. There were eight adverbs (2%), five verbs (1%), and four response particles (1%). In general, the analysis of the data revealed that different types of English lexical elements could be embedded in Arabic morphosyntactic frames. Example (8) is a bilingual CA/EN utterance that contains one English noun:

(8)

i:h	naggasat	min	al-points	mali
yes	deduct-Past	from	Def-points	mine

'Yes, it deducted my points.'

Adjectives are the second frequent English lexical items found in the data. An example of these adjectives would be *close* in the following utterance:

(9)

Ahmad	raħ	zʿaħajat	i-ǰiħrah	yaħni	iħna
Ahmad	was	victim	Def-fame	mean	we
rabiħ	wa:id	ħail	<b>close</b>	o	Ahmad
friends	much	very	close	and	Ahmad
taqadam	li	li?nih	yaħrif	muħanat	ħujoon
propose-Past	to me	because-1Sig	knows	suffering	Shujoon

ħino  
what

'Ahmad was a victim of fame, I mean, we are very much close friends and he proposed to me because he knows what Shujoon's suffering is.'

English adverbs occur only 8 times in the data. Example (10) shows how the English adverb *automatically* is inserted in the utterance:

(10)

la	hu	a-tasʿni:f	yadzi	<b>automatically</b>	yaħni
no	it	Def-classification	come	automatically	mean
a- sʿahafah	hi	illi	tigul	a-nas	huma
Def-press	it	that	say	Def-people	they
illi	yagulu	il-fatrah	a-zamaneih	illi	int
who	say	Def-period	Def-time	that	you

fi-ha  
in-it

'No, the classification comes automatically, I mean, the press tells. People are those who tell the period of time that you are in.'

Verbs seem to be the least inserted English item as compared to nouns and adjectives. The verb *block* is found in the data, as illustrated in example (11):

(11)

mu Neg	inh he	rad reply	rad response	mu Neg	kujeis good
ini that	ana I	aruḥ go	asi: disrespect	lih to-him	aw or
aṣmil make	lih to-him	<b>block</b> block			

'Even if he replies with a bad comment that does not mean that I will disrespect or block him.'

Most of the discourse markers in the data are *ok*. Example (12) contains an English discourse marker *ok*:

(12)

il-aṣmal Def-works	illi that	tuqadam given	li to me	ynasbni suit-Pres-me	agdar can-1Sig	aḡtiyil work-1Sig
fi-ha in-it	agdar can-1Sig	inih that		maḡalan for instance	yugu:lun say-1Pl	ibdaṣ creativity
mumaḡil actor	<b>ok</b> ok					

'The roles that are given to me, do they suit me? Can I perform them? Can they, for instance, say a creative actor? Ok.'

#### 4.2.1.2. **General description of English multi-lexical insertions**

The English insertions found in the corpus are not restricted to the single lexical insertions but include multi-lexical insertions. Table 5 indicates that the most frequent type of multi-lexical insertions is phrases.

**Table 5. EN multi-lexical elements in Arabic frames**

Type	Number	Percentage
Determiner phrases	27	63
Adjective phrases	8	19
Adverbial phrases	4	9
Verb phrases	2	5
Prepositional phrases	1	2
Conjoined phrases	1	2
Total	43	100

It can be noticed that the determiner phrases are the most used switches with 27 instances and constitute 63% of the EN multi-lexical elements. Example (13) contains an EN determiner phrase:

(13)

ana	kint	adris	<b>human resources management</b>	o	txaradzat
I	was	study	human resources management	and	graduated
b-tafwiq	baʕd				
with-excellence	too				

'I studied human resources management and I graduated with excellence too.'

Adjective phrases are the second-most occurring elements in this category and form 19% of the total 43 tokens. The utterance in example (14) contains the EN adjective phrase *sharp features*:

(14)

hw	<b>al-sharp features</b>	yaʕni	al-malamiḥ	al-ḥadah	daʕiman
	Def-sharp features	mean	Def-features	Def-sharp	always
btidi	al-ʕar				
represent	Def-evil				

'The sharp features, it means the sharp features, which always represent evil.'

After adjective phrases, adverbial phrases occurred only four times. Example (15) shows how the adverbial phrase *too much* is inserted in the utterance:

(15)

la no	wijlu:n how	mdʒahazah? prepare-1Sig?	ʕla by	fikrah the way	is-suʔal Def-question
haða this	awal first	marah time	ini I	ansʕidim Get shocked	min-h of-it
yaʕni mean	<b>too much</b> too much	ma Neg	adri know-1Sig	ħasiet feel-Past-1Sig	nafsi myself
ini that	ʕindi Have-1Sig	tʕajrah jet	xasʕah private		

'No, how am I prepared? By the way, this is the first time that I get shocked by this question. I mean it is too much. I don't know I felt that I have a private jet.'

There are two occurrences of EN verb phrases, which form 5%. The verb phrase *worked hard* in example (16) is used by one of the subjects of the study:

(16)

a:na I	abadan never	mub Neg	gilt lack	θiqah confidence	fi in
nafsi myself	wa and	la Neg	ʕei thing	bas but	inh that
a:na I	ħaseit feel-Past-1Sig	inh that	a:na I	<b>worked hard</b> worked hard	yaʕni mean
wijd huge	biðalt made	madʒhud effort	inih that	a:na I	atʕalam learn
kil every	ʕei thing				

'I never...it's not a lack of self-confidence but I felt that I worked hard. I made a tremendous effort so that I learn everything.'

The multi-lexical insertions include one occurrence of prepositional phrase and a conjoined phrase. Example (17) contains the prepositional phrase *in everything*:

(17)

ana                    ʔuʔmin                    bi-h                    aḥyanan                    w                    mub  
I                    believe                    in-it                    sometimes                    and                    Neg

**in everything**  
in everything

'I believe in it sometimes and not in everything.'

Example (18) contains the conjoined phrase *flu and everything*. It has the words *flu* and *everything* that are combined by the conjunction *and*:

(18)

i:h                    zkam                    **flu**                    **and**                    **everything**  
yes                    cold                    flu                    and                    everything

'Yes, cold, flu, and everything.'

#### 4.2.2. Colloquial Arabic/standard Arabic codeswitching

The nature of codeswitching between CA and SA in the corpus is of great interest since there are two distinctive types of mixed utterances according to the dominant variety: SA-dominant utterances, and CA-dominant utterances. SA-dominant utterances are utterances where SA mainly constitutes their structure and forms 76% of a total of 477 mixed utterances as shown in Table 6. CA-dominant utterances, on the other hand, are CA utterances that have SA insertions and form only 24% of the total utterances of CA/SA.

**Table 6. Utterances according to the dominant language**

Type	Number	Percentage
SA dominant utterances	361	76
CA dominant utterances	116	24
Total SA/CA utterances	477	100

##### 4.2.2.1. SA-dominant utterances

As shown in Table 6, SA-dominant utterances form 76% of a total 447 CA/SA utterances. Table 7 indicates that those utterances have at least one CA phonological,

lexical, multi-lexical or sentential element. Some utterances have even more than one CA element, which explains why the number of CA elements 658 is almost double the total utterances with CA elements 361. The majority of the embedded CA elements are phonological and lexical which form 58% and 33% respectively of the total tokens. The mixed utterances have 42 instances (6%) of CA multi-lexical elements. CA sentential elements seem to be the least embedded elements in SA-dominant utterances and form only 3% of a total of 658 inserted elements.

**Table 7. Distribution of CA elements in SA-dominant utterances**

CA elements	Number	Percentage
Phonological instances	381	58
Lexical instances	215	33
Multi-lexical instances	42	6
Sentential instances	20	3
Total	658	100

As indicated in Table 7, the CA phonological elements constitute almost more than a half of the total 648. The analysis reveals that 49% of those phonological elements is the change of the emphatic voiced stop /d<sup>ʕ</sup>/ to the interdental fricative voiced /ð<sup>ʕ</sup>/. This phonological process is known as consonant lenition (Odden, 2005). The alternation of this particular sound is noticed among Arabic speakers in Iraq and the Arabian Peninsula (Al-Sharif, 2017). This is confirmed in this study's corpus. We can see in example (19) how this sound was pronounced colloquially as /ð<sup>ʕ</sup>/ instead of the standard /d<sup>ʕ</sup>/:

(19)

kla	a-t <sup>ʕ</sup> arafein	yu:hawi:lan	an	yastayilan	al-wað <sup>ʕ</sup>
both	Def-sides	try	to	take advantage of	Def-situation
min	adʒil	taħqi:q	makasib	lahum	
in	order to	achieve	gains	to-them	

'Both sides are trying to take advantage of the situation in order to achieve their gains.'

#### 4.2.2.1.1. General description of CA lexical insertions

It is interesting to notice that the types of inserted CA lexical elements in SA frames are different from the types of the embedded English lexical insertions in Arabic frames (Table 8 & Table 4).

**Table 8. CA single words in SA-dominant utterances**

Type	Number	Percentage
Person pronouns	77	36
Relative pronouns	44	21
Demonstrative pronouns	41	19
Verbs	26	12
Adverbs	9	4
Nouns	6	3
Response particles	4	2
Interrogatives	3	1
Conjunctions	3	1
Negative particles	2	1
Total	215	100

The data analysis showed that different types of CA elements could be embedded in SA morphosyntactic frames. Table 8 shows that the majority of inserted lexical elements are pronouns, which form 76% of a total of 215 CA lexical elements found in the corpus. The most common type of pronouns is the person pronouns with 77 tokens (36%). In example (20), we can see a mixed utterance that contains one first person pronoun /iħna/ instead of the SA form /naħnu/:

(20)

fi	al-ħaqiqah	<b>iħna</b>	yadzib	an	naŕu:d
in	Def-fact	we	have	to	return
qalilan	ila	al-wara?			
a little	to	Def-back			

'In fact, we have to go back a little bit.'



After person pronouns, relative pronouns and demonstrative pronouns are the most frequent CA insertions; they constitute 21% and 19% respectively of the total number of CA insertions. Example (21) shows one CA relative pronoun /illi/ instead of the SA relative pronoun /allati/. There is also the dummy verb /fi:/ that is used to mean 'there are' instead of the SA form /hunak/:

(21)

<b>fi:</b>	ʕidat	ʕanasʕir	illi	ʔdat	ila
there	many	factors	that	lead-Past	to
<b>ðalik</b>					
that					

'There are many factors that led to that.'

The CA forms of demonstrative pronouns are frequently used in the SA dominant utterances. For instance, the CA demonstrative pronoun /haði/ is used instead of the SA /haðihi/ as shown in example (20).

(22)

<b>haði</b>	alʔumu:r	allati	nuhawil	an	nuðakir-hum
those	Def-matters	that	try-1PI	to	remind-3PI
<b>biha</b>	li-muraʕatiha				
of-3Sig	to take into consideration				

'Those matters that we try to remind them of to take into consideration.'

After pronouns, verbs frequently occur in SA dominant utterances and form 12% of the total number of CA lexical borrowings. We can see the CA verb /jifna/ instead of the SA /raʔeina/ in the following example:

(23)

<b>wa</b>	haða	ma	tam	fiʕlan	ʕala
and	this	Neg	was	indeed	on
<b>ardʕ</b>	al-waqiʕ	wa	<b>jifna</b>	nataʔdʒ	al-ħiwar
ground	Def-reality	and	see-Past-1PI	outcomes	Def-dialogue

'And this is what was done indeed on the ground.'

As shown in Table 8, adverbs constitute only 4% of the total CA lexical insertions in SA speech. Example (24) contains a CA adverb /xisʕah/ instead of the SA form /xusʕu sʕan/ :

(24)

<b>xisʕah</b>	fi	ʕahd	al-ʔnfitah	illi	taʕijʕuhu
especially	in	era	Def-openness	that	live-3Sig

al-Bahrain  
Def-Bahrain

'Especially in the openness era in which Bahrain lives.'

The results revealed that nouns were rarely replaced with the CA forms. There are only six CA nouns and they constitute 3% of the lexical insertions. Example (25) contains a CA noun /hala/ instead of the the SA form /ʔhlan/ :

(25)

<b>hala</b>	wa	ʕukran	ʕla	al-ʔstidʕafah
hello	and	thank	on	Def- hospitality

'Hello and thank you for the hospitality.'

Table 8 indicates that there are four response particles, which form 2% of the CA lexical insertions. For instance, the CA particle /i:h/ replaces the SA form /adʒal/ in example (26):

(26)

<b>i:h</b>	yu:dʒad	itisʕal	beina-na	wa	bein
yes	there is	contact	among-us	and	with

Ahrar Al-Sham  
Ahrar Al-Sham

'Yes, we are in contact with Ahrar Al-Sham.'

