The Bedouin-Sedentary Dichotomy in Najd:

A Sociolinguistic Study

Nasser Mohammed Alajmi

PhD

University of York

Language and Linguistic Science

June 2019

Abstract

This study is a sociolinguistic investigation of the dialect contact between the Bedouins (Bs) and Sedentary groups (Ss) in central Najd in Saudi Arabia. Three linguistic variables have been examined quantitatively. The variables are: the merger /u/, /i/ > /i/ (B /u/ as in *kull* and S /i/ as in *kill* 'all'), the 3SM suffixed pronoun (as in B *kita:b-ih* and S *kita:b-ah* 'his book') and the imperfect prefixes (as in B *ya-ktib* and S *yi-ktib* 'he writes'). The data is based on sociolinguistic interviews of 23 Bs, grouped by age and tribe, and 13 Ss, grouped by age. Additionally, the relative sociolinguistic salience of the variables has been measured using the *Social Category Association Test* (SCAT), online identification task and multiple interviews.

The interview data has been statistically analysed using *generalised linear mixed model* in *R*. The data show that Bs are converging towards the S dialect, to varying degrees across the examined linguistic variables. The Bs' convergence on the S variant of the merger (/i/), exhibiting the highest rate, is conditioned by the factors of tribe, education and level of contact. The Dawasir tribe were more likely to converge on /i/ than the other tribes, while speakers with low levels of education and contact with the Ss were less likely to converge on /i/. The Bs' convergence on the S variant of the 3SM suffixed pronoun, which is quite low, is conditioned by age. Younger Bs were more likely to converge on *-ah*. As for imperfect prefixes, the Bs did not show any convergence. The interview data for the Ss, on the other hand, show consistent use of their variants, except for 4 speakers who code-switched to Standard Arabic. The data from the SCAT, online identification task and multiple interviews reveal that the merger is salient (it is a marker), while the other variables are not (indicators).

The quantitative data and qualitative observations of other linguistic variables indicate that a Najdi koine has not formed; however, it is expected to be complete after two or three generations and will most likely exhibit S and Standard Arabic features, but not B.

To my wife Maitha,

my daughters Lina and Lama

&

my parents

Table of Contents

List of Tables

12

List of Figures

Acknowledgments

I thank Allah Almighty for granting me the inspiration, time, patience and health to finish this work.

I would like to express my sincere gratitude to my supervisor Prof. Paul Kerswill for his continuous support, patience, motivation, and immense knowledge. Also, he was more of a friend than a supervisor.

I would also like to thank my wife who supported me throughout my post-graduate studies. Thanks to my mother and father for always supporting me with prayers and love. Thanks to all my brothers and sisters for helping me whenever I needed. Special thanks go to my eldest brother (Saad) who has always been my safety net.

Finally, I would like to thank my cousin Mohammed for helping me in recruiting speakers. My deepest gratitude also goes to Abu-Nayif, who has helped me recruit a large number of speakers.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Phonetic Transcription and Transliteration Symbols

Arabic transliteration in this thesis is mostly in IPA. However, some symbols are substituted for the convenience of the reader. Below I provide these alternative symbols and their IPA equivalents.

Alternative	IPA
/i/	/ɪ/
/u/	/ʊ/
/y/	/j/
/x/	/χ/

Note: in all Arabic examples, the open front vowel /a/ and the open back vowel [ɑ] are both written as /*a*/ because it is the italicised version of /a/, and not because we want to indicate that all vowels in the examples are open back. The open back vowel [ɑ] in Arabic is an allophone of /a/. This allophonic variation is irrelevant to the present study. Therefore, the non-italicised version of /a/ will be written as /a/, while the italicised versions will be written as /*a*/.

List of Abbreviations

1	First person
2	Second person
3	Third person
M	Masculine
F	Feminine
C	Common gender
S	Singular
P	Plural
D	Dual
IND	Indefinite
B	(noun, singular) a Bedouin person
B	(adj.) Bedouin
Bs	(noun, plural) Bedouins
S	(noun, singular) a Sedentary person
S	(adj.) Sedentary
Ss	(noun, plural) Sedentary individuals
WB	(adj.) Western Bedouin
WBs	(noun, plural) Western Bedouins

Introduction

The population of central Najd (present-day Riyadh Province) is mainly comprised of Bedouins (Bs) and Sedentary groups (Ss), all of whom descend from Arab tribes that have inhabited the Arabian Peninsula since the pre-Islamic period and follow the same religion/sect (Islam/ Sunni). The Bs were nomads who used to roam the desert, while the Ss were farmers who settled in small villages scattered across the area. Thus, the major difference between these two groups was the lifestyle they led, not ethnicity or religion. Although members of each group did not essentially live in proximity to one another, it was found that each group spoke a rather similar dialect (Doughty, 1924; Ingham, 1994).

The classification of Arabic dialects into B and S is not novel but has been employed by early Arab grammarians, e.g., Ibn Jinni, as early as 1002 AC (Versteegh, 2001, p. 184). This dichotomy is found in all parts of the Arab world. Palva (2006, p. 606) lists some of the linguistic features shared by all or most dialects of each group. The features that have received most attention, some of which are found in regions where there is rural/urban sociolinguistic variation, are the realisations of /q/ (B or rural /g/ and urban S /ʔ/) and B retention of the interdental fricatives /θ/, /ð/ and /ðˤ/ (urban S /t/, /d/ and /dˤ/ respectively). In the case of Najd, however, both dialect groups (B and S) use /g/ and have retained all interdental fricatives. In fact, based on the linguistic criteria used to distinguish B from S dialects in most parts of the Arab world, the Najdi S dialect can be characterised as B. Thus, the linguistic B/S dichotomy in Najd is subtle and unique, compared to other regions. The linguistic differences between B and S dialects in Najd are reflected in the B retention of the opposition between the short vowels /u/ and /i/ (the merger), the realisation of the 3SM suffixed pronoun (B *-ih* and S *-ah*) and the vowel of the imperfect prefixes (B /a/ as in *ya-* and S /i/ as in *yi-*).

The massive urbanisation, associated with the oil boom in the 1970s, has led the two groups to abandon their old lifestyle and move to the city (see section 1.4). Some pursued education, even after high school, while others occupied jobs that matched their skills. Despite the social gap and differences in lifestyle between the Bs and Ss, homogeneity has played a critical role in the social integration between the two groups. While some Bs choose to isolate themselves and cluster in particular neighbourhoods, others choose to blend with the Ss and even marry into their group. It should be noted that in Najd nowadays, the categorisation of a tribe or group as B, for example, does not mean that they pursue a nomadic lifestyle, but it means they are affiliated with a tribe that used to lead a B lifestyle (see section 1.3.4).

Now that we have established that the Bs and Ss speak different dialects and that they have been in contact with one another for more than four decades, we raise the following research questions:

- What is the linguistic outcome of the massive urbanisation in Najd?
- Is there any dialect accommodation? If yes who is converging to who?
- What are the independent social variables governing this convergence?

The socio-political and/or socio-demographic set-up of each region determines the linguistic outcome of urbanisation and the direction of dialect convergence, if there is any. In North Africa, the massive migration of Bs to the urban centres caused a major shift towards the migrants' dialect (Miller, 2004). In the East, the major shift from Christian dialect towards Muslim dialect in Iraq (Abu-Haidar, 1991) and Shia towards Sunni dialect in Bahrain (Holes, 1986) is mainly driven by political reasons. On the other hand, in old urban centres, like Cairo and Damascus, where the rate of urbanisation was slower, the rural internal migration has resulted in dialect levelling, but the urban dialects are maintained. The unique socio-political set up of Najd makes it difficult to predict the linguistic outcome of dialect contact between

the Bs and Ss. There are no religious or political factors involved. All groups are Sunni Muslims, and no group is considered politically dominant. Demographically, no group is considered a minority. In term of social rank, each group is proud of its heritage, and both lifestyles are equally prestigious.

Due to the scarcity of sociolinguistic studies on Najd, we need to establish whether there is dialect convergence and, if there is any, its direction. Furthermore, the linguistic differences between the dialects under study have only been marginally mentioned in the literature.[1] Therefore, before we examine the dialect contact between the B and S dialects, we also need to establish the linguistic variables. That is, we need to collect data from both groups before we attempt to detect any pattern of dialect accommodation/shift.

As a native Najdi speaker who has been observing both Bs and Ss for over 2 years, I have noted that Bs tend to accommodate to the Ss, and not vice versa. Therefore, the sample for the main study includes more Bs than Ss. The Bs can be referred to as the main sample while the Ss function as a kind of control group. We need a greater number of Bs in order to have sufficient data for statistical analysis of the convergence patterns.

The B sample was mainly grouped by tribe and age, each of which is treated as a social factor. Additionally, the factors of education and level of contact (with the Ss) were considered. The sample of the Ss, on the other hand, was grouped by age, and the factor of style was added later. The reasons why Bs and Ss were not grouped by the same factors will be discussed in Chapter 3.

---

[1] Ingham (1994) included regional dialect differences in an appendix to his work on Najdi Arabic, only a few of which are related to the differences between the B and S dialects of central Najd. The majority of these differences involved the regional variation between Northern, Mixed-central and Central Najdi dialects (see section 1.5.4).

The data from each group is treated separately, in an attempt to discover any patterns of dialect accommodation. If the Bs, for example, show convergence in one variable, we investigate which social factor(s) might be explanatory for this variation.

At a later stage in this research, it was noted that the sociolinguistic salience of the variables might have played a role in the variation. In other words, one group showed variation in one variable more than the others because it is salient. The salience of the variables was measured using the *Social Category Association Test* (SCAT)[2], an online identification task and multiple interviews.

The thesis is structured as follows:

In Chapter 1, the first part will include a review of the history and demography (population and lifestyles) of Najd. In the second part, the linguistic situation in Saudi Arabia, including the dialects of Najd, will be reviewed; and the linguistic variables under study will be introduced.

In Chapter 2, we will review: the theory of the *Arabic Koine* by Ferguson (1959a), diglossia in the Arab world, social variables in Arabic sociolinguistics, the sociolinguistic situation in Saudi Arabia, and the original contribution of the current study.

In Chapter 3, the method of data collection (the sociolinguistic interview), social factors (age, tribe, education and contact), sampling and data analysis will be discussed.

