1   Introduction

This book is intended as a textbook and as a resource for research on variation and change in Arabic. The book was designed with a wide range of readers in mind, including students of Arabic linguistics, scholars in linguistics who may or may not know Arabic, and anyone who wishes to expand their knowledge about sociolinguistic theory and methodology as applied to Arabic data.

In selecting the topics to be covered, we followed two principles: relevance to general sociolinguistic theory and availability of empirical data from Arabic-speaking communities. The material included derives from field research in dialectology and sociolinguistics. These sub-fields of linguistics, as we illustrate throughout the book, are interlinked. They are both grounded in field linguistics and focus on language variation. Dialectology, as the name suggests, is the study of dialects, which began as a formal endeavour in the nineteenth century. Traditional dialectology was concerned with describing dialects and accounting for regional variation across dialects. Works in traditional dialectology were presented in grammars, dictionaries, and linguistic atlases.

1.1   Arabic Dialectology

Around the same time as the first large-scale dialectological projects in Europe were carried out, interest in describing Arabic dialects was on the rise as well, with the first such description (of Egyptian Arabic) appearing in 1880 by Wilhelm Spitta.[1] Subsequently, Gotthelf Bergsträsser (1915) published the first linguistic atlas of Arabic dialects, for Syria and Palestine, followed by Jean Cantineau's (1940) atlas of Horan (see Chapter 8).

Many more descriptions of Arabic dialects followed, as well as (more recently) atlases covering the dialects of certain countries and regions (e.g. Syria, Yemen, Palestine, Morocco) and a four-volume lexical atlas of the whole Arab world (by Behnstedt and Woidich 2010–2021).

---

[1] A detailed history of Arabic dialectology is available in Behnstedt and Woidich (2013).

Traditional dialectology provides the fundamental linguistic background necessary for embarking on sociolinguistic research, while sociolinguistics goes beyond description to refine our understanding of dialectological data, highlighting connections to a general theory of language. Since the earliest days of sociolinguistics, research in the field has relied rather heavily on dialectology, with the link between the two fields gradually increasing over the years. The earliest sociolinguistic studies of English varieties in the United States (Labov 1963) and in the United Kingdom (Trudgill 1974) were anchored in dialectological descriptions and in a sense were conceived of as a natural progression from the traditional studies in dialectology.

1.2   The Variationist Approach to Arabic Sociolinguistics

In research on Arabic, the first sociolinguistic study was Clive Holes's investigation of Bahrain (published in 1987). This study embraces two crucial principles. Firstly, its treatment of variation proceeded on the basis of thorough knowledge and detailed dialectological description of the varieties analysed. Secondly, it approached variation and change in the vernacular as a case of interaction between locally based native varieties. Holes separated the macro-level aspect of variation in Arabic, namely standard versus vernacular, from the domain of variation as it is observed within the vernacular.

These are the principles of sociolinguistic analysis of Arabic that we adopt and advocate because they are in keeping with the foundations of sociolinguistic analysis of *any* natural language, Arabic included. This, unfortunately, is not how Arabic data are always approached. We still find analyses of variation in Arabic based on a preoccupation with diglossia, which assume a priori that change in a given Arabic dialect must be either towards Standard Arabic (in Arabic: *al-fuṣḥā*) features or away from them. Thus, we read about seemingly contradictory patterns within the same dialect, whereby one feature changes from a localised variant to one which is identical to a Standard variant (e.g. [ʧ] > [k]) and another from a localised feature that the dialect shares with the Standard towards a non-Standard variant (e.g. [θ] > [t]). The former is often referred to as 'standardisation' and interpreted as being motivated by the prestige of Standard Arabic. The latter is then referred to as 'colloquialisation' and said to be motivated by the prestigious dialectal norm. What such analyses fail to capture is that changes of the type [ʧ] > [k] and those of the type [θ] > [t] follow a single trajectory, which may be characterised as LOCAL-ISED > SUPRALOCAL. In the example cited above, both target variants, [k] and [t], are the realisations of these two phonemes in the supralocal prestigious dialects of the Levant (e.g. Damascus, Beirut, Jerusalem, Amman). These supralocal dialects happen to share many of their features with the Standard (e.g. [k] for /k/), while other features diverge from it (e.g. [t] for /θ/). The fact

that one of them, [k], also exists in Standard Arabic, is merely a coincidence. It is thus clear that the trajectory of change in localised dialects, such as the rural dialects of Palestine, is straightforwardly towards these supralocal urban variants.