CA interrogatives are found in the data and constitute only 1% of the lexical insertions. The CA interrogative pronoun /leij/ is used instead of /limaða/ as shown in example (27):

(27)

leif why	ma Neg	nakun be-1PI	wad <sup>ʕ</sup> hi:n clear	yatalaqu:n receive-3PI	daʕm supprot
min from	Iran Iran	wa and	yatalaqu:n receive-3PI	daʕm support	min from
ʕhizb Hizb	Allah Allah				

'Why aren't we clear? They receive support from Iran and Hizb Allah.'

Besides interrogatives, Table 8 indicates that conjunctions are among the less frequent lexical insertions with only three occurrences (1%). Example (28) has the CA conjunction /ʕafan/, which means 'because', instead of /likei/:

(28)

ʕafan so	la Neg	a <sup>ʕ</sup> had one	yafham understand	a-ðaka? Def- intelligence	al-ʕat <sup>ʕ</sup> ifi Def-emotional
ana that	al-marʔah Def-woman	ʕat <sup>ʕ</sup> ifiah Emotional	wa and	a-ʕei? Def-thing	mawdʒu:d exist
fi in	al-dʒinsein Def-genders	wa and	leisa Neg	fi in	a-nisa? Def-women
faqat <sup>ʕ</sup> only					

'So no one understands that emotional intelligence means a woman is emotional; and this thing exists in both genders and not in women only.'

The least frequent CA lexical insertions in SA dominant utterances are the negative particles. They occur twice in the corpus and only form 1% of the total 215 CA insertions. Example (29) contains the CA negative particle /muf/ instead of the SA /leisa/:

(29)

ma what	yaquluhu say-2Sig	li to me	zami:li counterpart	wara? behind	al-ʔbwab Def-doos
al-muʔlaqah Def-closed	<b>muf</b> Neg	mus <sup>ʕ</sup> ara <sup>ʕ</sup> h declared	amam In front of	al-kamera Def-camera	

'What my counterpart says to me behind closed doors is not declared in front of the camera.'

#### 4.2.2.1.2. General description of CA multi-lexical insertions

Table 9 shows the types of multi-lexical insertions in the SA-dominant utterances. The frequently used multi-lexical insertions are classified as cross constituents, which are structures that consist of two main parts. The first item in such structure is colloquial and the second item is standard but influenced either phonologically, as in example (30), or morphologically, as in example (31), by CA. Cross constituents form 71% of the total of 42 CA multi-lexical insertions. The second most frequent type is tense phrases that occur four times in the corpus and form 10%. There are three complementizer phrases found among the CA switches, two determiner phrases, two prepositional phrases, and only one occurrence of CA verb phrase.

**Table 9. CA multi-lexical elements in SA-dominant utterances**

Type	Number	Percentage
Cross constituents	30	71
Tense phrases	4	10
Complementizer phrases	3	7
Determiner phrases	2	5
Prepositional phrases	2	5
Verb phrases	1	2
Total	42	100

Example (30) contains a cross constituent /illi tidʒimih/, where the first item is the CA relative pronoun /illi/ for the standard counterpart /allati/, and the second lexical item is standard but pronounced colloquially. It is clear that the first CA element phonologically influences the second element in this example, which is a way of transition from CA to SA that is noticed among the speakers in this data.

(30)

la Neg	nastatʕi:ʕ can-1PI	an to	namʕi: walk-we	wa and	naqu:l say-we
faqatʕ only	leisat Neg	hunak there	Qaʕidah Qaeda	dʕimn within	haða this
a-zaxam Def-momentum	al-iʕlami Def-media	a-rahi:b Def-horrible	illi that	tidʕimih Support-it	al-wilayat Def-United
al-mutaʕidah Def-States	fi in	iθbat proving	anna that	hunak there	Qaʕidah Qaeda
wa and	irhab terrorism				

'We can't simply walk and say there is no Qaeda within this horrible media momentum that the US supports in proving that there are Al-Qaeda and terrorism.'

Example (31) shows another incidence of cross constituents /illi inwaldat/, where the first item is the colloquial relative pronoun /illi/ followed by the verb /inwaldat/. This verb is interesting since the colloquial form of the passive verb applied to it by attaching the prefix *-in* to become /inwaldat/ instead of the standard form /wolidat/. The second item of the constituent in this example is morphologically influenced by the first CA element.

(31)

a:na I	ma Neg	aʕuf see	yeir except	anna that	aħzab parties
al-islam Def-Islam	a-siyasi Def-political	a-shi:ʕi Def-Shiit	illi that	<b>inwaldat</b> born-Pass	wa and
illi that	ʕaʕat live-Past	fi in	Iran Iran	hya it	aallati that
dʕaʕat come-Past	taħkum govern	baʕda after	al-ʔiħtilal Def-occupation		

'I don't see except the Shiite political parties, whose members were born and lived in Iran, are the ones that came to govern after the occupation.'

Moving to the second most frequent CA multi-lexical type, we can see that tense phrases occur four times. Example (32) contains a joined tense phrases /astanna wa axalli/ where both of the verbs are colloquial forms for the standard /antazʕir wa adaʕ/.

It is important to mention that the two verbs have implicit subjects; therefore, they are classified as tense phrases.

(32)

ila until	mata when	<b>astanna</b> wait-1Sig	<b>wa</b> and	<b>axalli</b> let-1Sig	Iran Iran
taqu:m do	bi-l-ʔʕmal of-Def-acts	haði? those			

'For how long should I wait and let Iran do those actions?'

CA complementizer phrases occur only three times in the corpus. An example on these phrases will be the complementizer phrase /ʃagul lik/, in example (33), which is another instance of cross constituents where the first item is a colloquial form of a question where the prefix *-ʃ* is attached to the verb /agul/, which is pronounced colloquially here. The standard form for the first item is /maða aqulu/ where we have the SA question word /maða/ followed by the verb in SA /aqulu/. The second item is pronounced colloquially as /lik/ and the standard pronunciation will be [laka].

(33)

yaʕni mean	<b>ʃagul</b> say-1Sig	<b>lik</b> to-you	fi in	kul all	al-buldan Def-countries
al-mutaħaðʕrah Def-civilized	muʕðʕam most	al-ʔanʃiʕah Def-activities	a-ħaqaʕyah Def-cultural	madʕu:mah supported	min by
al-qitʕaʕ Def-sector	al-xasʕ Def-private	wa and	la Neg	buda must	an to
nastafi:d draw on	min from	tadʕarib experiences	al-ʔxri:n Def-others		

'I mean what can I say to you? In all of the civilized countries, most of the cultural activities are supported by the private sector. We must draw on the experiences of others.'

After complementizer phrases, we can see that prepositional phrases are among the least frequent CA multi-lexical insertions. The colloquial prepositional phrase /fi:ha/, which means *of it*, is used instead of the standard counterpart /biha/ in example (34):

(34)

wa and	liðalik thus	hunak there	ihtyatʿat precautions	kaθi:rah several	alʔan now
taqum take	<b>fi:-ha</b> of-it				

'And thus there are several precautions Kuwait takes now.'

It is interesting to note that determiner phrases are among the least frequent CA multi-lexical insertions in SA dominant utterances, whereas they are the most frequent EN multi-lexical insertions in Arabic utterances (Table 5 & Table 9). Example (35) contains the determiner phrase /a-ssafrat illi kinna mawdʒu:di:n fiha/. This phrase has the colloquial plural form /a-ssafrat/ for the standard form /al-asfar/, and the colloquial relative pronoun /illi/ instead of the standard /allati/. Then comes the colloquial construction /mawdʒu:di:n fiha/, which cannot be used in this particular context. The standard form is /kunna mʕan/. The standard counterpart for the colloquial bolded phrase in this example is /al-asfar allati kunna fiha mʕan/.

(35)

kana be-Past	lana for-us	mawaqif experiences	kaθi:rah many	fiha In-it	taħadiyat challenges
wa and	baʕðʕ-ha some-3PI	mudʕhikah funny	ayðʕan too	xilal in	<b>a-ssafrat</b> Def-strips
illi that	<b>kinna</b> be-1PI	<b>mawdʒu:di:n</b> exist	<b>fiha</b> In-it		

'We had many experiences that included challenges and some of them were funny in the trips that we were part of.'

The corpus shows one instance of CA verb phrase in SA-dominant utterances. In example (36), we can see the verb phrase /gilt ma fih/, which means 'I said there is no', where the first item of this phrase is the verb /gilt/, which was pronounced colloquially instead of the standard pronunciation [qultu]. As for /ma fih/, it is a colloquial construction of the negative particle with the dummy verb /fi/ that was used instead of the standard form /la yudʒad/ or /leisa hunaka/.

(36)

ana                    **gilt**                    **ma**                    **fih**                    tadʒni:s                    siyasi  
I                        say-Past                    Neg                    there                    naturalization                    political

'I said there is no political naturalization.'

#### 4.2.2.1.3. General description of CA sentential insertions

In Table 7, CA sentential elements form 3% of the total number of CA elements in SA frames. The results indicate that those CA sentential elements are either subordinate clauses or main clauses. Table 10 shows the frequency of every type of CA sentential insertions. There are 13 CA main clauses, which are the most frequent sentential elements in the corpus at 65%. 35% of the elements are subordinate clauses (seven insertions). As found in the results of the multi-lexical elements, the sentential SA elements are also mostly cross constituents.

**Table 10. CA sentential elements in SA-dominant utterances**

Type	Number	Percentage
Main clauses	13	65
Subordinate clauses	7	35
Total	20	100

Starting with main clauses, example (37) contains a CA main clause /haðula ahali/ which is a cross constituent. The pronoun /haðula/, which means *those*, is the colloquial form of the standard /haʔulaʔ/. The noun /ahali/, which means *my people*, is the colloquial pronunciation for the standard /ahli/.

(37)

wa                    ana                    aʕrif                    a-nwab                    **haðula**                    **ahali**  
and                    I                    know                    Def-representatives                    those                    people-my  
  
f-ana                    mustaʕid                    an                    urahin                    bikul                    ma  
so-I                    ready                    to                    bet                    of-all                    what  
  
amlik                    ʕla                    haða                    al-ʔmir  
own-1Sig                    on                    this                    Def-matter

'And I know the representatives, those are my people, so I'm ready to bet everything I own on this matter.'



After CA main clauses, there are the CA subordinate clauses, which occur seven times in the corpus. In example (38), there is the CA subordinate clause /illi niʕi:ʕih o niʕuf/. Here we have a cross constituent where the first lexical item is colloquial and the rest are standard but pronounced colloquially. The colloquial relative pronoun /illi/ is used here instead of the standard form /allaði/. The verbs /niʕi:ʕih/ and /niʕuf/ are pronounced colloquially instead of the standard pronunciation /naʕi:ʕuhu/ and /narah/. Even the conjunction *and* is pronounced as [o] instead of the standard form [wa]. The SA equivalent for /illi niʕi:ʕih o niʕuf/ would be /allaði naʕi:ʕuhu wa narah/.

(38)

haða	masʕdar	wa	al-masʕdar	al-axar	hwa
this	source	and	Def-sources	Def-other	it
al-waqiʕ	illi	niʕi:ʕih	o	niʕuf	man
Def-reality	that	live-1PI-Pres-3Sig	and	see-1PI	who
hum	al-mutadʕanisi:n	fi	al-Baħrein		
they	Def-naturalized citizens	in	Def-Bahrain		

'This source and the other sources are the reality that we live, and we see who are the naturalized citizens in Bahrain.'

#### 4.2.2.2. CA-dominant utterances

In Table 6, 24% of the utterances in the corpus are CA/SA mixed utterances where CA is the dominant language. From Table 11, it can be noticed that, unlike the inserted CA elements, the majority of SA elements here are either multi-lexical or sentential. There are 75 SA multi-lexical elements that form 62%, and 33 sentential elements that form 28% of the total of 120 SA elements. The lexical and phonological SA elements constitute only 10% of the embedded elements with 5% for SA lexical insertions and 5% for the phonological insertions. It is important to mention that the number of SA insertions 120 (Table 11) exceeds the number of CA-dominant utterances 116 (Table 6), because there are some CA-dominant utterances that have more than one SA insertion. Only SA multi-lexical and sentential insertions will be discussed since they have an adequate number of tokens as compared to the phonological and lexical insertions.

**Table 11. Distribution of SA elements in CA-dominant utterances**

SA elements	Number	Percentage
Phonological instances	6	5
Lexical instances	6	5
Multi-lexical instances	75	62
Sentential instances	33	28
Total	120	100

**4.2.2.2.1. General description of SA multi-lexical insertions**

Table 12 reviews the types of SA multi-lexical elements found in CA-dominant utterances. The common type that forms 61% of the multi-lexical elements is the determiner phrases with 46 instances. The second common type is the tense phrases with 15 instances equaling 20%. Prepositional phrases appear six times in the corpus and constitute 8%. There are five adverbial phrases (7%) and three conjoined phrases (4%).