Chapters 4, 5 and 6 are each dedicated to a linguistic variable. In each chapter, the linguistic history of the variable is first presented. Second, a review of the forms of the variable in modern Arabic dialects is provided. Third, the variable in Najdi dialects and the envelope of

---

[2] Llamas et al. (2016).

variation are discussed. Lastly, descriptive statistics and the statistical analysis for the variable are given.

In Chapter 7, the first part will shed some light on the role of salience in dialect contact, different models of salience and the two approaches to measuring salience (*criteria-list* and *experimental*). The second part deals with salience in the present study: methodology, data collection and results.

In Chapter 8, we will discuss the sociolinguistic relevance of the quantitative results and provide qualitative observations on the other linguistic variables (1.5.5.4). Also, the formation of the Najdi koine and the sociolinguistics of diglossia in Najd will be discussed in depth.

Chapter One: Introduction and Background

This thesis investigates the linguistic outcome of dialect contact between the Bedouins (henceforth, Bs) and Sedentary groups (henceforth, Ss) of Najd. Each group led rather different lifestyles before the 1970s, and the contact between the groups was limited. Since the oil boom and massive urbanisation in Saudi Arabia, the Bs and Ss have given up their old lifestyles and moved to the city seeking jobs that would provide them with a stable income and therefore a more secure life. For more than four decades, Bs and Ss have lived side by side in the city; however, little is known about the linguistic outcome of this contact.

In the first part of this chapter, the history and demography of Najd, including the population and lifestyles, will be discussed. In the second part, the linguistic situation in Saudi Arabia, including the dialects of Saudi Arabia and Najd, will be reviewed.

1.1 A Brief Political History of Najd, Saudi Arabia

A decade after the fall of the Second Saudi States (1818 - 1891), King Abdul Aziz reclaimed Riyadh, currently the capital city of Saudi Arabia, from Ibin Rasheed in 1902. Later, he gradually started taking over other parts of the central region of the Arabian Peninsula, Najd, and declared himself Sultan of the *Najd Sultanate* in 1921. Six years later, he took over Hijaz, the west coast of the Arabian Peninsula where the two holy cities (Makkah and Madinah) are located, from the Ottoman Empire and declared himself king of *the Najd and Hijaz kingdom*. In 1932, King Abdul Aziz officially declared the country as the Kingdom of Saudi Arabia (Al-Uthaymin, 1999).

1.2 Location and History of Najd

'Najd' in Arabic literally means high-level land, and as the name suggests it is a relatively large plateau in the middle of the Arabian Peninsula. It used to refer to the central area stretching north to Iraq and Jordan, west to Hijaz, east to Al-Ahsa, and south to Najran on the borders with Yemen (Figure. 1.1). The geographical borders between Najd and the surrounding territories were not clear, but rather there were uninhabited desert tracts that minimised contact between Najd and the other territories (Ingham, 1994, p. 1). Later in 1992, King Fahad, son of King Abdul Aziz and the fourth king of Saudi Arabia, divided the kingdom into 13 administrative provinces (Figure 1.2), none of which was named Najd (Al-Uthaymin, 1999). The province that includes most of Najd area is now called *Riyadh*, named after the capital city. Although two other provinces encompass some parts of Najd, people nowadays perceive the province of Riyadh as Najd. In fact, some people use the terms 'Riyadh' and 'Najd' interchangeably; but it should be noted that the term 'Najd' is now used less frequently.

Figure 1.1: *Map of the Arabian Peninsula showing Najd and the surrounding territories.*

Figure 1.2: *Map of Saudi Arabia administrative regions.*

1.3 The Demography of Najd

Before the oil boom and urbanisation, the inhabitants of Najd were always grouped by the lifestyle they led, B or S. Bs were the tribes that constantly roamed the deserts while Ss were the tribes that settled in small villages scattered across the region (Oppenheim et al., 2004). There is no record of any groups living in Najd other than the Bs and Ss. In fact, while censuses of Najd in the past thirty years have grouped the population by citizenship, gender, age, education, marital status, etc., older estimates of the population of Najd (e.g., Hamza, 1931) grouped the population according to lifestyle (B and S). Hamza (1931) estimated that Najd's population was comprised of 1,300,000 Bs and 800,000 Ss (cited in Ashwan, 1990). The reason why Hamza adopted this division is that the population of Najd is homogeneous (all groups are descendants of Arabic tribes that inhabited the Arabian Peninsula since the pre-Islamic period) and follows the same religion and sect (Sunni Muslims). Another reason is that although members of each group do not necessarily live in proximity to one another,

they tend to share the same traditions, habits and more surprisingly a dialect (Doughty, 1924).[3]

Due to scarcity of studies on this topic, much of the commentary in the following sub-sections below is based on Al-Sowayan (2010) and the interviews of the elder Najdi speakers with whom *lifestyle in the past* was a preferred topic of discussion.

1.3.1 The lifestyles and cultural dichotomy between Bs and Ss

It should be first noted that the S lifestyle is not an evolution of the B. In other words, Ss were not originally Bs who progressed to modernity before others; but rather the two ways of life co-existed for centuries. The view that Bs' lifestyle is less sophisticated than that of the Ss' has been suggested by some scholars, such as Ibn Khaldoon, who based their views of the S civilization on those of the Fertile Crescent; however, the Bs and the Ss of Najd are equally developed (Al-Sowayan, 2010). The difference between them is just that each group has adopted the lifestyle they think best harnesses the scarce resources of the dry desert.

The differences between the two lifestyles resulted in an abundance of some resources and shortage of others, which resulted in the need for mutualism (i.e., trading goods). Bs would visit S villages annually, or more often, to trade wool and dairy products, or even sheep, for dates and other fresh produce. Other than these trading encounters, friendly visits between the two groups were not frequent due to a deep-rooted social gap.

Each group admires traits of the other group and disapproves of others. Bs in the past have always thought less of the S lifestyle. As documented in the form of poetry, Bs have disapproved of the fact that Ss drink cow milk (instead of milk from camels), eat chicken

---

[3] That is not to say there are no cultural and linguistic differences within B tribes or S groups.

(instead of hunted meat) and do not raid other tribes (Al-Sowayan, 2010)[4]. Nowadays most Bs appreciate how civilised and educated the Ss appear to be compared to some Bs. Ss, on the other hand, have always viewed the Bs as less civilised and rude. They have also criticised the extinct raiding tradition of Bs and described it as theft and betrayal (Al-Sowayan, 2010).

Because each group disapproves of the other's lifestyle, a group is less likely to change its lifestyle as it is more than just an adaptation to the environment. A group, however, may be forced to change lifestyle due to ecological reasons. For example, on rare occasions in the past, severe droughts caused wells to dry up and forced settled farmers to abandon their villages and adopt a semi-nomadic lifestyle in order to survive. When the drought is over, they either go back to their farms or decide to build a new settlement. On the other hand, a B tribe or family might lose their livestock due to an epidemic, such as scabies, and be forced to adopt a semi-sedentary lifestyle in which they settle at the periphery of a village and work there in exchange for food (Al-Sowayan, 2010).

As stated earlier, each group has its own traditions and values. In the following two subsections, each group will be discussed in more depth.

1.3.2 The lifestyle of the Bs

The word *Badu* (Arabic word for Bedouins) is derived from the root *bada* which translates as 'to start'. From the same root, the words *bida:'y* 'primitive' and *ba:diah* 'the place where Bs live' are derived. The B lifestyle has existed since the pre-Islamic period and has been always characterised as simple and primitive. For over a millennium, many civilisations have flourished and fallen apart in the Arabian Peninsula, and yet the Bs have preserved their

---

[4] Raiding other tribes and taking their livestock is one of the extinct traditions of the Bs of the Arabian Peninsula. Usually, a group of the tribe warriors raid another tribe at night and take their livestock which is usually a few kilometers away from the tents. If they succeed, a counter raid by the victim tribe to retrieve their livestock is commenced.

distinctive way of life. In addition, they share and show pride in the same principles and values, some of which are honesty, hospitality, bravery and chivalry

The Bs of Najd lived in tents made from wool. Each family had two tents, one for men and one for women, but sometimes they own one tent which they split using a cloth barrier. Each family had a herd of camels, sheep or both. As for their diet, they lived on dates, milk and whole wheat bread. Arabic coffee has a special place in their hearts, as they have their own rituals of brewing and serving it to their guests.

The Bs are not always moving from one place to another, but their stays are conditioned by the availability of water and pasture. Since Najd is one of the driest areas in the Arabian Peninsula, Bs needed to move frequently. They moved and settled as a tribe or a sub-group of a tribe, because small groups are vulnerable to raids from other B tribes.

1.3.3 The lifestyle of the Ss

The word *ħadˁar* 'the Sedentary people' is derived from the root *ħdˁr* which means 'civil'. From the same root, *taħadˁdˁar* 'become civilised or lives in the city' and *ħa:dˁirah* 'the place where the Ss live' are derived. The term *ħadˁar* 'the sedentary people' emerged to distinguish their lifestyle from that of the Bs. The term broadly covers any group of people who do not lead a nomadic way of life. In our study, however, the *ħadˁar* are the groups of people who lived in villages across the Arabian Peninsula, more specifically Najd, and whose lifestyle can be characterised as rural, rather than urban.

Most of the villages in Najd existed before Saudi Arabia was founded. They were comprised of mud houses and farms, all run by the local families, not foreign workers. Although the infrastructure of these villages was simple, it gave credence to the idea that the life of Ss was more sophisticated than that of the Bs.

As farmers, the diet of the Ss was more diverse than that of the Bs'. It consisted of vegetables, dates, whole wheat bread, grain, cow dairy products, and chicken. Cows are valued by the Ss and almost every household has one cow or more.[5] The Ss valued honesty, hospitality and bravery as much as the Bs. However, they did not hunt or participate in raiding other tribes or villages (Al-Sowayan, 2010).