Another misleading outcome of a 'standardisation' approach (i.e. the assumption that variation revolves around approximating to features present in Standard Arabic) is that it distracts from the sociolinguistically relevant points. For example, staying with Palestinian Arabic, the variable (q) has no fewer than four different variants: [ʔ], [k], [kˤ] and [g]. Research shows that the trajectory of change is systematically towards [ʔ], which is characteristic of the socially dominant city dialects. Speakers of Palestinian Arabic (as is the case in many other speech communities) generally use Standard [q] in a closed set of lexical items, which are understood to be borrowings from Standard Arabic (e.g. /taqriːr/ 'report'). Analysing variation in this case in terms of standardisation would only account for cases where the speakers vary between [q], on the one hand, and any of the other variants, on the other. Not only is this a false interpretation of the use of this variable, but crucially it distracts from the sociolinguistically relevant variation. What is relevant and socially meaningful in this community is the alternation between the vernacular variants [ʔ], [k], [kˤ], [g].

Additionally, there is a small number of speakers who belong to a well-defined social group (the Druze), for whom [q] is the vernacular variant of /q/. The Druze, too, in interdialectal contexts, tend to shift away from their traditional realisation, even though it is identical to that of Standard Arabic. And when they do, the target variant is, once again, [ʔ]. Another example of [q] being part of a traditional Palestinian dialect comes from a study by El Salman (2003) among speakers of Palestinian descent from the village of Tirat Haifa currently resident in Jordan. Here, too, the results show that younger speakers diverge from the traditional [q] variant to the variants used in Jordan, namely [g] and [ʔ].

This is not at all to say that Standard Arabic is irrelevant to Arabic-speaking communities in countries where it is the official language. Native speakers of Arabic of all backgrounds and creeds revere the Standard variety in various ways. It represents a pan-Arab norm, although its utility as a lingua franca between speakers of different dialects is often exaggerated in the literature and in popular perception.[2] It undoubtedly functions as a stylistic device and is the norm used in formal written and spoken domains. The use of Standard Arabic in these domains is worthy of investigation in its own right, and such research has its own decades-long tradition (e.g. the foundational works of Badawi 1985; Mitchell 1978; El-Hassan 1977). However, as a non-native variety,

---

[2] On this point see Holes (2004: 5).

Standard Arabic does not play a role, nor does it have a normative effect on the structure of variation in the core areas of the grammar of the vernacular (phonology, morphology, and syntax).

Understanding the structure of variation and trajectory of change in Arabic begins with the micro-attributes of the local dialect and the community in which this dialect is spoken, i.e. it begins with the specific and gradually expands to more generalisable trends. Basic to this bottom-up approach is the assumption that generalisations, if there are any, can best be identified through an aggregate of locally based studies in a given region. Importantly, this approach is driven by empirical data.

1.3   Diglossia and Code-Switching

An important, novel contribution to the incorporation of diglossia into the discourse about language variation and change in Arabic is Lotfi Sayahi's 2014 book, *Diglossia and Language Contact*: *Language Variation and Change in North Africa*. In this book, Sayahi makes two important contributions. Firstly, he takes a fresh look at the classic definitions of diglossia, starting with Karl Krumbacher (1902), through to Jean Psichari (1928), William Marçais (1930), Ferguson (1959), and finally, Penelope Eckert, who in a 1980 article analysed diglossia, based on data from Gascon,[3] but with broad applications for general sociolinguistics. Secondly, he couches his analysis of diglossia in the speech communities of the Maghreb within contemporary sociolinguistic theory. He does so through illustrating how diglossia interplays with other phenomena such as code-switching, borrowing (lexical and structural), bilingualism, and language variation.