**Table 12. SA multi-lexical elements in CA-dominant utterances**

Type	Number	Percentage
Determiner phrases	46	61
Tense phrase	15	20
Prepositional phrases	6	8
Adverbial phrases	5	7
Conjoined phrases	3	4
Total	75	100

We can see here that determiner phrases are the most frequent SA multi-lexical insertions in CA-dominant utterances. It is interesting to note that this type of multi-lexical insertions is the most frequent EN insertions in the EN/Arabic utterances as well. An example of a SA determiner phrases is /istiqrar dʒami:l/ in (39). This phrase stands out in this utterance because the constituent is SA as compared to the rest of the colloquial elements.

(39)

hi	ʕindha	<b>istiqrar</b>	<b>dʒami:l</b>	fi	isratha
She	has	peace	beautiful	in	household-her
w	bnatha	al-ḥilwi:n	Allah	yaḥfiðʕhum	liha
and	daughters-her	Def-sweet	Allah	save-them	for-her
Alḥla	w	Ezein			
Alḥla	and	Ezein			

'She has a beautiful peace in her household with her sweet daughters Alhala and Ezain may Allah save them.'

After determiner phrases, tense phrases are among the most frequent SA multi-lexical insertions. Example (40) contains the SA tense phrase /aqif lahu iḥtiranman wa taqdiran/, where the colloquial equivalent would be something like /awagif lih iḥtiram o tagdir fi ha-l-mauðʕuʕ/.

(40)

i:h	kan	mumin	bi-a-salfah	iman	kamel
yes	was-1Sig	believer	in-Def-topic	belief	complete
wa Allah	ana	alhagigah	<b>aqif</b>	<b>lahu</b>	<b>iḥtiranman</b>
by God	I	truly	stand up-Pres	for-him	respect
<b>wa</b>	<b>taqdiran</b>	fi	haðihi	a-zzawiyah	
and	honor	in	this	Def-matter	

'Yes, he had complete faith in this topic. By God, I truly stand up to respect and honor him regarding this matter.'

The third most common type of SA multi-lexical insertions are prepositional phrases. Example (41) contains the SA prepositional phrase /li-l-kaḥiri:n/, where the colloquial form would be /l-wajid/, /l-wadʒid/, or /l-kiḥi:r/.

(41)

zaʕalt get angry-1Sig	dazeit send-Past-1Sig	ana I	zaʕlanah mad	ana I	mitðʕajjah upset
o and	haða this	msabib cause	xuf fear	yugul say-1Sig	lidʒ to-you
ma Neg	ħad one	yasmaʔ listen	<b>li-l-kaθiri:n</b> for-Def- many-Pl	ħisbalhum think-3Pl	abafðʕaħhum 1Sig-will-expose-3Pl

'I got angry! I sent [a text message] I am mad! I am upset and this makes them afraid. They say to you no one listens to many people. They think I will expose them.'

The results revealed that SA adverbial phrases can be embedded in CA-dominant utterances. In example (42), the colloquial form of the SA adverbial phrase /maratan oxra/ would be /marah θaniah/.

(42)

aw and	yibtidi start-3Sig	ikarir repeat-3Sig	<b>maratan</b> once	<b>oxra</b> again	ʕaʃan because
dʒiði this	ya if	leit only	tirdʒaʕ return	alθalaθeen Def-thirteen	ħalqah episodes
haði this	wayid very	ħilwah sweet			

'And it started to repeat once again, therefore, if only the thirteen episodes could be re-aired, this would be very nice.'

The least common SA multi-lexical switches are conjoined phrases. Example (43) contains the conjoined phrases /bi-dʒʔizat afdʕal ʕardʕ mutakamil wa afdʕal muxridʒ/. The phrases here are pronounced completely in SA, unlike the previous string of words. For instance, the colloquial pronunciation for /bi-dʒʔizat/ and /mutakamil/ would be /ib-dʒayzat/ and /mitkamil/.

(43)

riħt go-1Sig-Past	ʃarkit participate-1Sig-Past	bi-mahradʒan	al-qahirah	i-tadʒribi:	aw
ʃift Saw-1Sig	Suleiman Al-Bassam Suleiman Al-Bassam	kan was	mʃarik participant	b-ʔsim Of-name	britʃania UK
aw and	faz won	<b>bi-dʒʔizat</b> of-award	<b>afdʃal</b> best	<b>ʃardʃ</b> performance	<b>mutakamil</b> intergrated
<b>wa</b> and	<b>afdʃal</b> best	<b>muxridʒ</b> director			

'I participated in Cairo's festival for experimental theatre and I saw Suleiman Al-Bassam who was representing UK and won the best integrated performance and the best director awards.'

#### 4.2.2.2.2. General description of SA sentential insertions

As shown in Table 11, sentential elements form 28% of the SA insertions. The analysis indicates that different types of SA sentential elements, such as main clauses, subordinate clauses, and formulaic expressions can be embedded in CA utterances. The nature of codeswitching from CA to SA is different than it is from SA to CA. Unlike the CA sentential elements, in the SA dominant utterances, the data shows that, when the speaker is mainly using CA, the embedded SA sentential elements seem to be longer and more complicated as there are some mixed utterances with unclear frame. Table 13 shows that 73% of the SA sentential elements are main clauses. Formulaic expressions are found among the inter-sentential insertions and form 12% of a total of 33 sentential elements. Those formulaic expressions are mostly verses from the Holy Quran, idioms, or proverbs inserted in the subjects' speech. There are four unclear mixed utterances that constitute 12% of the total number of SA inter-sentential elements 33. Those are mixed utterances with unclear frames, as will be shown in the examples below. The subordinate clauses are the least occurring insertions and form 3% with only one occurrence.

**Table 13. SA inter-sentential elements in CA-dominant utterances**

Type	Number	Percentage
Main clauses	24	73
Formulaic expressions	4	12
Unclear utterances	4	12
Subordinate clauses	1	3
Total	33	100

Starting with main clauses, the SA/CA utterance in example (44) contains the SA main clause /ana la amlik illa/. The CA equivalent would be the clause /ana ma li illa/.

(44)

ana I	la Neg	amlik own-1Sig	illa except	al-ʕeif Def-living	mali my
ana I	ʕindi have	tʕabʕan Of course	kint be-past-1Sig	aʕtiyil work	fi at
qanat Channel	Dubai Dubai	ana I	bint daughter	Al-qanah Def-Channel	sitʕf sixteen
sanah year	w and	baʕdha After that	axɔt take-past	taqaʕud retirement	fa-maʕaʕf so-payment
i-taqaʕud Def-retirement	tisʕat nine	alaf thousands	dirham dirham		

'I don't own except my living. Of course I worked at Dubai Channel. I was the daughter of the channel for sixteen years and after that I retired. So the pension is nine thousand dirhams.'

The results reveal that four of the SA inter-sentential insertions found in CA-dominant utterances are formulaic expressions such as /maθna wa θulaθ wa rubaʕ wa fi ayah wa lan aʕdilu/ in example (45):

(45)

<b>maθna</b> two	<b>wa</b> or	<b>θulaθ</b> three	<b>wa</b> or	<b>rubaʕ</b> four	<b>wa</b> and
<b>fi</b> in	<b>ayah</b> verse	<b>wa</b> and	<b>lan</b> Neg	<b>taʕdilu</b> Treat-fairly	fahað so-this
maʕnatih meaning-its	sʕariħ direct	yaʕni in other words	ma Neg	taxið take-2Sig	ila except
waħdah one					

‘(two or three or four) and in a verse (You will never treat fairly) so this has a direct meaning. In other words, you only marry one person.’

There are some inter-sentential insertions with unclear frame where speakers switch from CA to SA and keep the conversation in SA. The utterance in example (46) shows a mixed utterance with an unclear frame. The SA insertion is /bi-ʔusrah raʔiʕah dʒidan mutafahimah dʒidan yadʒmaʕuna al-ħub mutafahimi:n raʔiʕi:n/.

(46)

li-llah for-Allah	alħamd Def-thank	rab-na God-our	subħanh glorified	wa and	taʕala exalted
ʕawzʕni bless-Past-1Sig	<b>bi-ʔusrah</b> with-family	<b>raʔiʕah</b> wonderful	<b>dʒidan</b> very	<b>mutafahimah</b> understanding	<b>dʒidan</b> very
<b>yadʒmaʕuna</b> bring-Pres-1Pl	<b>al-ħub</b> Def-love	<b>mutafahimi:n</b> understanding	<b>raʔiʕi:n</b> wonderful	yaʕni mean-1Sig	aʕtigid believe-1Sig
ði this	ʕaʒan because	tarbyat upbringing	baba baba	yaʕni mean-1Sig	baba baba
Allah God	yarħam-ih have mercy- him	kan was	lama when	maθlan for example	ykalim-ni talk-to-me
aw or	ħad anyone	min of	awlad-ih Kids-his	yaʕni mean-1Sig	ħad anyone
yugu:l say-3sig	li- ħad to-anyone	min of	awladi-h kids-his	miθl like	ma what
tifazʕalt say-3sig	ma- ħad Neg-anone	bi-gu:l will-say	al-kilmah Def-word	ði this	

‘Our God blessed me with a very wonderful and understanding family. Love brings us together. [They are] understanding. [They are] wonderful. I believe this is because of my dad’s upbringing I mean my dad, may God have mercy on him, was when talking to me or to anyone of his kids I mean someone says to one of his kids like what you just said. No one would say this word.’

As shown in Table 13, there is a single occurrence of a SA subordinate clause. Example (47) contains the relative clause /fi tilaka al-ħuqbah/. The colloquial equivalent would be /illi fi ðak il-wagt/.

(47)

Kint was-1sig	fi in	masraħ theatre	al-xali:dʒ Def-gulf	al-ʕarabi Def-Arabian	bidiet start-past-1sig
maʕ-ahum with-3PI	ka-muʔasisah as-founder	fi in	al-masraħ Def-theatre	θum then	maθalt performed-1Sig
aɣlab most	al- masraħyat Def-plays	<b>allati</b> that	<b>kanat</b> were	<b>fi</b> in	<b>tilka</b> that

**al-ħuqbah**  
Def-era

'I was in the Arabian Gulf theatre. I started with them as a founder of the theatre then I performed in most of the plays that were in that era.'

### 4.3. Sociolinguistics aspects of style variation

This section will be devoted to describing style variation among the study's speakers by investigating the distribution of utterance types to see the effect of three main factors: occupation, gender, and age.

#### 4.3.1. Style and occupation

In the present study, the results show that occupation has a pivotal influence on style. This can be seen from the different distribution of utterance types in the politicians' and actors' data. As illustrated in Table 14, politicians frequently used two types of utterances: the mixed SA/CA with 358 utterances (51%) and SA with 297 utterances (43%). The less common types are multilingual SA/CA/EN with 24 utterances (3%), mixed SA with EN with 11 utterances (2%), and monolingual CA with 7 utterances (1%). It is noticeable that politicians minimized their use of English and mainly focused on SA and CA. They never mixed English with CA, as seen in Table 14. They only mixed it albeit minimally with SA or with both SA and CA.



**Table 14. Distribution of utterance types for politicians**

Type	Number	Percentage
SA	297	43
CA	7	1
EN	0	0
SA/CA	358	51
SA/EN	11	2
CA/EN	0	0
SA/CA/EN	24	3
<b>Total</b>	<b>697</b>	<b>100</b>

Actors, on the other hand, showed a completely different style. As shown in Table 15, the two salient types of utterances in their data are pure CA with 1662 utterances (78%) and mixed CA and EN with 273 utterances (13%). Next, the mixed SA and CA utterances form (6%) of total of 2111. The less frequent occurring types are pure SA with 17 utterances (1%), monolingual EN with 15 utterances (1%), and the three mixed codes SA/CA/EN with 25 utterances (1%). In contrast to the politicians, actors maximized their use of English switches and minimized SA switches. It is also clear that English was never mixed with SA here.

**Table 15. Distribution of utterance types for actors**

Type	Number	Percentage
SA	17	1
CA	1662	78
EN	15	1
SA/CA	119	6
SA/EN	0	0
CA/EN	273	13
SA/CA/EN	25	1
<b>Total</b>	<b>2111</b>	<b>100</b>

#### 4.3.2. Style and gender

In order to see if gender has any effect on style, the utterances for each gender in the two groups were counted. To remind the readers, there were three males and three females in the politicians group; and six males and six females in the actors group. As shown in Table 16, mixed utterances of SA/CA form half of the speech of both males and female politicians. There are 173 (50%) SA/CA utterances in male politicians' data, and 185 (51%) SA/CA utterances for female politicians. Interestingly, the results showed that male politicians used more SA which constructs 47% of a total of 343 of their data,

while SA forms 39% of a total of 354 of female politicians' speech. Multilingual utterances, utterances that have all three codes, are the third common type in the speech of female subjects with 22 utterances (6%). The fourth frequent type in the females' data is the mixed utterances of SA and EN, which forms (2%). Male politicians, however, used less English switches and this can be seen in Table 16, where the two types of utterances that have English insertions form only 1%. It is clear that the percentage varies between four main types of utterances (SA, SA/CA, SA/CA/EN) in females' speech, while the salient types in males' data are two: SA and SA/CA. CA utterances form 1% of the speech of both genders.