1.3.4 Tribal affiliation

The population of Najd, both B and S, can be further subdivided into tribes. Tribal affiliation has been and still is one of the strongest social bonds in the Arab world and more noticeably among Bs. Some would even risk their lives for others just because they are simply from the same tribe (Kressel, 1996). The B tribes that inhabited Najd include, but are not limited to, Qahtan, Ajman and Dawasir, originally from the southern parts of the Arabian Peninsula, and Itban, Subai' and Suhul from the mid-west. As for Ss, there is quite a large number of sub-tribes, but the main four are Tameem, Fadhil, Ayid and Banu Zaid (Oppenheim et al., 2004).
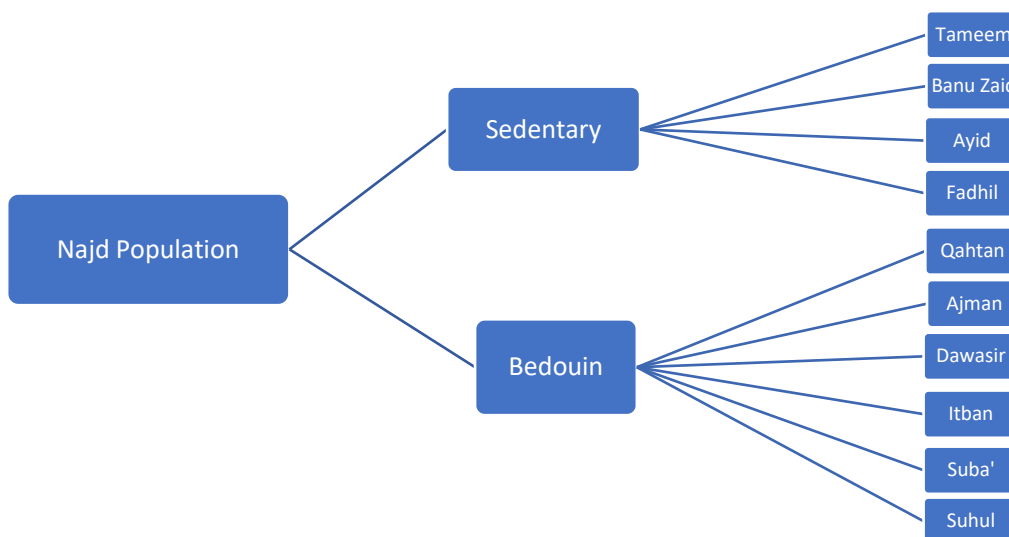
Figure 1.3: *Tribal affiliation of Najdi population. (Based on Oppenheim et al., 2004).*

---

[5] Cows are one of the symbols that set the Bs aside from the Ss. Bs do not own or breed cows because they do not tolerate a harsh climate like sheep or camels. Bs in the past did not consume meat or dairy products of cows because they looked at them as a symbol of Sedentarism. In modern Saudi Arabia nowadays, the majority of dairy products in the market are from cows which most people, Bs or Ss consume. As for beef, however, some Bs up to this day refuse to consume it because of the same reasons as their ancestors.

The origin of a tribe, particularly the B ones, is essential to this study. Although Bs nowadays are scattered across cities of Najd, and needless to say are not pursuing a nomadic lifestyle anymore, they categorise themselves by referring to tribal affiliation rather than their current place of residence (Miller, 2007). The effect of dialect contact (in our case with S dialect) is inevitable in most cases, but in the case of B tribes in Najd, as well as in Saudi Arabia, such effect is resisted by tribal identity. Not only does tribal structure help preserve the traditions and heritage of each social group, but it helps preserve the archaic features of their dialect as well (Bassiouney, 2009, p.113).

The B tribes that are investigated in this study are those which are originally from the southern parts of the Arabian Peninsula. The other B tribes, originally from the western regions of the Saudi Arabia (henceforth, WB), are included for the sake of comparison. Although both groups are B, and share most linguistic features of Najdi B dialect, they are set apart because WB tribes share one linguistic feature with the Ss, as we shall see in chapter 4.

The last issue we will discuss here is tribal genealogy and its linguistic relevance. For two of the examined linguistic variables (Chapter 4 and 6), we need to refer to the genealogy of tribes in order to account for why particular tribes share certain linguistic features. We will not attempt to discuss these linguistic features here, as they are examined in the respective chapters. Instead, we will present the tribal genealogy of the examined tribes, both B and S (based on Al-Hugail, 1967). The B tribes that are originally from the southern parts of the Arabian Peninsula are originally from Yemen. Ajman and Qahtan are from an ancient tribe called *Hamadan*, while Dawasir are descendants of *Azd*. The WB tribes, on the other hand, are descendants of *Hawazin*, an ancient tribe that inhabit the mid-western parts of the Arabian Peninsula (see Figure 1. 4 below). As for the Ss, they are mainly descendants of the well-known tribe of *Tamim* and few others.
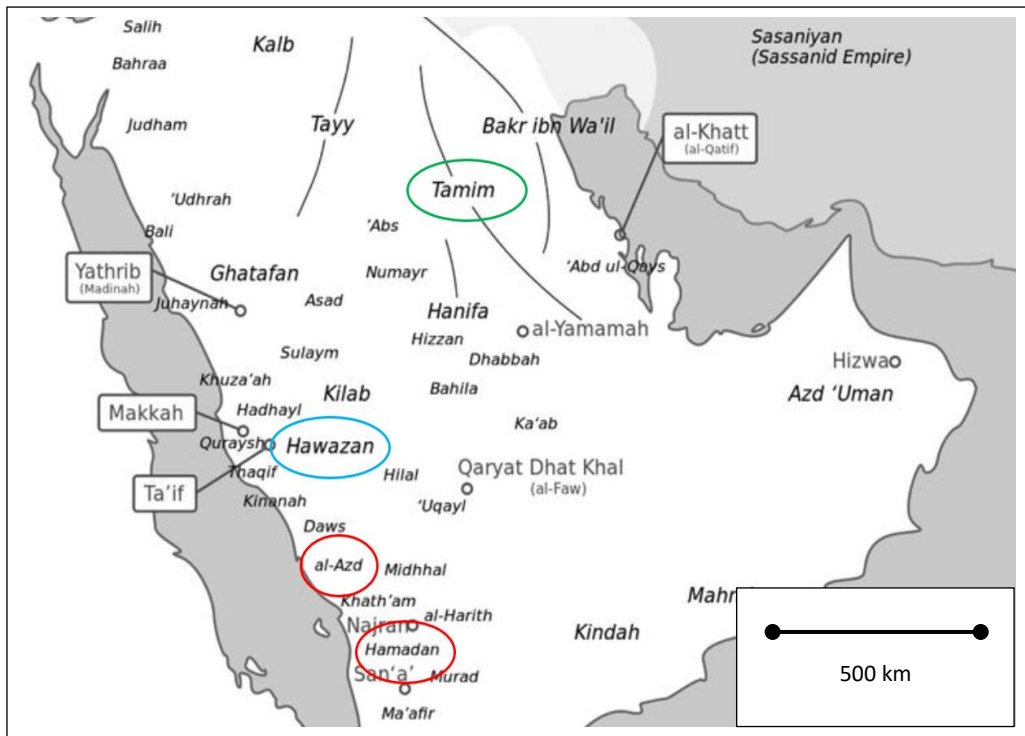
Figure 1.4: *The distribution of the Arab tribes in pre-Islamic times (adapted from Al-Sharkawi, 2017, p. 2).*

1.4 The Oil Boom and Massive Urbanisation

The rapid rate of urbanisation in Saudi Arabia is associated with the oil boom in the 1970s (Figure 1.5). The oil revenues have boosted the GDP of Saudi Arabia, some being channelled to the building of the infrastructure of the country and providing free healthcare and education. The government provided subsidies for wheat, gas and other products, and the income of individuals was tax free. The promise of this prosperous lifestyle gave the Ss and the Bs no choice but to give up their old way of life. The government offered various types of jobs, with priority being given according to citizenship, education and experience, but not being either B or S.

The generation born in the 1950s, whether B or S, has experienced life before and after the oil boom. The Ss of this generation have had basic education in the villages, also known as *kuttab*, in which they learned how to read and write and memorised religious texts. When they moved to the city, they were able to go to schools and pursue their education. The Bs, on the

other hand, were illiterate and when they moved to the city they worked as government drivers or joined the military. Most of the them learned to read and write later while a few were able to pursue their primary education in adult schools. The second generation, born in the 1970s-1980s, had full access to free basic education and college. According to the Saudi Arabian General Authority for Statistics (2016),[6] more than 70% of this age group have completed high school, and more than 60% have obtained college degrees.

Although the majority of the population of Saudi Arabia is now urban (figures 1.5 and 1.6), some members of each group refuse to abandon their heritage and traditions. Quite a large number of Bs still own camels or sheep, mostly for leisure and sometimes for commercial purposes. Some others whose grandparents were farmers (i.e. Ss) also own and run small farms for the same purpose. Due to inflation and other economic reasons, owning livestock or a farm nowadays costs more than it yields, which indicates that it is more a question of identity than just money for Bs and Ss.
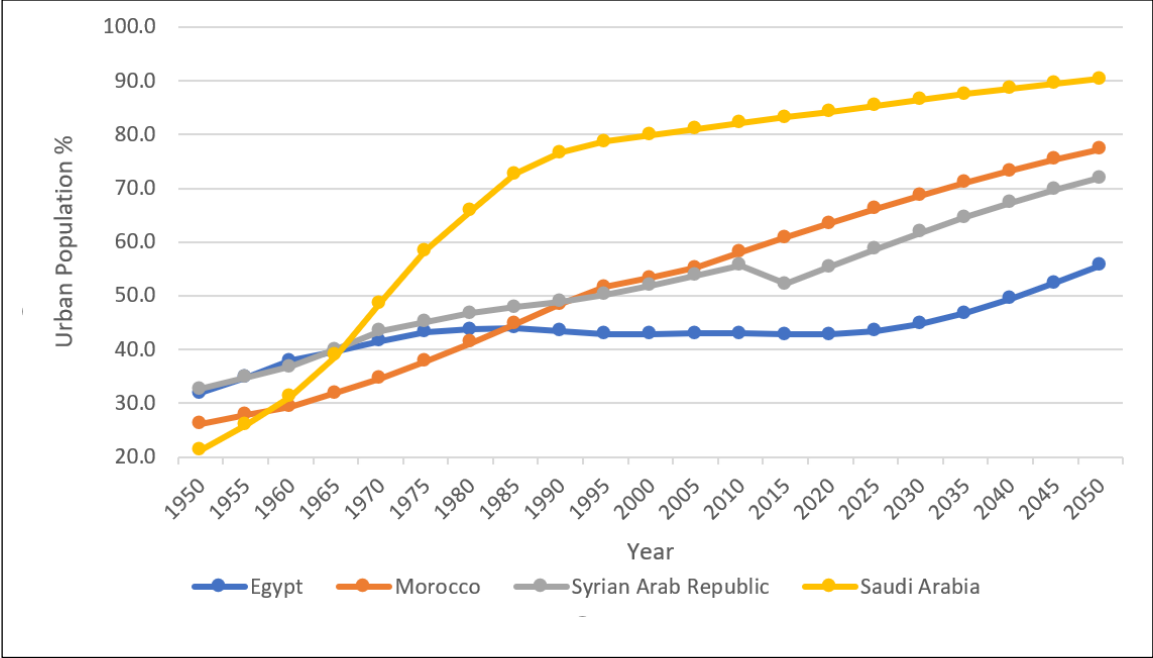
Figure 1.5: *Percentage of urban population in Saudi Arabia, Egypt, Syria and Morocco. Source: United Nations (2018).*