Sayahi distinguishes between diglossia and bilingualism in the context of the Maghreb. This distinction is based on a widely accepted premise: diglossia is within Arabic varieties, whereas bilingualism involves Arabic and some other language. This distinction lends itself to insights on code-switching as well.

Stemming from this distinction is Sayahi's (2014) reference to 'diglossic code-switching' and 'bilingual code-switching', both of which can be oral or written. Other scholars, e.g. Lahlou (1991), only consider the latter, i.e. the discursive alternation between Arabic and another language, as code-switching. This follows the classic definition of the code-switcher as being a bilingual. According to this definition, monolinguals, such as speakers of Arabic whose everyday speech may include elements of their vernacular in combination with elements from Standard Arabic (but with no interference of a foreign language) *cannot* be code-switchers (see Lahlou 1991).

---

[3] Gascon is a Romance language spoken in south-west France.

Code-switching in Arabic has been studied in detail particularly in the North African context. Studies such as Bentahila and Davies (from 1983 onwards), Lahlou (1991), and Ziamari (2003, 2007) in Morocco and Sayahi (2011) in Tunisia explore the post-colonial effects of French as a component in everyday speech, alongside Arabic. Another type of code-switching is between Arabic and European languages, e.g. Dutch, in Europe itself, among Arabic-speaking immigrants (e.g. Boumans 1998, Boumans and Caubet 2000).

Bentahila et al. (2013: 327) summarise the work in this field as follows:

[W]e can draw a distinction between two general approaches to codeswitching. On one hand, we have the strictly formalistic models, focusing largely on structure, formulating absolute generalizations, and claiming to identify universal principles; on the other, there are more holistic, interdisciplinary approaches that take a wider view, acknowledging the relevance of many other variables and drawing on insights from fields such as sociolinguistics, psycholinguistics, pragmatics, discourse, and conversation analysis.

We refer the reader to Sayahi's (2014) book for full coverage of the relationship between code-switching and language variation and change.

1.4   The Link to Historical Linguistics

Another important sub-discipline to which sociolinguistics is connected is historical linguistics, as both are concerned with the study and understanding of language change. Historical linguistics has traditionally analysed language change retrospectively, i.e. after the change had been completed, while sociolinguistics has devised methods to document and explain language change as it progresses. It was William Labov who in the 1960s pioneered the study of language change in this manner. In addition to allowing for the analysis of change in progress, through such means as the apparent time construct, Labov incorporated the social context of language use in the community as an essential component of linguistic theory (see Chapters 2, 7).

1.5   Variation and Change

Language variation and language change are inextricably linked, in the sense that a change from form A to form B implies a stage in which both A and B coexisted, i.e. a stage of variation. From this perspective, what we think of as 'historical change', e.g. the transition from Middle English to Modern English, is the cumulative effect of incremental stages of language variation. An important goal of sociolinguistics is to explain the mechanisms and the social forces that propel these changes, alongside purely linguistic factors. In practice, sociolinguists are primarily concerned with changes which are ongoing. It is thus necessary to provide thorough explanations of social variation in order to

understand change. Over the decades, scholars such as Penelope Eckert have further refined the theory to include the analysis of *social meaning*. While this theoretical innovation has continued to focus on language as the main object of study, it added an additional dimension to the interpretation of linguistic variation, arguing that the usage of different linguistic variants mirrors speakers' social positioning, in much the same way as do other performative behaviours, such as clothing and make-up.

Sociolinguistic research is an empirically based scientific discipline, in the sense that it extrapolates generalisations from attested facts. In order to do this, it is necessary to measure patterns of variation according to objective criteria, such as quantitative analysis of the frequency of forms within and across different social groups. It is important to emphasise that this is a means to an end, not an end in itself. The end, in this case, is to identify the sociolinguistic factors that underpin language variation in a community.

1.6   Layout of the Book

The book begins with a chapter on methodology, which provides a concise introduction to the principles and concepts involved in the design and execution of sociolinguistic research on Arabic. In this chapter we also introduce social variation by discussing age, one of the fundamental speaker variables that sociolinguists include in sampling and analysing variation and change.