**Table 16. Distribution of utterance types for politicians by gender**

<b>Gender</b>	<b>Type</b>	<b>Number</b>	<b>Percentage</b>
Male	SA	160	47
	CA	3	1
	EN	0	0
	SA/CA	173	50
	SA/EN	5	1
	CA/EN	0	0
	SA/CA/EN	2	1
	<b>Total</b>	<b>343</b>	<b>100</b>
Female	SA	137	39
	CA	4	1
	EN	0	0
	SA/CA	185	52
	SA/EN	6	2
	CA/EN	0	0
	SA/CA/EN	22	6
	<b>Total</b>	<b>354</b>	<b>100</b>

As for the effect of gender on the linguistic style of actors, Table 17 shows that the majority of the speech of both genders is in CA. Male subjects used more CA, 869 utterances (84%), compared to actresses who have 793 utterances (74%) in their data. The second common type of utterances is the mixing of CA and EN. Actresses tend to codeswitch more (CA/EN, 16%) than actors do (CA/EN, 10%). After the mixed CA/EN utterances SA, the mixed SA/CA utterances are the most frequent occurring type with 65 utterances (6%) for actresses and 54 utterances (5%) for actors. Similar to female politicians, actresses showed more varying types of as the mixed SA/CA/EN constituted 2%. The use of SA between the two genders is the opposite of that of politicians'. As for SA utterances, they are more frequent in the speech of actresses with 15 utterances (1%) and only two utterances for actors (0%).

**Table 17. Distribution of utterance types for actors by gender**

Gender	Type	Number	Percentage
Male	SA	2	0
	CA	869	84
	EN	1	0
	SA/CA	54	5
	SA/EN	0	0
	CA/EN	106	10
	SA/CA/EN	5	1
	<b>Total</b>	<b>1037</b>	<b>100</b>
Female	SA	15	1
	CA	793	74
	EN	14	1
	SA/CA	65	6
	SA/EN	0	0
	CA/EN	167	16
	SA/CA/EN	20	2
	<b>Total</b>	<b>1074</b>	<b>100</b>

### 4.3.3. Style and age

The actors group has age as an additional variable. The figures in Table 18 indicate that CA constitutes 79% of the speech of both age groups. The two groups differ when it comes to the second common type of utterances, where the older group has the mixed type SA/CA with 104 utterances (11%). The younger group, however, has the CA/EN codeswitching utterances as the second most frequent type which form 18% of their speech. The older group also tended to mix English either with CA, which forms 7%, or with CA and SA, which forms 2% of the total of 961 utterances.

**Table 18. Distribution of utterances types for actors by age**

Gender	Type	Number	Percentage
Older	SA	11	1
	CA	759	79
	EN	3	0
	SA/CA	104	11
	SA/EN	0	0
	CA/EN	63	7
	SA/CA/EN	21	2
	<b>Total</b>	<b>961</b>	<b>100</b>
Younger	SA	6	1
	CA	903	79
	EN	12	1
	SA/CA	15	1
	SA/EN	0	0
	CA/EN	210	18
	SA/CA/EN	4	0
	<b>Total</b>	<b>1150</b>	<b>100</b>

#### 4.4. Summary of the results

In this chapter, I presented the results of the analysis of the structural and sociolinguistic aspects of style variation. Starting with the structural aspects, the findings revealed that the corpus has 2808 utterances that are classified under seven categories: SA, CA, EN, SA/CA, CA/EN, SA/EN, and SA/CA/EN. The focus of this study is on two types of codeswitching. First, we have English/Arabic codeswitching that forms 12% of the corpus. Second, there is SA/CA code mixing that constitutes 17% of the total number of utterances (Table 2). While examining my data, I have noticed that the nature of SA/CA mixing is of great interest since it differs among the subjects and this difference is based on occupation. To clarify, politicians have SA/CA utterances where SA is dominant and CA is the embedded variety, whereas actors have SA/CA utterances with CA as the dominant variety and SA as the embedded one.

The analysis of the morphosyntactic aspects of each type of codeswitching showed that there are 419 English insertions in English/Arabic utterances (Table 3), 658 CA insertions in SA-dominant utterances (Table 7), and 120 SA insertions in CA-dominant utterances (Table 11). As for English/Arabic utterances, the most common type of insertions is lexical elements, which form 80% of the total number of elements of 419. Next, there are the multi-lexical insertions and the sentential insertions, where each

type forms 10% of the total number of EN elements (Table 3). As for the SA-dominant utterances, phonological instances constitute 58% of the total number of CA elements followed by the lexical insertions at 33%. Lexical and multi-lexical insertions are the least common insertions and form 6% and 3%, respectively (Table 7). The morphosyntactic distribution of SA in CA-dominant utterances exhibits a different nature compared to the previous English/Arabic and SA-dominant utterances. In CA-dominant utterances, the SA multi-lexical feature is the most frequently occurring type with 75 instances 62% followed by 33 sentential insertions which form (28%) of the total of 120 CA elements (Table 11). SA phonological and lexical elements are the least occurring here with only six instances for each type.

The second main part of this chapter is about the description of the sociolinguistic aspects of stylistic variation. There are three factors in this study: occupation, gender, and age. The results indicate that occupation plays a pivotal role in stylistic variation. Although politicians tried to maintain using SA in their speech, insertions of CA were highly frequent and this can be seen in the number of SA/CA utterances (358), which form 51% of their speech (Table 14). Actors, on the other hand, mainly use CA (Table 15). They use English insertions more than SA insertions in their speech. As for gender, both female and male politicians have the same linguistic behavior except for females who use more English insertions (Table 16). Actresses also have the lead in using both EN/Arabic and SA/CA utterances as compared to actors (Table 17). Actors have age as an additional factor. The results show the younger group adapt a speech style where English is the main embedded language, whereas SA is the main embedded language for the older group (Table 18).

## **Chapter 5.**

### **Discussion and conclusion**

This chapter interprets the findings presented in the previous chapter. Section 5.1 discusses the overall pattern of language use among the study's subjects. Subsections 5.1.1 and 5.1.2 cover the structural characteristics of this pattern. In section 5.2, the effect of social factors will be explained and compared to previous studies. The last section, 5.3, summarizes the main findings of the study and provides directions for future research.

#### **5.1. The structural characteristics of the overall pattern of language use**

It is clear that there are four types of utterances that are frequent in the corpus. The types are arranged from the most to least frequent as follows (Table 2): CA (59%), SA/CA mixed utterances (17%), SA utterances (11%), and Arabic/EN utterances (10%). There are limited numbers of studies in the literature that focus on diglossia and codeswitching simultaneously. Researchers focus either on codeswitching between English and Arabic (Atawneh, 2007; Al Batoush, 2014; Ismail, 2015; Al-Hayek, 2016) or the diglossic switching between SA and colloquial varieties (Alshamrani, 2008; Bassiouney, 2009; Al-birini, 2011; Ababtain, 2013; Dashti & Dashti, 2015). Codeswitching and diglossia are two salient linguistic behaviors in the corpus. Myers-Scotton's (1993a) Matrix language-Frame model (MLF) is used to analyze both types of mixed utterances. In the next sections, the findings of the structural characteristics will be discussed.

##### **5.1.1. Diglossia**

Not only do the results show that the SA/CA mixed utterances are different in terms of the dominant language but also the insertions of the embedded language vary. There is no doubt that the formality of setting is a major influence on the use of diglossic switching (Holes, 1995) in this study. In SA-dominant utterances, where the subjects codeswitched from SA to CA, phonological and lexical elements constitute 91% of the

CA insertions (Table 7). Even when we look at the CA lexical elements, we find that 76% of them are CA personal, relative, and demonstrative pronouns (Table 8). In contrast to SA-dominant utterances, multi-lexical and sentential elements form 90% of the SA insertions in the utterances dominated by CA (Table 11). The two types of mixed utterances indicate that diglossic switching is employed differently to craft different linguistic styles. SA is a dominant language in a more formal setting and an embedded language as the setting gets less formal.

In his study of SA/CA codeswitching, Albirini (2011) finds that diglossic switching occurs in all three forms of discourse: religious, political, and sports announcing. He states that code choices are not only restricted to the formality of the context, as Ferguson (1959) states, but it “depends largely on the function to be performed and its relation to the High or Low code” (p. 558). This is similar to the findings of this study. Diglossic switching is found in both formal (politicians’ data) and informal (actor’s data) settings. The present study is another indication that SA and CA are not in a complementary distribution as stated in Ferguson’s classic definition of diglossia. Furthermore, the results of the present study confirm Hary’s (1996) notion of the relationship between SA and CA. Pure SA and pure CA do not exist in Arabs’ speech. What exists is a mixture of both varieties, which is found in the speech of this study’s subjects.

### **5.1.2. Codeswitching**

The data reveals that English/Arabic utterances form 12% of the corpus and occur on the lexical and the phrasal levels (Table 2 & Table 3). According to Myers-Scotton (1993a), there is one language, i.e., the matrix language (ML), that dominates in the interaction of many languages by supplying the mixed utterances with morphosyntactic frames. The matrix language of Arabic/English mixed utterances is Arabic. English in these cases is the embedded language (EL), which has two functions. First, it provides the multilingual utterances with content morphemes such as nouns, verbs, prepositions, and adjectives. Second, it provides EL islands, which are monolingual well-formed phrases. The findings concerning English insertions are consistent with the findings of previous studies on Arabic/English codeswitching such as Alrowais (2012), Ben Nafa (2013), Hamouda (2015), and Kniaz (2017), who found that nouns and adjectives are the most frequent lexical switches. Gardner-Chloros (2009)

states that English nouns are grammatically self-contained characters; therefore, they are the most inserted elements in Arabic speech since they are structurally less difficult to be integrated.

The results reveal that there is a relationship between English insertions and the dominant variety. For example, English insertions occur more in CA than in SA in the data of this study, which is similar to the findings of Mahmoud (2013), who examined the influence of English on MSA by collecting and analyzing data from books, articles, manuals, reports, recommendations and TV programs. Another example on the relationship between English insertions and the dominant variety can be seen in the nature of the insertions themselves. The results show that English insertions mostly maintain their original linguistic features if they occur in CA contexts. When occurring in SA contexts, however, they are elements that are referred to as *cultural borrowings* by Myers-Scotton (1993a). Cultural borrowing forms “represent objects and concepts new to the ML culture” (p. 169). They have previously undergone a process of Arabization where they are phonologically and morphologically altered to suit Arabic morphosyntactic structure and then have become part of the Arabic lexicon.

A good example of a cultural borrowing would be the Arabized word /lu:jisti:/ for the English adjective *logistic* in the SA utterance in example (1). This word has undergone a process of change on two linguistic levels. Phonologically speaking, the English stress is absent in the Arabized form /lu:jisti:/. Besides phonological changes, the English adjective suffix /-ic/ is omitted from the Arabized form to be compatible with Arabic morphological rules regarding adjectives.

(1)

al-wilayat	al-mutaḥidah	tuqadim	daʕm	istixbarati	wa
Def-states	Def-united	prodiver-Pres	support	intelligence	and
<b>lu:jisti:</b>	li-dwal	a-taḥaluf	fi	ʕamalyat-ih	fi
logistic	for-countries	Def-coalition	in	operations-its	in

al-yemen  
Def-Yemen

‘The US provides intelligence and logistic support for the coalition for its operations in Yemen.’



## **5.2. Social characteristics and persona construction**

It is hypothesized that the two groups will use style as a means to construct two distinct identities: a socially conservative identity for the politicians and a more cosmopolitan one for the actors. It is further hypothesized that SA will be present in the speech of politicians for its association with conservative values, and the actors will use CA as well as English because these two codes are part of a more popular persona. The findings support these hypotheses. The next sections will discuss the results of the social characteristics of the styles of politicians and actors in light of Eckert's (2004) framework.

### **5.2.1. Style and occupation**

#### **5.2.1.1. *The conservative persona***

Albirini (2016) states that, in order to construct national identities, speakers differentiate between two varieties: SA and CA. Arabs perceive SA as a variety that denotes Arab unity; thus it is firmly related to the ideology of Pan-Arab nationalism. CA, on the other hand, is restricted in use to certain regions and is connected in the Arabs' minds with the historical, political, and economic decline of their ethnicity.

This language ideology is shown in the conservative persona that politicians project by their speech style as they try to maintain the use of the ultimate SA. This is clearly indicated by the percentage of pure SA utterances that form 43% of their data (Table 14). Although they are expected to solely use SA in such formal settings (Ferguson, 1959), CA is also found in their data. The results show that 51% of the overall utterances are diglossic (Table 14). The structural description of the diglossic utterances in the previous chapter give us some insights that helps explain the unexpected CA occurrences as well as the nature of diglossia in their speech. There are two types of diglossic utterances: CA-dominant utterances and SA-dominant utterances. Almost all the diglossic utterances in the politicians' data are SA-dominant. When taking even a closer look at the distribution of the CA elements in those utterances, we find that 91% of the embedded CA elements are either phonological (58%) or lexical (33%), while the remaining 9% are multi-lexical (6%) and sentential (3%) as shown in Table 7. These results indicate that politicians attempt to adhere to SA and the majority of the CA use is limited to phonological and lexical features. Another indication of the politicians' attempts

to avoid CA is the cross constituents that form 70% of their CA multi-lexical elements (Table 9). It is noticed that when they switch to CA, they gradually transform to SA as mentioned in example (30) in the previous chapter.