---

[6] In the report titled: *Detailed results of Riyadh Province (general population and housing census)*, Table 10-1.
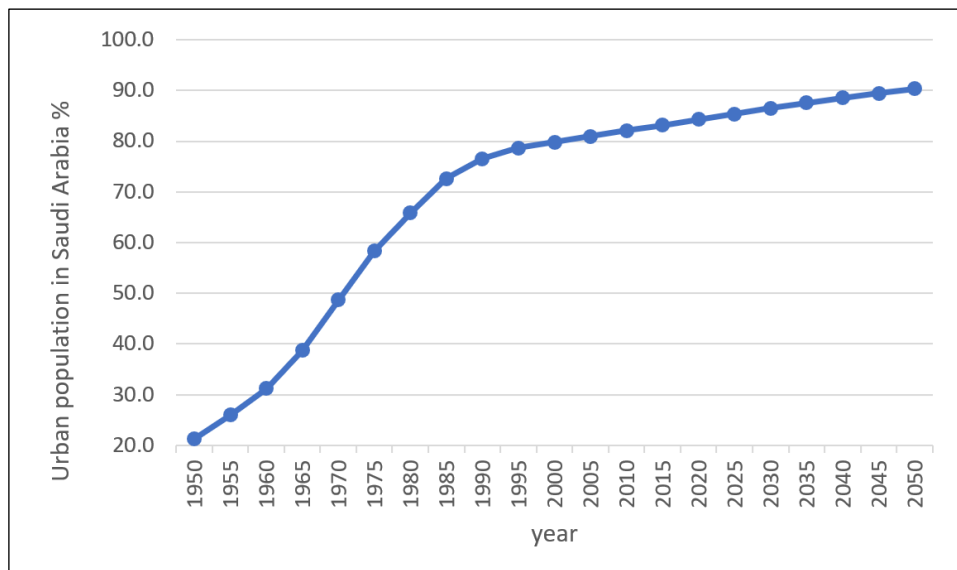
Figure 1.6: *Percentage of urban population in Saudi Arabia. Source: United Nations (2018).*

It should be noted that the urbanisation of the Bs started a little before the oil boom. After King Abdul Aziz unified the kingdom of Saudi Arabia in 1932, he planned on sedentarising the Bs in order to facilitate his role as ruler. He built small settlements, referred to as *hijar* (the plural of *hijrah*, literally translates to 'the place of migrants'), and provided fertile land and farming tools. The Bs resisted at the beginning, but they eventually settled in these villages. While a few of the settled Bs adopted farming as a new lifestyle, the majority continued to pursue a semi-nomadic lifestyle in which they still had livestock but did not move around as much. To this day, these *hijar* are scattered across Saudi Arabia and are mostly inhabited by one B tribe that first settled there.

1.5 The Linguistic Situation in Saudi Arabia

The Arabic language is spoken by more than 200 million people in 22 countries. Modern Standard Arabic, which has a written form, is the official language used in media, education, religion, official documents, newspapers, etc. The spoken form of the language is often known as 'Colloquial Arabic' (Versteegh, 2001), with varying degrees of mutual

intelligibility across the Arab world. Each spoken variety in a country is loosely labelled as a separate Arabic variety, such as 'Egyptian Arabic', but a broader and more systematic classification can be suggested based on shared linguistic features. Versteegh (2001) classifies Arabic varieties geographically into Maghreb, Egyptian, Syro-Lebanese, Mesopotamian and the Arabic variety of the Arabian Peninsula.

1.5.1 Arabic varieties of the Arabian Peninsula

There are many views on the origin of Classical Arabic, but most of them agree that it emerged from the Arabian Peninsula (Rabin, 1951). Based on data collected by Arabic philologists, Rabin (1951) recognises ancient dialects of the Arabian Peninsula as the dialects of Yemen, Himyar, Northern Yemen, Hudhail, Hijaz and Tayyi'. Tracking the evolution and/or extinction of these dialects is marginal to this study. As mentioned in 1.3.4, the ancient dialects are mentioned in this research in relation to the historical background of the current dialects and account for why some linguistic features are present in the dialect of a group of tribes and absent from others.

The current dialects of the Arabian Peninsula, including Yemen, Oman, United Arab Emirates, Qatar, Bahrain, Kuwait and Saudi Arabia, have been broadly classified by Johnstone (1967a) into South-western Arabian, Omani, North Arabian and Hijazi. South-western Arabic includes the dialects spoken in Yemen and the south-western regions of Saudi Arabia; Omani Arabic covers Oman; and Hijazi Arabic is spoken on the west coast of Saudi Arabia. The Northern Arabian variety, relevant to this study, is taken to include the central and northern regions of the Arabian Peninsula and the East Coast. Johnstone's (1967a) classification is based on shared linguistic features, but the region has changed. Dialects have drifted apart in the last 5 decades as the political borders of the vast deserts of the peninsula have formed.

1.5.2 Dialects of Saudi Arabia

The dialects of Saudi Arabia form a continuum with those of Yemen in the south, the Emirates and Bahrain to the east, and Jordan and Iraq to the north; and they can be grouped into South-western, Eastern, Najdi and Hijazi (Ingham, 2009). The dialects of the southwestern regions can be further grouped into the dialects of the highlands, Tihamah valleys and coastal Tihamah; the first two have more features in common than either has with the last one (Prochazka, 1988). The highlands and Tihamah valleys are inhabited by the same tribes, who are mostly agricultural, though some are nomadic, while coastal Tihamah, separated by a desert tract, has a different tribal structure and culture (Prochazka, 1988). The Hijaz dialects comprise the urban varieties, spoken in the holy city of Makkah and in Jeddah, and the dialect of the B tribes, most of whom live on the periphery of these cities (Al-Essa, 2008). Urban Hijazi dialect is the linguistic outcome of contact between Arab and non-Arab immigrants who have been settling in these cities for generations (Al-Ahdal, 1989). It is the most innovative dialect in Saudi Arabia, resembling features of varieties spoken outside the Arabian Peninsula. In fact, Ingham (1971) characterises it as being of the Egypto-Levantine type. Eastern dialects of Saudi Arabia, and the Gulf in general, can be classified according to religious sect, Sunni or Shia (Holes, 1984). The Sunni dialect is spoken in Al-Hasa while the Shia dialect is spoken in Al-Qatif (Ingham, 2009). Najdi Arabic refers to the dialects spoken in the central regions of the Arabian Peninsula and stretching up to the Syrian desert (Prochazka, 1988; Ingham, 1994). Ingham (2009), however, distinguishes Central Najdi spoken in Riyadh province from Northern Najdi spoken by Shammar and Anizah tribes in Hayil and further northern areas of Saudi Arabia.
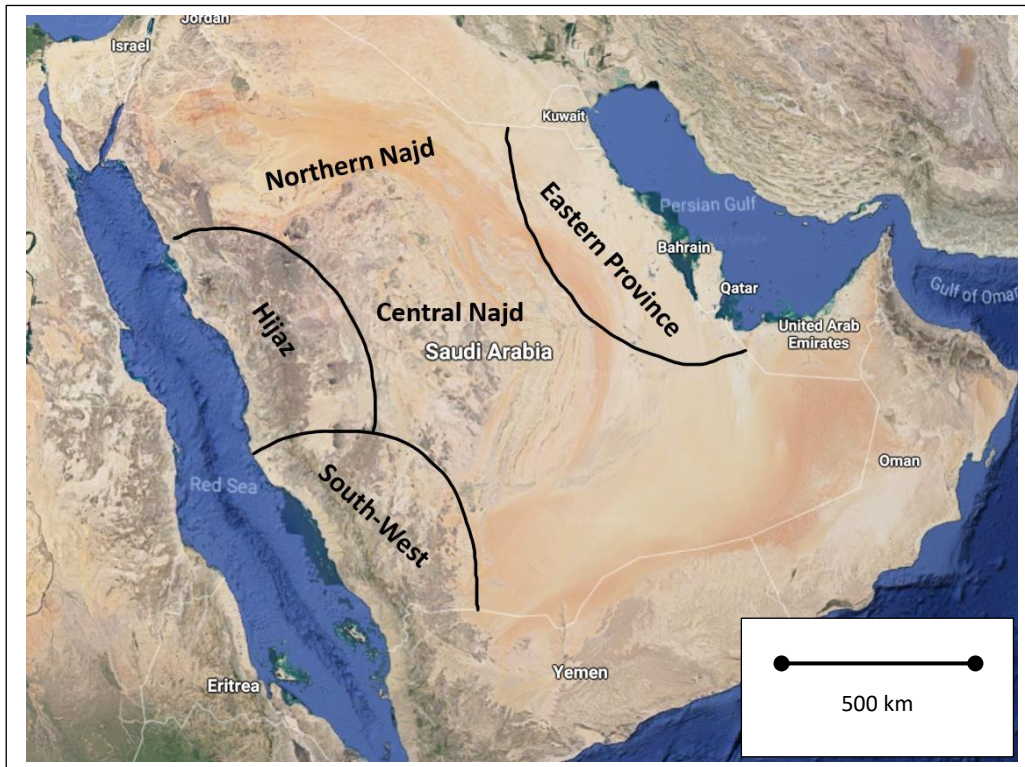
Figure 1.7: *Regional varieties of Arabic spoken in Saudi Arabia, based on Google Maps and Ingham (2009).*

To highlight the main linguistic features of the dialects of Saudi Arabia, it can be stated, with a degree of generalisation, that the dialects of the west and east coast (Hijazi and Eastern) are more innovative than the South-western and Najdi dialects (both central and northern). The most significant innovative feature of Hijazi and Eastern dialects is the absence of the interdentals. /θ/ and /ð/ in Hijazi are realised as /t/ and /d/, e.g., *tala:tah < θala:θah* 'three', *danab < ðanab* 'tail'; while /θ/ is reported as /f/ in some Eastern dialects, e.g., *fala:fi:n < θala:θi:n* 'thirty', (Ingham, 2009). Another major sound change in Saudi dialects is the merger of emphatics /ðˤ/ and /dˤ/. They have merged as /ðˤ/ in almost all Saudi dialects, while in Urban Hijazi dialect they merged to /dˤ/ (Ingham, 1971; 2009). Another linguistic feature shared by all Saudi dialects, including Urban Hijazi, is the change from /q/ to /g/, as opposed to /q/ to /ʔ/ in Cairene and Levantine Arabic. Lastly, the Saudi dialects, along with many

dialects around the Arab world, have lost the medial Hamzah (glottal stop), except in words borrowed from Classical Arabic.