Chapter 3 is dedicated to gender, another central speaker variable. Research in Arabic sociolinguistics has only recently begun to incorporate an interdisciplinary approach to analysing gender as a social category in a manner comparable to research on other languages. In this chapter we dispel some myths about Arabic and caution against relying on stereotypes and unwarranted generalisations. We advocate an approach to analysing gender in Arabic-speaking communities that considers it a non-discrete social category, whose social meaning is locally constructed.

Chapters 4 and 5 deal with different ways to stratify speech communities, addressing methodological as well as analytical issues. We begin with a discussion of education, a category that has been used widely in the analysis of Arabic data but whose exact relevance to variation and change in Arabic has been poorly understood. We follow with various measures of social stratification and discuss large-scale categorisations such as socio-economic class, as well as smaller-scale groupings such as social networks.

We have dedicated a separate chapter to religion and ethnicity as social variables because we believe they may play a more significant role than previously recognised, especially in an ever-changing political climate. Chapter 6 presents a fresh approach to these two social factors, reviewing classic and more recent studies in Arabic sociolinguistics where one or both of these factors were found

to play a role in structuring variation. We connect findings from micro- and macro-level studies, e.g. phonological change coupled with language shift. We present analyses and re-analyses of Arabic data that go beyond the linguistic differences themselves and focus on the underlying causes and factors that foster sociolinguistic variation along religious and/or ethnic lines.

Chapter 7 deals with language change. It lays the theoretical ground for the study of this process, which is central to sociolinguistics. These theoretical aspects are illustrated with an array of empirically tested investigations of several Arabic-speaking communities. We emphasise and elaborate on the connections between historical linguistics and sociolinguistic approaches to the study of language change.

Chapters 8 and 9 cover topics in variation and diffusion of linguistic features across space. We provide extensive examples of traditional dialect geography by reviewing the seminal works of the early Arabic dialectologists. We proceed by introducing recent studies, highlighting the continuity and innovations in the description and cartographic representation of regional variation and the theoretical insights that emanate from them. In Chapter 9 we focus specifically on language contact (and dialect contact) as a necessary precursor to diffusion across space.

We have made every effort to render this book accessible to a wide readership while maintaining a high standard of academic rigour, factual accuracy, and sound intellectual reasoning.

1.7   Further Reading

The following list of recommended reading includes articles and chapters that introduce Arabic sociolinguistics and critically survey research in the field. These resources approach the subject matter from a perspective similar to the approach adopted in this book. We list them here in chronological order to reflect the evolution of the field.

> Holes (1995) – A discussion of koineisation in the dialects of three Arab cities.
>
> Haeri (2000) – On the nexus between Arabic sociolinguistics and linguistic anthropology.
>
> Owens (2001) – A critical review of research in Arabic sociolinguistics to date.
>
> Miller (2007) – An overview of linguistic developments in different types of urban centres.
>
> Owens (2011) – An introduction of Arabic sociolinguistics for scholars in Semitic linguistics.
>
> Owens (2013) – A historico-philosophical perspective on Arabic.

8   Introduction

Al-Wer (2013) – A critical review of research in Arabic variationist sociolinguistics focussing on methods and analytical frameworks.

Horesh and Cotter (2015) – A survey of works to date focussing on the sociolinguistics of Palestinian Arabic.

Horesh and Cotter (2016) – A critical survey of variationist research in Arabic, intended for a broad audience of linguists.

Al-Wer and De Jong (2018) – A concise introduction to macro-sociolinguistic aspects of Arabic and the geographical classification of its varieties.

Holes (2018) – An extensive, summative account of historical Arabic dialectology.

Al-Essa (2019) – A critical synthesis of works on phonological and morphological variation in Arabic dialects.

Al-Wer and Horesh (2019) – An epistemology of Arabic sociolinguistics.

Haeri and Cotter (2019) – An update and addendum to Haeri (2000).

Herin (2019) – A coherent illustration of the concept of traditional dialects as applied to Arabic.

Horesh (2021) – An up-to-date evaluation of variationist sociolinguistics contextualised within Arabic linguistics more broadly.

Al-Wer et al. (2022) – A synthesis of recent findings from several Arabic vernaculars and their implications for a general theory of language change.