The structural analysis of the politicians' data indicates that diglossia is a salient characteristic of their speech. The results are similar to those of Mazraani's (1997) who thoroughly studies speech variation in three famous Arabic politicians: the Egyptian president Gamal Abdel Nasser, the Iraqi president Saddam Hussein, and the Libyan leader Muammar Al-Gaddafi. All three of them share almost the same style in that they switch between SA and CA. In a more recent study, Alshamrani (2008) studies diglossia in Arabic TV stations and points out that, although SA is expected in political contexts, his results show that SA and CA are mixed in the speech of guests in political programs. Versteegh (2014) attributed this style to the fact that politicians talk to an audience from all walks of life, therefore, switching to CA is crucial to make sure that their messages are conveyed and well understood. However, they do switch back to SA to promote Arab unity and pride. Pride and Arab unity are culturally and nationally very crucial concepts that politicians strive to project to Arabs in the Middle East.

Although politicians have a good command of English, they are aware of the language ideology stating that in order to construct their political personae, they should not show any interest in using foreign languages. Albirini (2016) asserts that there is a strong link between identity in the Arab world and the Arabic language. Speaking Arabic serves as a criterion that determines who Arabs are. Through the Arabic language, speakers can project many identities, whether national, ethnic, or religious. After the decline of Arab nationalism in the 1990s, the Arabic language became a key element in constructing an ethnic identity that is associated with cultural and historical heritage. Arabs believe that their ethnicity is threatened by the incursion of borrowings from other languages, such as English. This is well reflected in the politicians' speech, where there is not a single occurrence of a pure English utterance. In the process of persona construction, politicians distance themselves from English except for some words, i.e., cultural borrowings, that are Arabized and have become part of the Arabic linguistic repertoire such as the word in the above example (1).

### **5.2.1.2. *The popular persona***

Code choice is of a different nature in the speech style of actors compared to politicians. As seen in Table 15, CA is the commonly used variety with 1662 utterances (79%) of a total of 2111. This confirms the observation of previous researchers that CA is the variety of mass entertainment media (Badawi, 2006; Al-Sobh et al., 2015). Using CA has certain communicative effects that cannot be gained by using SA. By using CA, the speaker wants to strengthen their personal connection with people and to deliver audience-inclusive messages (Stadlbauer, 2010). Those communicative effects are important for actors since one of their career's tasks is to keep a strong relationship with the audience. Therefore, a speech style that is mainly dominated by CA helps actors construct a popular persona. The present results regarding CA being the main variety used by actors seem to be consistent with other research. For example, Alshamrani (2008) investigates Arabic diglossia and the settings under which it is used in three selected Arabic TV stations. He finds that actors in an interview program used CA 87% of the time. In another study, Dashti and Dashti (2015) carry out a qualitative study on the diglossic interaction of Kuwaiti presenters in Kuwaiti TV channels. They find that CA is mainly used in less formal programs such as sport programs.

In addition to monolingual CA utterances, the findings indicate that codeswitching and diglossia are two characteristics of their speech style. It is interesting to note that actors tend to switch between English and CA (13%) more than they do between SA and CA (6%) (Table 15). In constructing their personae, English appears to be preferred to SA as an embedded language. This finding is similar to that of Dashti and Dashti (2015) who spot codeswitching between CA and English in beauty programs in their study. The two linguistic behaviors, codeswitching, and diglossia, will be further examined in the following sections as gender and age play pivotal roles in how actors use language to project their images.

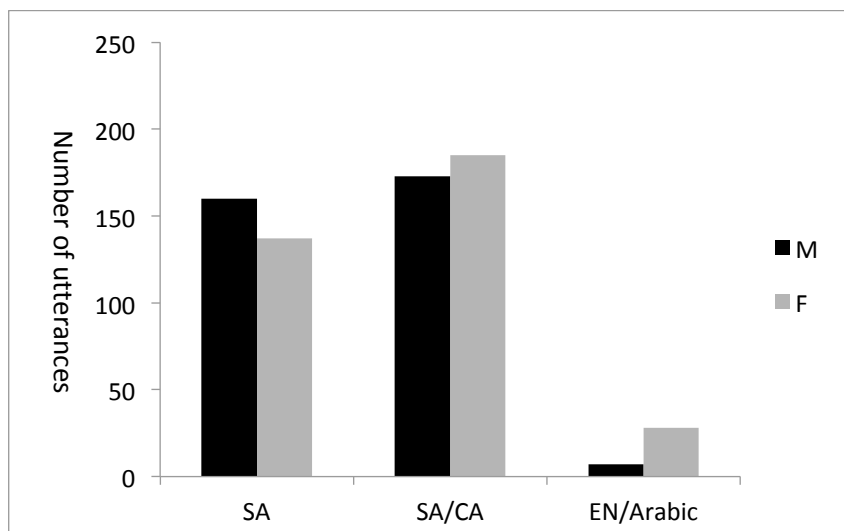
### **5.2.2. Style and gender**

The role of gender as a factor in sociolinguistics has been recognized for decades (Abd-el-Jawad, 1981; Sadiqi, 2003b; Atawneh, 2007; Holmes & Wilson, 2017; Sayahi, 2011; Al Batoush, 2014; Ismail, 2015). Gender differences in language use range from instances where the two genders speak different languages to other

instances where the difference is just in linguistic features such as pronunciation and vocabulary (Holmes & Wilson, 2017). Eckert (2012) insists that correlations between style and social factors are not the best ways to explain or even understand stylistic variation. Instead, gender must be understood and studied in its social context or, in other words, gender must be studied by investigating the social practices of speakers. Since politicians and actors display two distinguishable styles, they will be separately discussed in the next section.

#### **5.2.2.1. *Gender and the politician persona***

The results indicate that male and female politicians use the same speech style but with small differences. Although those differences are not significant, they are similar to the results of previous studies on linguistic variation in the Arab world where women use less standard forms compared to men (Schmidt, 1974; Salam, 1980; Abd-el-Jawad, 1981). The findings show that female subjects lead in using the mixed utterances of the three codes. In fact, a female politician from the United Arab Emirates uses six English insertions that have Arabic equivalents. Those minor gender differences, illustrated in Figure 2, might be attributed to the fact that in this study female politicians discussed domestic-related topics, such as improvements introduced to the country's different sectors, that are of interest to their citizens, whereas the male politicians discussed topics that are concerning on an international level, such as the political conflicts threatening the sovereignty of the Gulf Cooperation Council countries, and they attract a larger audience (other governments). Both female and male politicians use linguistic style to construct a political persona, but the intensity of this identity decreases when the speech is delivered to their citizens. This is noticed by Coupland (2000) who states that stylistic variation is used by the speakers to project an identity that either reduces or increases the social and cultural differences with another identity (the interlocutors' identity).

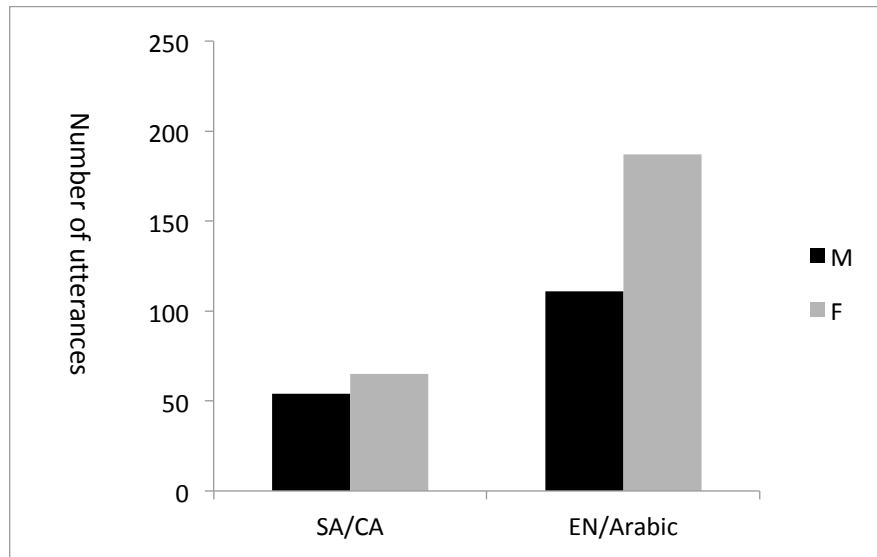


**Figure 2. Distribution of monolingual SA, diglossic, and codeswitched utterances for politicians by gender**

Being less conservative when talking to a local audience and more conservative when talking to an international audience is found in the speech of a former president of Murcia, the autonomous region in southeastern Spain (Hernández-Campoy & Cutillas-Espinosa, 2012). She is known for her heavy use of vernacular forms in formal and informal settings. Her linguistic behavior is a topic of heated debate among people of Murcia as well as the media. Hernández-Campoy and Cutillas-Espinosa (2012) quantitatively and qualitatively analyzed her speech using speaker design. The results showed that she uses more dialectal features in her speech to a local audience in order to project a social identity. However, more standard features are found in her speech to a national audience to project a more political persona.

**5.2.2.2. Gender and the popular persona**

The findings show that there is no significant difference in the frequency of diglossic utterances among actors and actresses, but they behave differently in terms of codeswitching. Figure 3 illustrates that the actresses have more codeswitched utterances than the actors. Therefore, the focus in this section will be on gender differences regarding the distribution of English/Arabic utterances.



**Figure 3. Distribution of diglossic and codeswitched utterances for actors by gender**

As Figure 3 shows, actresses codeswitch 25% more than actors. Moreover, the English insertions found in their data are linguistically complex and advanced. Actors' insertions, on the other hand, are mostly Arabized and restricted to the jargons usually used in their domain such as stage, action, and drama, etc. To understand this distinct use of codeswitching among actors and actresses, their social practices should be examined (Eckert, 2012). In addition to acting, many actresses in the Arabian Gulf region entered the world of fashion and beauty. This is due to the fact that actresses feel threatened by the rising phenomenon of social media influencers and fashion/beauty bloggers whose large fan base grants them many opportunities whether it is to collaborate with well-known brands and companies worldwide or to break into the world of acting. Social media influencers "represent a new type of independent third party endorser who shape audience attitudes through blogs, tweets, and the use of other social media" (Freberg, Graham, Mcgaughey, & Freberg, 2011, p. 90). They "work to generate a form of "celebrity" capital by cultivating as much attention as possible and crafting an authentic "personal brand" via social networks, which can subsequently be used by companies and advertisers for consumer outreach" (Hearn & Schoenhoff, 2016, p. 194).

In a recent study that investigates the impact of traditional celebrities vs. social media influencers endorsements on advertising effectiveness, it is concluded that "participants identify more with influencers than celebrities, feel more similar to

influencers than celebrities, and trust influencers more than celebrities” (Schouten, Janssen, & Verspaget, 2019, p. 17). With their large number of followers that exceeds the actresses’ followers, social media influencers have not only taken advertising opportunities in the Gulf region from actresses, but also succeeded in getting acting roles in major productions. This can be seen in the complaints of many actresses about the new casting criterion that most directors and production companies have followed lately, which focuses on fame rather than talent (Alenizi, 2019; Jomah, 2018; AlArabiya, 2016; Hijab, 2016). An actress comments on this by stating that many actresses have started to imitate the lifestyle of social media influencers, as they felt deprived of the fame that they used to have in the past (Aleilah, 2019).

In order to compete with social media influencers and fit in this industry that is largely dominated by English as an international language of communication (Albirini, 2016), actresses need to project a more cosmopolitan image that can help them reach and keep contact with the growing social media users from different countries worldwide. By expanding their fan base and showing their competence in English, they will not only be appealing to fashion and beauty brands that choose influencing and popular figures with large number of followers for advertising contracts or collaboration to promote their products, but they will also maintain their acting career. The shift of actresses’ focus to beauty and fashion can be seen from the considerable time they spend and the great efforts they make on photo sessions or make-up tutorials that they share with their fans on social media. An actress, for instance, launched her own lash line and others use their social media accounts to advertise some beauty products. This shift has caused an issue in the acting field where actresses are criticized for indulging in the social media world at the expense of their profession (Hijab, 2016).