1.5.3 Najdi Arabic

It has been suggested that Najdi Arabic is the most conservative contemporary spoken variety of Arabic, and this has been attributed to the area being relatively geographically isolated (Ingham, 1994, p.5; Versteegh, 2001, p.193). Among many conservative features that Najdi Arabic shares with other Arabic varieties, such as retention of interdentals and dual noun suffix, it retains highly archaic features of Classical Arabic, which include internal passive formation as in *yaktib* 'writing' > *yiktab* 'being written' and indefinite suffix *-in*, also known as *tanwīn*, as in *kitab-in* 'a book' (Ingham, 1994; Palva, 2006). It should be noted that these archaic features are proposed to have been retained by only Bs of Najd; however, data from the current study shows that tanwīn is used by S speakers as well. More discussion of theses features is provided in last section of this chapter and Chapter 8.

There are several descriptive studies of Najdi Arabic (e.g., Abboud, 1964, 1978, 1979; Al-Sweel, 1987). Ingham (1994), however, stands out as the most comprehensive descriptive account of the phonology, morphology and syntax of Najdi Arabic. It should be noted that he collected the data for this book in 1977, before the massive urbanisation. In addition, he considers Najd as stretching to the south as far as Najran on the borders with Yemen and to the north as far as the deserts of Syria and Iraq. In the current study, however, we limit ourselves to the variety spoken in Riyadh Province (i.e., Central Najd).

1.5.3.1 Phonology of Najdi Arabic

In this sub-section, the main phonological features and the phonemic inventory of Najdi Arabic are presented, all taken from Ingham (1994). It should be noted that he used Classical Arabic as a baseline for the changes undergone in Najdi Arabic. Also, it should be

emphasised that these changes might be shared by other Arabic varieties, inside and outside the Arabian Peninsula. The main changes to the phonemic inventory of Najdi Arabic are as follows:

- The pharyngealised voiced plosive /dˤ/ and fricative /ðˤ/ merged to / ðˤ/. Therefore, Classical Arabic *dˤaħik* 'to laugh' and *ðˤahar* 'to appear' are both realised in Najdi Arabic as *ðˤaħik* and *ðˤahar*.

- The medial glottal stop is lost and compensated with a long vowel, except for words borrowed from Classical Arabic. Thus, Classical Arabic words *raʔs* 'head' and *biʔr* 'well (*n*)' are produced as *ra:s*, and *bi:r* in Najdi Arabic. Ingham (1994) pointed out that words like *saʔal* 'to ask', in which the glottal stop is sandwiched between two short vowels, are produced as *saʕal* or avoided and replaced by *nishad* which is another word for 'ask'. The researcher, however, observed that these forms are now used by uneducated older speakers and that the younger educated generation of Najd, who are exposed to Modern Standard Arabic, tend to use the glottal stop in *saʔal* and all relevant words.

- The affrication of /ɡ/ and /k/ to /dz/ and /ts/ in front-vowel environment is one of the distinctive features of Najdi Arabic. The affricated form /dz/ is the result of a series of fronting processes involving Classical Arabic /q/ (i.e., q > *g* > dz). Examples of /ɡ/ > /dz/ are *gidir > dzidir* 'pot' and *tˤiri:g > tˤiri:dz* 'road'. The affrication of /k/ to /ts/, however, is another variant of the archaic sound change /k/ > /s/ (a phenomenon well known as *kaskasah*), which is still used in Najd (Johnstone, 1963). The form /ts/ is used in front-vowel environments as well as in replacement of the 2SF suffixed object/possessive pronoun -*k*. Examples of /k/ > /ts/ are *kalb > tsalb* 'dog' (front vowel environment) and *kita:bi-k > kita:bi-ts* 'your book. feminine' (2SF suffixed pronoun).

- Ingham (1994) noted an overlap between /u/ and /i/ as in *kull/kill* 'all' and characterised it as non-distinctive. This variation, however, is one of the most salient linguistic features distinguishing B dialect from S in Najd, and it will be extensively discussed in Chapter 4.

Table 1.1: *Consonants of Najdi Arabic (adapted from Ingham, 1994)*

	Labial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyn.	Glottal
Plosives emphatic	b			t d tˤ			k g	(q)		ʔ
Affricates					dʒ					
Fricative emphatic		f	θ ð ðˤ	s z sˤ	ʃ			χ ʁ	ħ ʕ	h
Nasals	m			n						
Trill				r						
Lateral				l						
Glide	w					j				
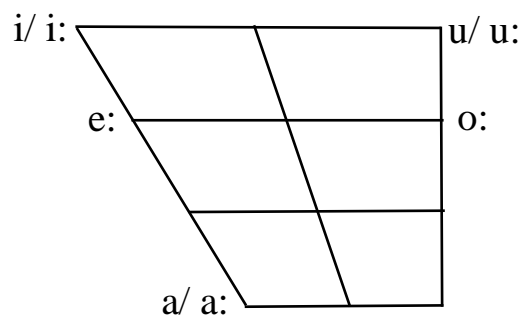
Figure 1.8: *Vowels of Najdi Arabic (Adapted from Ingham, 1994).*

1.5.3.2 Morphology of Najdi Arabic

Najdi Arabic, as well as other Arabic varieties, is well known for its non-concatenative morphology. In brief, a word in nonconcatenative morphology is composed of the following elements (Moore, 1990, p.64):

A) A consonantal root carrying the semantic properties. These roots are commonly triliteral (i.e., consisting of three consonants such as *ktb* 'to write' and *ksr* 'to break'), but they can be bi-literal (e.g., *md* 'to stretch', *mr* 'to pass by') or less likely quadriliteral (e.g., *dħrj* 'to roll').

B) A vocalic pattern which alternates according to voice (active or passive), aspect (perfective and imperfective), and sometimes across dialects (e.g., Najdi *kitab* 'book' compared to Hijazi *katab*). Examples of the vocalic patterns are *i-a* as in *k-i-t-a-b* 'he wrote' and *i* as in *yaks-i-r* 'he is breaking'.

C) A template in which both consonants and vowels are inserted. These templates (also known as verb forms) serve different semantic functions, such as the causative form CVCCVC as in *kattab* 'caused someone to write'. The verb forms of Najdi Arabic are shown in Table 1.2. All verb forms from1-10 are inherited from Classical Arabic. However, Najdi Arabic has developed three additional distinctive verb forms–potential *ntvfaʕʕal*, potential reciprocal *ntvfaːʕal* and pretensive *tvfeːʕal* (Al-Sweel, 1987).

D) Affixes with various functions such as the marking of person, number, gender, etc. The morphophonology of these markers changes according to aspect and verb form (as seen in Table 1.2). Additionally, some affixes change across dialects, and more relevantly in the dialects under study (e.g., B *ya-ktib* compared to S *yi-ktib*). This variation is studied thoroughly in Chapter 6. In Table 1.3 the verb conjugations of Najdi Arabic are provided.

Table 1.2: *Verb forms of Najdi Arabic. (the subject in all verb forms is 3SM), from (Al-Sweel, 1987)*

Perfective verb forms	Imperfective verb forms	Gloss
1) kitab, laʕab	*ya-ktib, ya-lʕab*	'write', 'play', respectively
2) ʕallam	*y-ʕallim*	'tell'
3) ra:sal	*y-ra:sil*	'message'
4) ʔadxal	*yi-dxil*	'cause to enter'
5) taʕallag	*yi-taʕallag*	'hang-reflexive'
6) tara:sal	*yi-tara:sal*	'message-reciprocal'
7) indʒalad	*y-indʒilid*	'being beaten'
8) istitar	*y-istitir*	'to cover oneself'
9) istafta:	*y-istafti:*	'to ask for fatwa'
10) iħmarr	*y-iħmarr*	'to become red'
11) ---	*yi-ntaraggaʕ*	'can be patched'
12) ---	*yi-ntifa:ham*	'can be negotiated'
13) taħe:wal	*y-taħe:wal*	'pretend to be cross-eyed'

Table 1.3: *Verb conjugations of Najdi Arabic (using the root ktb 'to write')*

	Perfective conjugations	Imperfective conjugations	
	B/S	B	S
1SC	*kitab-t*	*ʔa-ktib*	*ʔa-ktib*
2SM	*kitab-t*	*ta-ktib*	*ti-ktib*
2SF	*kitab-ti*	*ta-ktib-i:n*	*ti-ktib-i:n*
3SM	*kitab-Ø*	*ya-ktib*	*yi-ktib*
3SF	*ktib-at*	*ta-ktib*	*ti-ktib*
1PC	*kitab-na*	*na-ktib*	*ni-ktib*
2PC	*kitab-tu*	*ta-ktib-u:n*	*ti-ktib-u:n*
3PC	*ktib-au*	*ya-ktib-u:n*	*yi-ktib-u:n*

1.5.4 Najdi Arabic dialects

Ingham (1994, p.4) groups the dialects of Najdi Arabic as "1- The speech of the sedentary population of the areas of Central Najd (i.e., the districts of al-Ārid, al-Washim and Sudair), of Qasīm and Jabal Shammar to the north and Najrān and Bīsha to the south. 2- The speech of the main bedouin tribes of those regions i.e., Anizah, Utaibah, Subai, Suhul, Bugum, Dawasir, Harb, Mutair, 'Awāzim and Rashayidah in the centre, Shammar and Dhafīr in the north and Ghatan, Āl Murrah and 'Ājman in the south and east. 3-the speech of émigré bedouin tribes of Syrian desert and the Jazirah of Iraq of 'Anizah and Shammar extraction." He asserts that these dialects are classified as Najdi in a general sense based on shared linguistic features distinguishing them from neighbouring dialects. He then provides another regional sub-grouping as follows:

1. "Central Najdi. The dialects of Central Najd as described above and the central bedouin tribes also the 'Anizah of the Syrian desert.