For actors, the situation is different since the fashion and beauty industry does not intensively attract men in the region and therefore they do not need to compete in this field. Although the beauty and cosmetic industry is usually linked to femininity, the situation has changed in recent decades when many cosmetics companies started to launch men-specific products. Mounghkem and Surakiatpinyo (2010) state that the men’s beauty market is considered new in relation to the women’s market. In addition to being relatively new, there are many cultural differences regarding male consumption in which the percentage of male cosmetic consumption in Western countries is higher than it is in the developing countries (Khan et al., 2017; Souiden and Diagne, 2009). In spite

of the fact that many studies conclude that men show positive attitudes towards cosmetic consumption, societal beliefs play a major role in the type of consumed products (Erosy et al., 2015; Khan et al., 2017).

On Boutiqaat, an online beauty and fashion shop store launched from Kuwait in 2015, there are 408 celebrities and social media influencers who display their personal recommendations to customers. Every celebrity has their own page or virtual boutique where they promote and recommend cosmetic products and fashion items to their fans. Men form only 2.9% of the website celebrities and all of them are neither actors nor from the Arabian Gulf region. As for actresses, they are from the Gulf region and form 18% of Boutiqaat's registered celebrities. The remaining 79.1% are social media influencers, famous makeup artists, and fashion and beauty bloggers. This reveals some facts about the different social practices between actors and actresses regarding interest in the beauty industry. Actresses use a distinctive style that is characterized by linguistically advanced English insertions in order to construct a cosmopolitan persona that enables them to compete with other social media influencers and bloggers and to fit in the beauty industry.

The present study explains the reason behind avoiding codeswitching by the male participants in the study of Al Kaddour and Kaddoura (2019). In their study of the usage of codeswitching and code-mixing, they find that male speakers never codeswitch to English and the results of the structured interviews show that they actually consider codeswitching a feminine feature. From the above explanation of the social practices of both genders in the present study, excessive use of English insertions is perceived by men as a feature of feminine style that they avoid in constructing their identities.

### **5.2.3. Style and age**

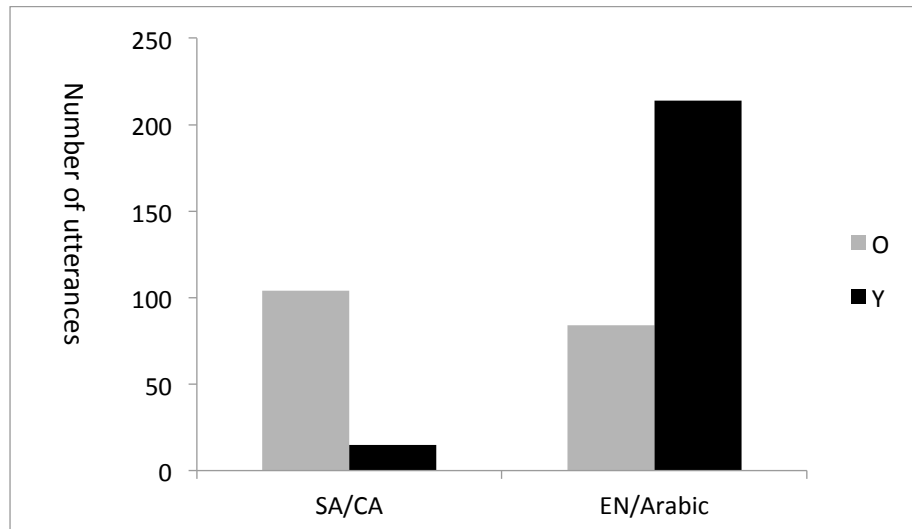
Age is an extra grouping factor that is used for actors in this study. There are several studies about age and linguistic style in the Arabic world. Al-Kaddour and Kaddoura (2019) study codeswitching and code-mixing in the speech of 20 Emirati speakers. They found that younger speakers codeswitch between English and Emirati Arabic more than the older group. Qudah (2018) investigates diglossia in Jordanian accounts on social media. He found that the younger generation uses Jordanian Arabic more than SA. As shown in Figure 4, the results indicate an opposite trend regarding the



usage of codeswitching and diglossia between the two different age groups of actors and actresses. The older group integrates SA in their speech to project a traditional identity. Arabic speakers are sentimentally attached to SA, which is strongly rooted in and related to the culture of the Arabs and the Muslim faith (Al-Wer, 1997). Most of the SA sentential insertions are found in the older actors' data. Those insertions are basically formulaic expressions such as verses from the Quran, idioms, and proverbs. Although the findings also show that the older group's data contains English insertions, they are mostly Arabized jargons which are directly connected to their profession. Their English usage resembles that of the politicians. There is only one actress, from the United Arab Emirates in the older group whose speech involves 20 lexical and multi-lexical English insertions that are not Arabized or related to the acting career. This finding is similar to the unexpected occurrence of six English insertions by a female politician who is also from the United Arab Emirates. This could be used as an indication that the role of English in the United Arab Emirates might be different than it is in the rest of the Gulf Cooperation Council countries.

It is noticeable that the younger group tends to embed English insertions more than SA in their speech. Their data contains only 15 utterances that are diglossic. The usage of English more than SA among them might be due to an effort to construct a globalized persona. To establish this identity, an individual should reflect some desirable traits, such as modernity and sophistication, which are more attained by English (Albirini, 2016). In fact, two speakers of the younger group expressed their ambitions of becoming Hollywood stars and being recognized internationally in their career. To achieve such ambitions, English is an important linguistic resource that helps them to craft a speech style from which they can project their modernized and globalized personae.

This shift from SA to CA and foreign languages among younger generations is discussed by many Arab linguists as being a threat to SA (Al-Sobh et al., 2015). The attitude of younger generations towards English as a necessary language in constructing a more modernized persona, the reform of educational systems in the Gulf Cooperation Council countries where English is mandated as the language of instruction in post-secondary education, and the important function of English as an international language are contributing factors that can initiate a change for the role of English in the Arabian Gulf region.



**Figure 4. Distribution of diglossic and codeswitched utterances for actors by age**

### 5.3. Conclusion

The purpose of this study has been to investigate style variation in two contrasting groups of public figures in the Arabian Gulf region: politicians and actors. My goal was to assess how linguistic style can be used by speakers as a tool to construct their identities. Another important goal was to study the rising phenomenon of English borrowings in the speech of Arabs in this region. I undertook this study due to the linguistic changes that have been taking place in most of the Gulf Cooperation Council countries where English has become the language of instruction in post-secondary education and the international language of communication of the Gulf citizens with people worldwide online. I wanted to investigate how SA, CA, and English are used and to what extent each language is represented in the subjects' speech. I wanted to know what connotations each language can offer in the process of persona construction. By choosing two very distinct groups in terms of overall style, I gained a clearer understanding of how the interplay between the two varieties of Arabic and English was utilized to project certain images. From my analysis of the gathered data, I discovered many interesting data about the range of identities that the subjects constructed through their distinct linguistic styles and the link between social practices, social characteristics, and language use in the process of self-representation.

The findings in general confirm the hypotheses. The subjects deploy SA, CA, and English to construct two recognizably different personae: a conservative persona for politicians, and a cosmopolitan one for actors. More than half of the politicians' utterances are a mixture of SA and CA, while they also use SA to a great extent. Actors, on the other hand, adopt a different style in which CA is the dominant code, but they also use utterances that are a mixture of CA and English borrowings (12% of their speech). These findings can be seen as an indication of a linguistic change in the region: CA forms are being inserted more frequently into SA, not only in the speech of popular figures such as actors but also in that of more conservative speakers such as politicians. The appearance of English borrowings in the utterances of the actors is a strong indication that the local linguistic repertoire is expanding, which is a significant finding on its own. It is also another piece of evidence that linguistic norms in the Arab Gulf region might be shifting towards the vernacular. Qudah (2018) expresses his concern regarding the increasing use of vernaculars, as they are not mutually intelligible and recommends that Arabic speakers must increase their use of SA. Al-Sobh et al. (2015) point out that SA faces serious challenges such as diglossic switching and shifting to foreign languages. Holes (2011), who seems to be the most positive observer of the linguistic change as compared to other Arab linguists, states, "Arabic is not dying, but rather ... the linguistic landscape of the Gulf, and perhaps that of the wider Arab Middle East, is certainly changing." (p. 144).

This study is evidence for the efficiency of speaker design, which treats speakers as proactive agents instead of reactive ones. This in turn helps interpret stylistic variation by considering all the ideological, cultural, or social factors that might affect speakers' choice of codes to project certain personae. I found that there was a distinguishable style for each of the following groups: politicians, older actors and actresses, younger actors, and younger actresses. Those styles vary in the use of SA, CA, and English in order to project and construct four personae that ranged from being very conservative to more cosmopolitan. Starting with politicians, their style was characterized by the use of monolingual SA utterances and SA/CA mixed utterances. The avoidance of English was clear in their data as they restricted it to lexical elements that either are Arabized or have no Arabic equivalents. The attempts to hold their interviews in SA, to resist using CA, and to avoid English as much as they could were important to construct a conservative persona that was needed to deliver their national and political messages. Then, it was

the older actors and actresses whose style was identified by the use of monolingual CA utterances, SA intra- and inter-sentential embedded elements, and the restriction of English use to their career jargons. This style helped them to construct a popular persona flavored with a traditional sense that enabled them to strengthen their connection with the fans and to express their emotional attachment to the Arabic culture.

The younger actors and actresses appeared to be the least conservative with language use. Younger actors used a style that I considered plain. In constructing their persona, they focused on CA with integrating lexical English borrowings. In terms of codeswitching between English and Arabic, their data showed more English borrowings compared to older actors and actresses but fewer and less advanced ones compared to the younger actresses. The most interesting style was found in the younger actresses' data which was identified by the use of CA and the expanded use of English embedded elements that were more complex and various compared to the rest of the subjects. When investigating their social practices, it was clear that younger actresses utilized English besides CA to project an image of a popular actress as well as a beauty and fashion influencer that can relate and be successful in the beauty industry.

The appearance of English borrowings in the data can mirror the importance of English in the region. This language was selected by the younger actresses and actors in the construction of their speech style to project a more modernized and sophisticated image that allows the younger actresses to fit in the beauty industry and to help both younger actresses and actors achieve international recognition that can enhance their acting career.

This study has made a number of useful contributions to stylistic variation research. It has highlighted some important findings about the relationship between language and persona construction in the Arabic social context. The present research has revealed that Arabic is not the only language required in the process of persona construction in the Arabian Gulf. The four conspicuous linguistic styles, that gradually transformed from being saturated with SA to being saturated with CA; and from avoiding foreign language insertions to embracing English insertions, added to our knowledge about the change in the linguistic landscape in general and the role of English in a region considered part of the Expanding Circle.

One thing we must keep in mind when looking at the subjects is that they are public figures and have certain roles to fulfill. Therefore, they do not necessarily represent the majority of speakers in the Arabian Gulf region. It would be better if the number of subjects for politicians and actors was equal, but the situation of politics in the Gulf region is relatively different which makes it difficult to find politicians from a younger age group. Future studies should investigate other social groups in the Arabian Gulf that have different image requirements (such as social media influencers, athletes, and businessmen/women) in order to determine how they use diglossic and codeswitching choices in the construction of their linguistic style. Furthermore, participants from other countries of the Gulf Cooperation Council could shed light on the regional differences in terms of language usage.

## References

- Ababtain, H. (2013). *Saudi Arabic diglossia and codeswitching in Twitter: Education and gender effect*. [Blog post]. Retrieved from <https://languageandsocieties.wordpress.com/2013/04/15/saudi-arabic-diglossia-and-code-switching-in-twitter-education-and-gender-effect/>
- Abd-el-Jawad, H. (1981). *Lexical and phonological variation in spoken Arabic in Amman* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.
- Abd-el-Jawad, H. (1987). Cross dialectal variation in Arabic: Competing prestigious forms. *Language in Society*, 16, 359-368.
- Abdul-Raof, H. (2013). *Subject, theme and agent in Modern Standard Arabic*. London: Routledge.
- Abu-Melhim, A. (1991). Code-switching and linguistic accommodation in Arabic. In B. Comrie & M. Eid (Eds.), *Perspectives on Arabic linguistics III: Papers from The Third Annual Symposium on Arabic Linguistics* (pp. 231-250). Amsterdam: John Benjamins.
- Adouane, W., & Johansson, R. (2016, May 23). Gulf Arabic linguistic resource building for sentiment analysis. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odiijk, & S. Piperidis (Eds.), *Paper presented at the Proceedings of the Tenth International Conference on Language Resources and Evaluation Conference (LREC)*, (pp. 2710-2715). Portorož, Slovenia.
- Ahmed, K. (2010). English in the Arab Gulf. *Asian Journal of University Education*, 6 (1), 1-12.
- AlArabiya. (2016, October 16). *Marwah Muhammed: I do not regret the harassment scene in Shad Balad* [Video file]. Retrieved from <https://www.youtube.com/watch?v=rXG03oym7Bo>
- Al-Batal, M. (2002). Identity and language tension in Lebanon: The Arabic of local news at LBCI. In A. Rouchdy (Ed.), *Language contact and language conflict in Arabic: Variations on a sociolinguistic theme* (pp. 91-115). London: RutledgeCurzon Press.
- Al Batoush, M. A. (2014). English words in colloquial Jordanian Arabic. *International Journal of Linguistics*, 6 (2), 98-108.
- Albirini, A. (2011). The sociolinguistic functions of codeswitching between standard Arabic and dialectal Arabic. *Language in Society* 40, 537–562.

- Albirini, A. (2016). *Modern Arabic sociolinguistics: Diglossia, variation, codeswitching, attitudes and identity*. London: Routledge.
- Aleilah, Y. (2019, July 15). Maram to Alanba: "Whoever beats his drum, they say we are first". *Alanba*. Retrieved from <https://alanba.com.kw/913546>
- Alenizi, F. (2019, June 6). Huda Salah: Positivity and pay are the reasons for my absence from acting. *Kuwait's Sahafah Net*. Retrieved from <http://kw.sahafahn.net/news5212116.html>
- Al-Hayek, R. (2016). *Arabic-English code-mixing by Jordanian University students* (Doctoral dissertation). Western Sydney University, Australia.
- Al Kaddour, N., & Kaddoura, R. (2019). The use of codeswitching and code-mixing by speakers of Emirati Arabic (EA). *Journal of Literature, Languages and Linguistics*, 52, 59-63.
- Alrowais, H. (2012). *Code switching between Arabic and English: Social motivation and structural constraints* (Unpublished MA thesis). Ball State University, Indiana.
- Al-Saidat, E. (1999). *The written and spoken Arabic of south Jordan: A study of diglossia* (Unpublished Ph.D. Dissertation). University of Mysore. India.
- Al-Saidat, E., & Al-Momani, I. (2010). Future markers in Modern Standard Arabic and Jordanian Arabic: A contrastive study. *European Journal of Social Sciences*, 12 (3), 397-408.
- Alshamrani, H. (2008). Diglossia in Arabic TV stations. *Journal of King Saud University*, 24, 57-69.
- Al-Sharif, N. (2017). *The sound of Aldadu in the Arabic language: A historical descriptive study* (Unpublished MA thesis). Islamic University of Gaza, Palestine.
- Al-Sobh, M., Abu-Melhim, A., & Bani-Hani, N. (2015). Diglossia as a result of language variation in Arabic: Possible solutions in light of language planning. *Journal of Language Teaching and Research*, 6 (2), 274-279.
- Al-Toma, S. (1969). *The problem of diglossia in Arabic*. Cambridge, MA: Harvard University Press.
- Al-Wer, E. (1997). Arabic between reality and ideology. *International Journal of Applied Linguistics*, 7 (2), 251-65.
- Asali, S. (2011). *Attitudes of Arab American speakers in the USA towards English-Arabic codeswitching* (Master's thesis). Middle East University, Jordan.

- Atawneh, A. (2007). English loanwords. In K. Versteegh, M. Eid, A. Elgballi, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 29-35). Leiden: Brill.
- Auer, P. (2002). *Codeswitching in conversation: Language, interaction and identity*. New York, NY: Routledge.
- Badawi, E. (2006). Foreword. Arabic for non-native speakers in the 21st century: A shopping list. In K. M. Wahba, Z. A. Taha, & L. England (Eds.), *Handbook for Arabic language teaching professionals in the 21st century* (p. ix-xiv). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Bakir, M. (1986). Sex differences in the approximation standard Arabic: A case study. *Anthropological Linguistics*, 28 (1), 3-9.
- Balamoti, A. (2010). *Code switching as a conversational strategy: Evidence from Greek students in Edinburgh* (MSc thesis). The University of Edinburgh, Edinburgh.
- Bateson, M. C. (2003). *Arabic language handbook*. Washington, DC: Georgetown University Press.
- Bassiouney, R. (2009). *Arabic sociolinguistics: Topics in diglossia, gender, identity, and politics*. Washington, DC: Georgetown University Press.
- Bell, A. (1977). *The language of radio news in Auckland: A sociolinguistic study of style, audience, and subediting variation* (Unpublished doctoral dissertation). University of Auckland, New Zealand.
- Bell, A. (1984) Language style as audience design. *Language in Society*, 13 (2), 145-204.
- Bell, A. (2001). Back in style: Reworking audience design. In J. R. Rickford & P. Eckert (Eds.), *Style and sociolinguistic variation* (pp. 139–169). Cambridge: Cambridge University Press
- Ben Nafa, H. (2013). *Codeswitching among Arabic-English adult bilinguals in the UK: Syntactic structures, and pragmatic functions* (Master's thesis). The University of Manchester, England.
- Bentahila, A. (1991). Standards for Arabic: One, two or many? *Indian Journal of Applied Linguistics* 17 (1), 69-88.
- Bentahila, A., & Davies, E. (1983). The syntax of Arabic-French codeswitching. *Lingua*, 59 (4), 301-330.
- Bentahila, Abdelâli, & Davies, Eirlys. (1998). Codeswitching: An Unequal Partnership? In R. Jacobson (Ed.), *Codeswitching worldwide* (pp. 25-49). Berlin, Mouton de Gruyter.



- Biber, D., & Finegan, E. (Eds.). (1994). *Sociolinguistic perspectives on register*. Oxford, England; New York, NY: Oxford University Press.
- Biber, D., & Conrad, S. (2019). *Register, genre, and style* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Bolinger, D. (1975). *Aspects of language*. New York: Harcourt Brace Jovanovich.
- Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge: Cambridge University Press.
- Boutiquaat. (n.d). Retrieved from <https://boutiquaat.com>
- Boyle, R. (2012). Language contact in the United Arab Emirates. *World Englishes*, 31 (3), 312–330.
- Bucholtz, Mary. (2015). The elements of style. In D. N. Djenar, A. Mahboob & and K. Cruickshank (Eds.), *Language and identity across modes of communication* (pp. 27-60). Berlin: De Gruyter Mouton
- Bucholtz, M., Bermudez, N., Fung, V., Edwards, L., & Vargas, R. (2007). Hella nor cal or totally so cal?: The perceptual dialectology of California. *Journal of English Linguistics*, 35 (4), 325-352.
- Chambers, J. (1995). *Sociolinguistic theory*. Cambridge, MA: Blackwells.
- Christoffersen, K. (2014). Does child codeswitching demonstrate communicative competence?: A comparison of simultaneous and sequential bilinguals. *Arizona Working Papers in SLA & Teaching*, 21, 20-40.
- Coulthard, M. (2014). *An introduction to discourse analysis* (2<sup>nd</sup> ed.). New York, NY: Routledge.
- Coupland, N. (2000). Language, situation and the relational self: Theorizing dialect-style in sociolinguistics. In P. Eckert and J. Rickford (Eds.), *Stylistic variation in language* (pp. 185-210). Cambridge: Cambridge University Press.
- Coupland, N. (2007). *Style: language variation and identity*. Cambridge: Cambridge University Press.
- Crookes, G., & Rulon, K. (1985). *Incorporation of corrective feedback in native speaker/on-native speaker conversation (Technical Report No. 3)*. Retrieved from [http://tesol.aua.am/tqd\\_2000/TQD\\_2000/Tq\\_d2000/Vol\\_22\\_4.pdf#page=132](http://tesol.aua.am/tqd_2000/TQD_2000/Tq_d2000/Vol_22_4.pdf#page=132)
- Cuvalay-Haak, M. (1997). *The verb in literary and colloquial Arabic*. Berlin: Mouton de Gruyter.

- Dahir, J. (1998). *Linguistic variation in Damascus Arabic: A quantitative analysis of men's and women's speech* (Unpublished Ph.D. Dissertation). New York University, New York.
- Dashti, A. (2015). The role and status of the English language in Kuwait. *English Today*, 31 (3), 28 – 33.
- Dashti, A., & Dashti, F. (2015). Is Kuwait TV diglossic? A sociolinguistic investigation. *European Scientific Journal*, 11 (2), 123-137.
- Di Sciullo, A., Muysken, P., & Singh, R. (1986). Government and code-switching, *Journal of Linguistics*, 22 (1), 1-24.
- Ebad, R. (2014). The role and impact of English as a language and a medium of instruction in Saudi higher education institutions: Students-instructors perspective. *Studies in English Language Teaching*, 2 (2), 140-148.
- Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. New York, NY: Teachers College Press.
- Eckert, P. (2000). *Linguistic variation as social practice*. Oxford: Blackwell.
- Eckert, P. (2004). The meaning of style. In E. Chun, L. Mahalingappa, & S. Mehus (Eds.), *Proceedings of the Eleventh Annual Symposium about Language and Society* (pp. 41-53). Austin, University of Texas.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12 (3), 453- 476.
- Eckert, P. (2009). Three waves of variation study: the emergence of meaning in the study of variation [PDF file]. Retrieved from <https://pdfs.semanticscholar.org/89d3/de3d9d756fd39cd0b2b1270c309feb4a49a7.pdf>
- Eckert, P. (2012). Three waves of variation study: the emergence of meaning in the study of variation. *Annual Review of Anthropology*, 41, 87-100.
- Edwards, J. (2008). The transcription of discourse. In D. Schiffrin, D. Tannen, & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 321 - 348). Oxford: Blackwell.
- Erosy, N., Yolal, M., & Batmaz, B. (2015). Cosmetic consumption of metropolitan males in a developing country. *Studia*, 60 (3), 5-22.
- Ferguson, C. A. (1959). Diglossia. *Word*, 15, 325-340.

- Firth, J. (1959). The treatment of language in general linguistics. In F. R. Palmer (Ed.), *Selected papers of J.R. Firth 1952–59* (pp. 206 - 209). Bloomington, IN: Indiana University Press
- Fishman, J. (1967). Bilingualism with and without diglossia: Diglossia with and without bilingualism. *Journal of Social Issues*, 23 (2), 29–38.
- Freberg, K., Graham, K., Mcgaughey, K., & Freberg, L. (2011). Who are the social media influencers? A study of public perceptions of personality. *Relations Review*, 37 (1), 90-92.
- Friginal, E., & Hardy, J. (2014). *Corpus-based sociolinguistics: A student guide*. New York, NY: Routledge.
- Fussell, B. (2011). The local flavour of English in the Gulf. *English Today*, 27 (4), 26-32.
- Gardner-Chloros, P. (2009). *Codeswitching*. Cambridge: Cambridge University Press.
- Giles, H., & Powesland, P. (1975). *Speech styles and social evaluation*. London: Academic Press.
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1-68). New York, NY: Cambridge University Press.
- Görlach, M. (1998). Varieties of English worldwide: Where we stand. *Links & Letters*, 5, 13-36.
- Gumperz, J. J. (1982). *Discourse strategies*. New York, NY: Cambridge University Press.
- Habash, N. (2010). *Introduction to Arabic natural language processing*. San Rafael, CA: Morgan & Claypool Publishers.
- Halim, N. S., & Maros, M. (2014). The functions of codeswitching in Facebook interactions. *Procedia - Social and Behavioural Sciences*, 118, 126-133.
- Hamouda, A. (2015). *Arabic-English code switching in the Egyptian talk show 'Shabab Beek'* (Master's thesis). Arizona State University, Arizona.
- Harrel, R., & Blanc, H. (1960). A linguistic analysis of Egyptian radio Arabic. In C. A. Ferguson (Ed.), *Contributions to Arabic linguistics* (pp. 3-67). Cambridge; MA: Harvard University Press.
- Hary, B. (1996). The importance of the language continuum in Arabic multiglossia. In A. Elgibali (Ed.), *Understanding Arabic* (pp. 69-90). Cairo: The American University in Cairo Press.

- Haugen, E. (1956). *Bilingualism in the Americas: A bibliography and research Guide*. Montgomery, AL: University of Alabama Press.
- Hearn, A., & Schoenhoff, S. (2019). From celebrity to influencer: Tracing the diffusion of celebrity value across the data stream. In: P. David Marshall & S. Redmond, (Eds.), *A Companion to celebrity* (pp. 194–212). Chichester: John Wiley & Sons.
- Heath, J. (1989). *From code switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. London: Kegan Paul International.
- Hernández-Campoy, J. M., & Cutillas-Espinosa, J. A. (2012). Speaker design strategies in political contexts of a dialectal community. In H. C. Manuel & J. A. Cutillas-Espinosa (Eds.), *Style-shifting in public: New perspectives on stylistic variation* (pp. 21-43). Amsterdam: John Benjamins.
- Hewson, C. (2008). Internet-mediated research as an emergent method and its potential role in facilitating mixed methods research. In S. N. Hesse-Biber & P. Leavy (Eds.), *Handbook of emergent methods* (pp. 543-570). New York, NY: Guilford Press.
- Hijab, M. (2016, November 14). Two thirds of the actresses can only fit for being fashionistas. *Alrai*. Retrieved from <https://www.alraimedia.com/Home/Details?id=87ea4760-dde4-4f4e-a833-21a6c812ba29>
- Hilgendorf, S. (2015). Plurality and World Englishes: The social realities of language use. *World Englishes*, 34 (1), 55–67.
- Hilgendorf, S. (2018). Beyond center and periphery: Plurality with inclusivity. *World Englishes*, 37 (3), 472-83.
- Holes, C. (1984). *Colloquial Arabic of the Gulf and Saudi Arabia*. London: Routledge & Kegan Paul.
- Holes, C. (1990). *Gulf Arabic*. London: Routledge.
- Holes, C. (1993). The uses of variation: A study of the political speeches of Gamal Abd al-Nasir. In M. Eid & C. Holes (Eds.), *Perspectives on Arabic linguistics: Papers from the Annual Symposium on Arabic Linguistics* (pp. 13-45 ). Amsterdam: Benjamins.
- Holes, C. (1995). Community, dialect and urbanization in the Arabic-speaking Middle East. *Bulletin of the School of Oriental and African Studies*, 58 (2), 270-287.
- Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Washington, D.C.: Georgetown University Press.