2. Northern Najdi. The dialect of Jabal Shammar and of the Shammar tribes of Northern Najd and the Jazirah.

3. Mixed Northern-Central. The dialect of Qasim and of the Dhafir tribe.

4. Southern. The dialect of Najrān and the Ghatān [also written as Qahtan] tribe of the south and of the Al-Murrah and 'Ājmān tribes of the east." (Ingham, 1994, p.5)

The first group, which is the focus of this study, includes the speech of the Bs and the Ss of Central Najd. The Anizah tribe, however, is excluded as their speech resembles the Northern Najdi variety more than that of the Central one.[7] The second group is spoken by Anizah and Shammar tribes in the northern regions of Saudi Arabia in the cities of Hayil and Ar'ar, as

---

[7] Anizah is one the of the biggest tribes in Saudi Arabia, with some of its members still living in Syria and Iraq. Since the dialect of the Bs relates to the area from which they originate (Ingham, 2009), the dialect of members of Anizah tribe who inhabited Central Najd for decades is still characterised as Northern Najdi.

well as other cities, all of which are more than 700 km from Riyadh. The third group (mixed Northern-Central) is the dialect spoken in Qassim province, and it is more Northern than it is Central. Ingham (1994, p.193-194) outlines the differences between the northern and central dialects and states that the dialect of Qassim exhibits northern features except for two that are central (see Table 1.4). The last group covers two tribes (Qahtan and Ajman) who still inhabit the southern regions of Saudi Arabia, but a considerable number of members of both tribes migrated to Central Najd some time before the 1950s. This makes the last group of dialects unnecessary for this project because the dialect of a B tribe relates to the place from which they originate (Ingham, 2009, p.124). In other words, the dialect of members of the Ajman or Qahtan tribe who live in Central Najd will be relatively similar to the dialect of those who currently live in southern regions. However, in case there are minor differences between these groups, only Ajmans and Qahtans from Central Najd are included in the study.

Table 1.4: *Differences between Northern, Qasimi and Central Najdi dialects (object suffixed pronouns), adapted from Ingham (1994, p. 193-194)*

Northern	Qassim	Central	Gloss
-an	*-an*	*-ni*	'me'
-ak	*-ak*	*-ik*	'you. m'
-uh/w	*-uh/w*	*-ih/h*	'him'
-ah	*-ah*	*-ha*	'her'
-kam	*-kum*	*-kum*	'you. pl'
-ham	*-hum*	*-hum*	'them'

In the following sections (1.5.5), the dialects of Central Najd (henceforth, Najd) will be discussed. The differences between the B and S dialects and the selected linguistic variables will be highlighted.

1.5.5 The B and S dialects of Najd: The dialects under study and the linguistic variables

The distinction between B and S dialects of Arabic has been noted since the pre-Islamic period (Versteegh, 2001, p.42). B dialects are more conservative in general while S dialects are more innovative given the nature of urban environments (i.e. high contact with other dialects). In the early Islamic period, the dialect of the Bs was regarded as the purest form of Arabic (Versteegh, 2001, p.43). B dialects nowadays are obviously not as conservative as they were before, but, with a degree of generalisation, they can be still regarded as more conservative than the S dialects (Palva, 2006, p.606)

Among all the B dialects of the Arab world, the B dialect of Najd is regarded as the most conservative. Versteegh (2001) points out that "the conservatism of the Bedouin dialects [of Najd] contrasts with the reduction or simplification of linguistic structure in those areas where there is intensive interaction between Bedouin and sedentary population, for instance, in southern Iraq, along the Gulf coast and even in Mecca, which has a mixed population with many immigrants from outside the peninsula." (p. 188). It should be emphasised that the S dialect of Najd is still conservative compared to other S dialects.

The B dialect of Najd is spoken by several B tribes (refer to Figure 1.3) who migrated to the central deserts of the Arabian Peninsula more than eight decades ago (four decades before the oil boom). The history and origins of each tribe are unique in themselves, but the study will be limited to the three major tribes of Ajman, Dawasir and Qahtan, originally from the southern regions of Saudi Arabia. The other group, WB tribes, speaks Najdi B dialect, except for one linguistic feature (the merger) which they share with S dialect. The S dialect, on the other hand, is spoken by Ss who used to inhabit small villages such as Hotat bani Tamim, Ad-Dilam, Al-Yamamah and Addir'iyah, all of which are now towns or parts of cities. Although

there are minor differences within the B tribes and across S villages, the study will focus on linguistic features that are shared by all Bs and those shared by all Ss.

1.5.5.1 Merger of /u/ with /i/

Prochazka (1988) noted the merger of /u/ with /i/ in the short vowel inventory of some dialects in Saudi Arabia, while the distinction is retained in others. According to Prochazka's dialect classification, in which he classifies dialects based on tribe and/or city, the word *gult* 'I said' is realised as *gilt* in the dialects of the Riwaili tribe (a sub-group of Anizah tribe) and the cities of Hayil, Sudair, Riyadh and Hofof; on the other hand, the vowel /u/ is retained in the dialects of Najran city, the Qahtan tribe and the south-west region of Saudi Arabia (Prochazka, 1988, p.17, 77). Ingham (1994, p.14) noted an overlap between /u/ and /i/ in Najdi Arabic as in *kull/kill* 'all' but regarded this variation as non-distinctive. This variation is not random, as we shall see in Chapter 4; rather, it is sociolinguistically conditioned.

Based on Prochazka's observations on the merger, it can be stated that there is a sociolinguistic variation in Najd between originally southern B tribes and S groups. He noted that /u/ is lost in the dialect of Riyadh (i.e., S) and retained in the speech of the southern B tribes (i.e., Qahtan and Ajman from Najran), some of whom migrated to Najd in the early 1900s. The WB tribes, on the other hand, have fully undergone the merger, like the Ss. The merger is the most relevant linguistic variable of this study, as it shows conditioned variation among Bs. More in-depth background and analysis of this variable is provided in Chapter 4.

Table 1.5: *Example of /u/ merger with /i/*

B dialect	S dialect	Gloss
xuð	*xið*	'take -Imperative'
ygusˤsˤ	*ygisˤsˤ*	'he cuts'
kursi	*kirsi*	'chair'

1.5.5.2 The 3SM objective/possessive suffixed pronoun (-ih / -ah)

Another alternation in the dialects of Bs and Ss is the realisation of the 3SM objective/ possessive suffixed pronoun. For the Classical Arabic word *kita:bu-hu* 'his book', the Bs of Najd produce *kita:b-ih* while Ss use *kita:b-ah*. Although Ingham (1994) did not mention this variation, Prochazka (1988, p.11) and Al-Essa (2008, p.19) reported both forms to exist in Najd, but neither of them associated this variation with the B/S variation. Data from the current study, however, shows that *-ih* is used by Bs while *-ah* is used by the Ss. This suffixed pronoun is one of the linguistic variables in this study. More discussion of this variation is provided in Chapter 5.

Table 1.6: *Example of the 3SM objective/possessive suffixed pronoun (-ih / -ah)*

B dialect	S dialect	Gloss
sayya:rt-ih	*sayya:rt-ah*	'his car'
ra:s-ih	*ra:s-ah*	'his head'
ya-ksr-ih	*ya-ksr-ah*	'he breaks it.m'
ydʒi:b-ih	*ydʒi:b-ah*	'he brings it'

1.5.5.3 The imperfect prefixes (ya-/yi-, ta-/ti-, na-/ni-)

The alternation between /a/ and /i/ in imperfect prefixes across Arabic dialects has been noted by the ancient Arabic grammarians who referred to it then as *talatalah* (Rabin, 1951, p.61). Ingham (1994, p.194) noted this alternation in Najd and associated it with the B and S dialects. The Bs use /a/ in all imperfect prefixes while Ss use /i/ in these prefixes, except the 1SC prefix which is always *ʔa-* (see Table 1.3). Therefore, the B forms *ya-ktib* 'he writes', *ta-ktib* 'she writes' and *na-ktib* 'we write' are realised as *yi-ktib*, *ti-ktib* and *ni-ktib* in the S dialect. This variation is one of the linguistic variables in the current study, and more discussion is provided in Chapter 6.

1.5.5.4 Other linguistic variables

This section highlights other linguistic features that were not studied quantitatively, because they were infrequent, or those which proved not to pertain to a particular group. The observations and findings concerning these features will be discussed in Chapter 8.

1.5.5.4.1 Tanwīn

One way to mark indefiniteness in Classical Arabic is to use the marker *-n* (also known as *tanwīn* or *nunation*) as in *rajulu-n* 'a man' or *kita:bu-n* 'a book'. Because tanwīn is lost in almost all Arabic dialects, retaining this marker is one of the striking conservative features of the B dialect of Najd (Ingham, 1994, p. 6; Palva, 2006, p. 606; Versteegh, 2001, p. 164). As anticipated by Palva (2006, p. 611), it is being lost from these dialects due to increasing contact with the S dialects; and one of our goals in this study is to test this hypothesis (see Chapter 8).

1.5.5.4.2 Guttural effect

In Najdi Arabic, the guttural sounds (/χ/, /ʁ/, /ħ/, /ʕ/, /h/) cause the originally Classical Arabic initial syllable CaG- (where G stands for guttural) to re-syllabify as CGa- (Johnstone, 1964, p.80; Ingham, 1994, p.19). This rule can be seen in nouns such as *ʃha-re:n* (Classical Arabic *ʃah-re:n*) 'two months', but it is mostly evident in imperfective verbs that begin with a guttural (e.g., *yħafir < yaħfir* 'he digs'). Recall (Table 1.3) that imperfective verbs are prefixed with *ʔa-*, *ya-*, *ta-*, or *na-* and therefore imperfective verbs that begin with a guttural have the initial syllable CaG.