- Holes, C. (2011). Language and identity in the Arabian Gulf. *Journal of Arabian Studies*, 1 (2), 129-145.
- Holmes, J., & Hazen, K. (2014). *Research methods in sociolinguistics: A practical guide*. Hoboken, NJ: John Wiley and Sons Incorporated.
- Holmes, J., & Wilson, N. (2017). *An introduction to sociolinguistics* (5<sup>th</sup> ed.). New York, NY: Routledge.
- Hudson-Edwards, A. (1984) Rediscovering diglossia. *Southwest Journal of Linguistics*, 7, 5 -15.
- Hymes, D. (1974). *Foundations in sociolinguistics: An ethnographic approach*. Philadelphia, PA: University of Pennsylvania Press.
- Irvine, J. (2001). Style as distinctiveness: The culture and ideology of linguistic differentiation. In P. Eckert and J. Rickford. (Eds.), *Stylistic variation in language* (pp. 21-43). Cambridge: Cambridge University Press.
- Ismail, M. (2015). The sociolinguistic dimensions of codeswitching between Arabic and English by Saudis. *International Journal of English Linguistics*, 5 (5), 99-109.
- Jacobson, R. (2001). Language alternation: The third kind of codeswitching mechanism. In R. Jacobson (Ed.), *Codeswitching worldwide II* (pp. 59-72). Berlin: Mouton de Gruyter.
- Johnson, K., & Johnson, H. (1998). *Encyclopedic dictionary of applied linguistics: A handbook for language teaching*. Oxford: Blackwell.
- Johnstone, T. (1967). *Eastern Arabic dialects studies*. London: Oxford University Press.
- Jomah, M. (2018, November 16). Ameerah Mohammed: Ibtisam (Ghosaon fi alwahal) is a different character. *Aljarida*. Retrieved from <http://www.aljarida.com/articles/1542293576055851900/>
- Kachoub, B., & Hilgendorf, S. (2019). Quinlingualism in the Maghreb?: English use in Moroccan outdoor advertising. *English Today*, 35 (3), 1-14.
- Kachru, B. (1966). Indian English: A study in contextualization. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.), *In memory of J. R. Firth* (pp. 255–287). London: Longman.
- Kachru, B. (1985) Standards, codification and sociolinguistic realism: the English language in the outer circle. In R. Quirk & H.G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11-30). Cambridge: Cambridge University Press.
- Katzner, K. (2002). *The languages of the world* (3rd ed.). London: Routledge.

- Kaye, A. S. (1970). Modern standard Arabic and the colloquials. *Lingua*, 24, 374-391.
- Khalifa, S., Habash, N., Abdulrahim, D., & Hassan, S. (2016). A large scale corpus of Gulf Arabic. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & Stelios Piperidis (Eds.), Paper presented at *the Proceedings of the Tenth International Conference on Language Resources and Evaluation Conference* (pp. 4282-89). Portorož, Slovenia. Retrieved from <https://www.aclweb.org/anthology/L16-1679>
- Khan, I., Dongping, H., Abdullah, M., Ahmad, Z., Ahmad Ghauri, T., Ghazanfar, S., & Wright, L. (2017). Men's attitude and motivation toward consumption of grooming products: A comparison of chinese and pakistani male consumers. *Cogent Business & Management*, 4 (1). <https://doi.org/10.1080/23311975.2017.1309783>
- Kortmann, B. (2005). *English linguistics: Essentials*. Berlin: Cornelsen.
- Kniaż, M. (2017). English lexical items in Egyptian Arabic. Codeswitching or borrowing? *Alicante Journal of English Studies*, 30, 185-210.
- Labov, W. (1963). The social motivation of a sound change. *Word*, 19, 273–309.
- Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Lea, W. (1973). Segmental and suprasegmental influences on fundamental frequency contours. In L. M Hyman (Ed.), *Consonant Types and Tones, Southern California Occasional papers in linguistics* (pp. 15–70). Los Angeles: University of Southern California Press.
- Lee, J. S. (2007). I'm the illest fucka. *English Today*, 23 (2), 54-60.
- Leung, C. (2006). *Codeswitching in print advertisements in Hong Kong and Sweden* (Unpublished master's thesis). Lund University, Sweden.
- Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). Oxford: Oxford University Press.
- Lipski, J. M. (2005). Codeswitching or borrowing? No sé so no puedo decir, you know. In L. Sayahi & M. Westmoreland (Eds.), *Selected proceedings of the Second Workshop on Spanish Sociolinguistics* (pp. 1-15). Somerville, MA: Cascadilla Proceedings Project.
- MacSwan, J. (1999). *A minimalist approach to intrasentential code switching*. New York, NY: Garland.

- Mahmoud, A. (2013). A linguistic perspective of the effect of English on MSA: Manifestations and ramifications. *Journal of King Saud University - Languages and Translation*, 25 (1), 35-43.
- Mallinson, C., Childs, B., & Herk, G. V. (2013). *Data collection in sociolinguistics: methods and applications*. New York: Routledge.
- Marçais, W. (1930). La diglossie arabe. *l'Enseignement public*, 97, 401-9.
- Mazraani, N. (1997). *Aspects of language variation in Arabic political speech-making*. Richmond, UK: Curzon.
- Mejdell, G. (2006). *Mixed styles in spoken Arabic in Egypt*. Leiden: Brill.
- Moessner, L. (2001) Genre, text type, style, register: A terminological Maze? *European Journal of English Studies*, 5 (2), 131-138.
- Moungkhem, C., & Surakiatpinyo, J. (2010). *A study of factors affecting on men's skin care products purchasing, particularly in Karlstad, Sweden* (Unpublished master's thesis). Karlstad University, Sweden.
- Mukenge, C., & Chimbarange, A. (2012). A discourse analysis of the use of code switching in the film Yellow Card. *International Journal of Linguistics*, 4 (4), 581-589.
- Myers-Scotton, C. (1989). Codeswitching with English: Types of switching, types of communities. *World Englishes*, 8 (1), 333-346.
- Myers-Scotton, C. (1993a). *Duelling languages*. Oxford: Oxford University Press.
- Myers-Scotton, C. (1993b). *Social motivations for codeswitching: Evidence from Africa*. Oxford: Clarendon Press.
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Nerghes, A. (2011). *The impact of codeswitching on persuasion: An elaboration likelihood perspective* (Master's thesis). Wageningen University, Netherlands.
- Obiamalu, G.O., & Mbagwu, D. U. (2007). Codeswitching: Insights from code-switched English/Igbo expressions. *Awka Journal of Linguistics and Languages*, 3, 51-53.
- Odden, D. (2005) *Introducing phonology*. Cambridge: Cambridge University Press.
- Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y Termino en Español: Toward a typology of codeswitching. *Linguistics*, 18 (7/8), 581-618.

- Poplack, S. (1981). Syntactic structure and social function of codeswitching. In R. P. Durán (Ed.), *Latino language and communicative behaviour* (pp. 169-184). Norwood, NJ: Ablex.
- Qafisheh, H. A. (1975). From Gulf Arabic into Modern Standard Arabic: A pilot study. *Journal of the American Association of Teachers of Arabic*, 8 (1), 36-39.
- Qafisheh, H. A. (1977). *A short reference grammar of Gulf Arabic*. Tucson, AZ: University of Arizona Press.
- Qudah, M. (2018). A sociolinguistic study: Diglossia in social media. *International Conference ICT for Language Learning*. Retrieved from <https://conference.pixel-online.net/ICT4LL/files/ict4ll/ed0010/FP/4139-ICL2733-FP-ICT4LL10.pdf>
- Sadiqi, F. (1995). The language of women in the city of Fes. *International Journal of the Sociology of Language*, 112, 63-80.
- Sadiqi, F. (2003a). *Women, gender, and language in Morocco*. Leiden: Brill.
- Sadiqi, F. (2003b). Women and linguistic space in Morocco. *Women and Language*, 26 (1), 35-43.
- Salam, A. M. (1980). Phonological variation in educated spoken Arabic: A study of the uvular and related plosive types. *Bulletin of the School of Oriental and African Studies*, 43 (1), 77-100.
- Salih, S. M. (2015). Re-evaluating diglossia: The case of Iraqi Arabic. *International Journal of Educational Investigations*, 2 (5), 103-117.
- Saville-Troike, M. (1982). *The ethnography of communication: An introduction*. Oxford: Blackwell.
- Sayahi, L. (2011). Current perspectives on Tunisian sociolinguistics. *International Journal of the Sociology of Language*, 2 (11), 1-8.
- Schilling-Estes, N. (2002). Investigating stylistic variation. In J.K.Chambers, P. Trudgill, and N. Schilling-Estes (Eds.), *The handbook of language, variation, and change* (pp. 375-401). Malden: Blackwell.
- Schmidt, R. (1974). *Sociolinguistic variation in spoken Egyptian Arabic: A re-examination of the concept of diglossia* (Doctoral dissertation). Brown University, Providence.
- Schouten, A., Janssen, L., & Verspaget, M. (2019). Celebrity vs. influencer endorsements in advertising: The role of identification, credibility, and product-endorser fit. *International Journal of Advertising*, 1-24. DOI: 10.1080/02650487.2019.1634898



- Scollon, R. (1974). *One child's language from one to two: The origins of construction* (Doctoral dissertation). University of Hawaii at Manoa, Honolulu, Hawaii.
- Shivtiel, A. (1991). The maze of Arabic. In A. S. Kaye (Ed.), *Semitic studies* (pp. 1435-1442). Wiesbaden, Germany: Harrassowitz.
- Silverstein, M. (1976). Shifters, linguistic categories, and cultural description. In K. H. Basso & H. A. Selby (Eds.), *Meaning in anthropology* (pp. 11- 55). Albuquerque: University of New Mexico Press.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23 (3- 4), 193- 229.
- Smith, L. (2015). English as an international auxiliary language. *Journal of English as a Lingua Franca*, 4(1), 159-164. (Reprinted from *RELC Journal*, 7 (2), (1976), 38–42.)
- Souiden, N., & Diagne, M. (2009). Canadian and French men's consumption of cosmetics: A comparison of their attitudes and motivations. *Journal of Consumer Marketing*, 26 (2), 97-109.
- Stadlbauer, S. (2010). Language ideologies in the Arabic diglossia of Egypt. *Colorado Research in Linguistics*, 22 (4), 1-19.
- Suleiman, S. (1985). *Jordanian Arabic between diglossia and bilingualism: Linguistic analysis*. Amsterdam: John Benjamins.
- Suleiman, Y. (1999). *Language and society in the Middle East and North Africa: Studies in variation and identity*. Richmond, UK: Curzon.
- Tamimi, Y., & I'lawi, N. (2006). Variation in northern Jordanian Arabic: Manifestations of standard origin. *Journal of Language and Linguistics* 5 (2), 230-245.
- Theodoropoulou, I. (2014). *Sociolinguistics of style and social class in contemporary Athens*. Amsterdam; Philadelphia, PA: John Benjamins.
- Versteegh, K. (2014). *The Arabic language* (2<sup>nd</sup> ed.). Edinburgh: Edinburgh University Press.
- Wardhaugh, R., & Fuller, J. (2015). *An introduction to sociolinguistics* (7<sup>th</sup> ed.). Chichester, West Sussex: Blackwell.
- Winford, D. (2002). *An introduction to contact linguistics*. Malden, MA: Blackwell.
- Wolfram, W., & Schilling-Estes, N. (2006). *American English: Dialects and variation*. Malden, MA: Blackwell.

- Wolfson, N. (1986). Research methodology and the question of validity. *TESOL Quarterly*, 20, 689-699.
- Zaidan, O, & Callison-Burch, C. (2014). Arabic dialect identification. *Computational Linguistics*, 40 (1), 171–202.
- Zhang, Q. (2005). A chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society*, 34 (3), 431- 66.
- Zhang, Q. (2008). Rhotacization and the ‘Beijing smooth operator’: The social meaning of a linguistic variable. *Journal of Sociolinguistics*, 12 (2), 201-222.