It is not reported in the literature whether CGa- forms are used by the Bs, Ss or both. In this study, we will investigate whether the Ss will apply the re-syllabification rule to imperfective verbs beginning with a guttural, or whether it will be blocked since the initial syllable for

these verbs in the S dialect is CiG-. Recall in section 1.5.5.3 that the Ss use the prefixes *yi-*, *ti-*, and *ni-*, and therefore the forms *yiħfɪr* or *yħafɪr* are both plausible. The answer to the question 'what form do the Ss use?' will be provided in Chapter 8.

1.5.5.4.3 Lowering of /e:/ and /o:/

In most contemporary Arabic dialects, the Classical Arabic diphthongs /ay/ and /aw/ have been monophthongised to /e:/ and /o:/ (or /u:/). Thus Classical Arabic forms *bayt* 'house' and *lawn* 'colour' are realised as *be:t* and *lo:n*. In Najdi Arabic, Ingham (1994) reported /e:/ and /o:/ as the monophthongised forms of the diphthongs. However, informal observations suggest that some Bs use lower versions of these vowels when followed by a short /a/. For the monophthong /e:/, some Bs produce *bɛ:ðˤa* 'white', *zɛ:nah* 'beautiful' and *ʃɛ:nah* 'ugly' instead of *be:ðˤa*, *ze:nah* and *ʃe:nah*. As for /o:/, on the other hand, it is realised as /a:/ in the same environment.[8] Examples are *sa:da* 'black', *na:sˤal* 'we arrive' and *ħa:la* 'cross-eyed, Feminine' (instead of *so:da*, *no:sˤal* and *ħo:la*).

1.5.5.4.4 Ħali:b vs ħili:b

As observed by the researcher, some of the B forms are stigmatised by the Ss. The most common stigmatised B words are *ħili:b* 'milk' and *ʕisˤi:r* 'juice' (instead of *ħalib* and *ʕasˤi:r*). Although this variation also exists in other lexical items, the stigmatisation seems to be associated with these two forms only. Najdi speakers do not overtly comment on this feature unless it occurs in these forms. More discussion and observations about this variation in the data are provided in Chapter 8.

---

[8] Ingham (1994, p.25) noted these forms but asserted that they are restricted to initial-weak verbs that begin with /w/ as in *ya:zan < wizan* 'to weigh'. However, this form is the B version of the Sedentary form *yo:zan* (< *ya:zan*). More discussion is provided in Chapter 8.

Table 1.7: *Examples of ħali:b vs ħili:b*

S dialect	B dialect	Gloss
ħadi:d	*ħidi:d*	'iron'
xabi:θ	*xibi:θ*	'wicked'
ħari:m	*hiri:m*	'women'

1.5.5.4.5 Internal passive

The formation of passive voice by alternation of the vocalic pattern is an archaic feature characteristic of Classical and Modern Standard Arabic. This feature is lost in all modern Arabic dialects, except for the B dialects of Najd (Ingham, 1982a, p. 188). It is however, expected to be lost due to contact with the S dialect (Palva, 2006, p. 611). It is difficult to investigate this feature quantitatively; however, based on our observations of the data we will confirm whether this feature is retained, in the process of being lost, or completely lost.

Chapter Two: Arabic Sociolinguistics

2.1 Introduction

In this chapter, we will review the topics that are relevant to this study. The first topic is the theory of the *Arabic Koine* by Ferguson (1959a). The Arabic koine, hypothesised to have formed as a result of the Islamic conquests, is believed to be an intermediate stage of Arabic from which the contemporary spoken dialects descend. References to the koine will be made in this research in regard to the differences between peninsular and non-peninsular dialects of Arabic. The second topic is diglossia in the Arab world (Ferguson, 1959b). Diglossia refers to the situation where there is a formal and an informal variety, both of which are used by the same speakers of one community. It is sociolinguistically relevant to this study because (a) the formal variety, Standard Arabic, shares two linguistic variables with one of the examined dialects, and (b) it is considered a possible variety to which speakers might code-switch. In the third section, a group of social variables in the Arab world will be reviewed. Lastly, the sociolinguistics of Saudi Arabia and the original contribution of the current study will be discussed.

2.2 Arabic Koine

Ferguson (1959a) argues that modern Arabic dialects share a striking number of linguistic features, when compared to Classical Arabic. However, while some features could be argued to have developed naturally from Classical Arabic (e.g., reduction of inflectional categories, loss of sounds like the glottal stop), other complex linguistic features are difficult to account for by simply assuming that all modern Arabic dialects have undergone the same changes by natural drift; see Miller (2006, p. 594) for more details. He then proposes that there is an intermediate stage of Arabic from which almost all modern Arabic dialects descend. This

intermediate variety is referred to as the Arabic Koine due to the circumstances that led to its formation.

The Islamic conquests started after the death of the Prophet Mohammed in 632 CE. The garrison towns in Iraq, Syria and Egypt were inhabited by Arabs from different tribes and non-Arabs. Due to the contact between these groups, a koine, in which all distinctive features of the different tribes' dialects were abandoned and many other features were simplified and reduced, is hypothesised to have formed (Al-Sharqawi, 2017; Ferguson, 1959a; Versteegh, 1984). A distinction is then to be made between non-peninsular dialects, which are descendants of the koine, and peninsular dialects, which developed naturally but were still influenced by the koine (Al-Sharqawi, 2017).

There is controversy as to whether modern Arabic dialects have evolved from the Arabic koine (e.g., Owens, 2006; Blanc, 1969). In the present study, however, we will adopt the theory of the Arabic koine. The theory is relevant to this study in respect to the typology of the linguistic variables. In each chapter on the linguistic variables, a comparison between the realisations of the variable in Najdi dialects and the other modern Arabic dialects is provided. Also, the distinction between peninsular and non-peninsular dialects is crucial to explaining the typological differences.

2.3 Diglossia in the Arab World

Diglossia refers to the situation where "two or more varieties of the same language are used by some speakers under different conditions … [In other words] many speakers speak their local dialect at home or among family or friends of the same dialect area but use the standard language in communicating with speakers of other dialects or on public occasions." (Ferguson, 1959b, p. 325). Ferguson refers to the standard language as the *high variety*, while the local variety is referred to as the *low variety*.

An example of the speech communities where this case is evident is the entire Arab world. Standard Arabic (or Al-Fuṣħa), considered the high variety, is spoken in formal settings. The preservation of the status of Arabic, and the language itself, is attributed merely to religious factors. Not only is it the language of the holy *Qur'an*, but also the language through which Islamic principles and practices are taught. Consequently, it is not only preserved in Arab Islamic communities, but it is also taught to non-Arab Muslims. It must be noted that the language of the holy book is considerably more sophisticated (morphologically and syntactically) compared to the Standard Arabic used nowadays. The language of the Qur'an is said to resemble Classical (archaic) Arabic while Modern Standard Arabic is the formal variety used in the present days (Fischer, 2006, p.398-399). Henceforth, Modern Standard Arabic will be referred to just as *Standard Arabic*.

On the other hand, in each Arab speech community, there is a colloquial spoken variety (low) which is radically different from Standard Arabic; hence, shifting from one variety to another is considered code-switching. The low variety is used in almost all informal conversations and encounters, and it is normally the first language that is acquired in childhood.

Ferguson (1959b) lists possible situations where the high or low varieties are normally used. However, for the case of Arabic this list has been criticised as being too categorical and impressionistic (Badawi 1973; Blanc, 1960; El-Hassan 1977; Hary 1996). More specifically, it is said to have "overlooked the range of sociolinguistic variation encountered in the speech of educated Arabic speakers." (Mahmoud, 1986, p. 239). This implies the following: educated speakers may alternate between Standard and colloquial variety on informal occasions; educated speakers might need to switch to the low variety in few formal situations if their audience is uneducated; uneducated speakers may fail to produce Standard Arabic in situations that require its use.

The list of possible situations where each variety is normally used in Najd is provided in Table 2.1. To account for sociolinguistic variation in the speech of educated speakers, an additional category (*acceptability of the other variety*) is added. This category shows whether the use of Standard Arabic is acceptable in situations where the colloquial variety is normally used, and vice versa. The sociolinguistic interview, as the situation where diglossia is relevant to this research, has been added to the list.

Table 2.1: *List of possible situations where each variety is normally used (adapted from Ferguson, 1959b)*

Situation/Context	Standard Arabic	Spoken Dialect	Acceptability of the other variety
Friday sermon (**khutbah**)	✓	X	X
Sermons in small mosques	✓	X	✓
University lecture	✓	X	✓
Political speech	✓	X	X
Newspaper	✓	X	X
Conversation with colleagues at work	X	✓	✓
Conversation with family and friends	X	✓	X
Shops	X	✓	X
Poetry	✓	✓	*NA*
(The sociolinguistic interview)	X	✓	✓

In Friday sermons, Standard Arabic is always used, and use of colloquial Arabic is not acceptable. In occasional sermons delivered in small mosques, in which the audience is usually older uneducated people, Standard Arabic is usually used, but shifting to the colloquial variety is very common. Conversation with colleagues is conducted in the low variety, but the use of Standard Arabic is acceptable, and in some situations preferred. When talking to family and friends, the use of the high variety is very unlikely, and those who do so will be the object of ridicule. As for poetry, two genres of poetry *Arabi* (i.e., in Standard

Arabic) and *Nabati* (i.e., in colloquial Arabic) have co-existed and been equally appreciated (Abu-Haidar, 2006b, p. 273).[9] Lastly, in the sociolinguistic interview, which is designed to meet the goal of this research, the use of Standard Arabic is discouraged; however, a few speakers view the situation as semi-formal and code-switched between Najdi and Standard Arabic. More discussion of this issue is provided in Chapter 8.

In studies which examined attitudes towards Standard Arabic as opposed to spoken dialects (e.g., Murad, 2007; Chakrani, 2011), it was found that educated speakers' attitude towards Standard Arabic ranked significantly higher than that of those who are less educated. Therefore, it is not surprising if educated speakers tend to speak Standard Arabic more frequently in situations where its use is deemed acceptable.

2.4 Sociolinguistic Variables in the Arab World

There are several extra-linguistic variables which influence language variation and change in the Arab countries. According to Bassiouney (2009, p. 123), these factors can be grouped into *semi-fixed* or *flexible* variables. Semi-fixed (or inherent) variables, e.g., ethnicity, religion and tribal affiliation, are difficult to change by individuals, depending on how salient these social variables are in a community. For example, in a country/region where sectarian divisions are salient (e.g., Sunni versus Shiite in Bahrain), it is essential to be associated with the group into which one is born. In most parts of the Arab world, 'semi-fixed sociolinguistic variables' are associated with 'social groups', which can help form a community of practice. On the other hand, flexible variables, e.g., social network, urban/rural and education, can possibly change. Within a social group, a change in any of the flexible variables is expected to induce language change. For example, Shiite individuals whose social network is loose, and which therefore

---

[9] In Abu Dhabi, two international poetry competitions, in the form of reality shows, are run every year. The first one, *Sha'ir Al-Malyoon* (literally translates to *Million's Poet*), is for colloquial poetry; while the second one, *Ameer Al-Shu'ara* (literally translates to *Prince of Poets*), is for Standard Arabic poetry.

brings them into contact with Sunnis, are expected to exhibit change in their speech at the age of adolescence.

In this chapter, we will only review the semi-fixed variables of ethnicity and religion. Under the flexible variables, we will review urbanisation only; tribal affiliation, social network and education will be reviewed in Chapter 3.

2.4.1 Semi-fixed sociolinguistic variables

2.4.1.1 Ethnicity

There is a controversy about what social parameters can be applied to determine whether a group forms a separate ethnicity. In the Arab world, particularly, it is not clear whether religious groups (e.g., Sunni and Shiite) are separate ethnic groups. Owens (2001, p. 434) argues that ethnicity in the Arab world is "any of a number of social parameters by which non-national social groupings are distinguished, including religion, shared history, skin colour, kinship, lineage and place of origin. The relevant criterion or criteria defining ethnicity may differ from place to place." Bassiouney (2009, p.98-99) disagrees, arguing that "religion is a more complicated and historically intertwined factor than it seems." According to her, ethnicity as a semi-fixed variable in the Arab world applies to the case of Jordan, where Jordanians and Palestinians are considered two separate ethnicities (nationalities), although both are Arabs. This classification is not attributed only to nationality, but because of a long-lasting political tension between the two groups (Al-Wer, 1999, p.39-40).

Using the realisations of the phonological variable /q/, the dialects of Jordan can be classified as the following (Cleveland, 1963; Abdel-Jawad, 1986):

- Urban Palestinian dialect ([ʔ] as in *ʔa:l* 'he said').
- Rural Palestinian dialect ([k] as in *ka:l* 'he said').
- Bedouin and rural Jordanian dialects ([g] as in *ga:l* 'he said').

The contact between the two ethnicities has been studied extensively (Abdel-Jawad, 1986; Al-Wer, 1999; Al-Wer, 2007b). It was found that in urban centres, where there is contact between the two groups, most dialect features of male Jordanians (e.g., [g]) are maintained because their social networks are very closed (Abdel-Jawad, 1986). Female Jordanians, on the other hand, are converging on the Palestinian (urban) variant [ʔ], as for them it is associated with modernity and prestige (Al-Wer, 1999). In Amman, the dialect contact between the two groups has resulted in the formation of Ammani koine (Al-Wer, 2007b). The koine consists of mainly Palestinian (urban) features, such as [t], [d], [dˤ] as opposed to Jordanian [θ], [ð], [ðˤ]. The realisations of /q/, however, have acquired stylistic variation, summarised as the following:

- [ʔ] is used among Palestinian boys (talking to each other).
- [g] is used among Jordanian boys.
- [ʔ] is used by both groups when talking to girls.
- [g] is used by both groups in mixed (Jordanian and Palestinian) boy groups (Al-Wer, 2007b, p. 66-67).

2.4.1.2 Religion

Religion as an independent variable is relevant to some parts of the Arab world, namely Iraq (Blanc, 1964), Egypt, Lebanon and most countries on the east coast of the Arabian Peninsula (Germanos and Miller, 2015; Versteegh, 2017). "Religion [in the Arab world] is important in terms of language variation and change only in the sense that it can create a close-knit community whose members feel for one reason or another that they are united by it." (Bassiouney, 2009, p.105). Bassiouney adds that religion is intertwined with political factors, which can then help predict the direction of language change.

The Sunni/Shiite sectarian division is the most prominent sociolinguistic variable in some regions of Saudi Arabia (Al-Qatif and Al-Ahssa), some parts of Kuwait, but more noticeably in Bahrain. Holes (1983) studied the correlation between Shiite/Sunni and B/S dialect variation in Bahrain. He posits that the origin of each group might be correlated with the sociolinguistic sectarian differences. The Sunni group are essentially Najdi Bs who recently migrated to Bahrain (late eighteenth century). The Shiites, who refer to themselves as *Baharna* (literally *Bahrainis*) to emphasise that they are the original inhabitants and separate themselves from the immigrants (the Sunnis), led a rural S lifestyle. Holes compared the data for both Sunni and Shiite groups with B and S dialects from various regions in the Arab world and concludes that Sunnis in Bahrain speak a B-type dialect while Shiites speak a S-type dialect.

According to Holes (1986; 2005), because the ruling family in Bahrain is Sunni, the Sunni dialect has acquired some prestige. Consequently, due to the increasing contact between the two groups, the Shiites are beginning to accommodate to the Sunnis. One of the salient variables is the realisation of /q/. The Shiite variant [q] (as in *qalam* 'pen') has become stigmatised, while the Sunni variant [g] has become associated with prestige and wealth.

Another place where there is linguistic variation by religion is Iraq. Abu-Haidar (1991) investigated the language differences between Christian and Muslim Arabic in Baghdad. Christians are a well-established community in Baghdad and their dialect is older than that of Muslims. The Christian dialect exhibits S features while the Muslim dialect is B. According to Abu-Haider, the Christians are converging towards Muslims, as salient features are adjusted towards the Muslim dialect. She explains that Standard Arabic is the high variety which is shared by the two groups. For Muslims, their dialect is the low variety, while for Christians, the Muslim dialect is another high variety. That is, Christian Baghdadis are triglossic, speaking their variety at home and with Christian friends, Muslim dialect in

everyday informal encounters, and Standard Arabic in formal settings (Abu-Haidar, 1991, p.143).

2.4.2 Flexible sociolinguistic variables: urbanisation

Urbanisation in Saudi Arabia, which is associated with the oil boom, has been reviewed in Chapter 1 (1.4). In this section, we will review urbanisation in the Arab world and the linguistic outcome of dialect contact in urban centres.

Miller (2004) summarises the evolution of the urban vernaculars into two types of processes:

> In the first case, the migration to the city produced a considerable transformation of the urban vernacular (Iraq, Mesopotamia, North Africa, Bahrain). The former urban vernacular became restricted to religious minorities and/or to women while the koineized Bedouin variety ended up as the dominant public urban standard language. In the second case, migration produced a levelling of the urban vernacular but without radical transformation (Cairo, Damascus) (p. 181-182).

The case of Bahrain and Iraq (Baghdad) have been discussed above. In North Africa, the Pre-Hilali urban dialects are found in old urban centres, while the Hilali B dialect is taking over in capital cities and new ones. The national urban koines of the capital cities, which are sedentarised B dialects, are spreading at the expense of old urban dialects which are now spoken by old women (Boucherit and Lentin, 1989; Caubet, 1998; Jabeur, 1987, 1995; Messaoudi, 1998, 2001: cited in Miller, 2004, p. 183). In the second case (as in Cairo and Damascus) the existing urban dialects were not subjected to bedouinization because the migrant population came from the neighbouring rural areas and their dialect was in most cases rural S. However, the main urban dialects, although maintained, were subjected to levelling

and koineization. For more discussion of the case of koineization and levelling of Cairene Arabic, see Miller (2004, p. 185-189).

2.5 Sociolinguistics of Saudi Arabia

There is a considerable number of studies that have examined language variation and change in most Arab countries.[10] As for Saudi Arabia, there are several sociolinguistic studies, some of which will be briefly reviewed.

Al-Jehani (1985) examined the dialect contact between the urban population of Makkah and the B tribes, particularly the realisations of the Arabic interdental fricatives (/θ/, /ð/, /ðˤ/). Given the rich history of the city as it attracts pilgrims from all around the world, some of whom have been settled in the city for several generations, it can then be stated that the urban dialect of Makkah is the result of contact between Arabic dialects and non-Arabic languages. The dialect of Makkah is characterised as Egypto-Levantine (Ingham, 1971), as discussed in section 1.5.2. One of the reasons that has led Ingham to such a characterisation is the realisation of the Arabic interdentals as [t], [d] and [dˤ] in the urban dialect of Makkah, as opposed to most dialects of the Arabian Peninsula in which they are retained in their original forms ([θ], [ð] and [ðˤ], respectively). In the dialect of B tribes in Makkah, who recently migrated from the periphery of the city, the realisations of these consonants as interdentals are generally retained; but Al-Jehani found that B speakers, particularly young and less educated ones, are converging on the urban stop variants.

Al-Ahdal (1989) examined the sociolinguistic variation in Makkah, focusing on prestige and style-shifting. The population of Makkah was grouped into tribal (mainly B tribes) and non-

---

[10] Here are some examples of these studies, listed by country: Egypt (Haeri, 1991; Schmidt, 1974); Iraq (Abu-Haider, 1991); Syria (Daher, 1999); Jordan (Abd-el-Jawad, 1986; Al-Wer, 1999, 2007b); Bahrain (Holes, 1983).

tribal (mainly Arabs and non-Arabs who do not belong to peninsular tribes)[11]. A three-level sociolinguistic prestige model was suggested: standard < national < local. The standard level represents Standard Arabic, which was switched to by both groups in formal contexts only. The national level is represented by the tribal dialect, since it is similar to the national variety of the country (Najdi variety). This level was adjusted to by some non-tribal speakers (conditioned by particular social factors) in some of the linguistic variables examined. The last level, local, was adjusted to by some tribal speakers when they desired to project their local urban identity.

A similar study to that of Al-Jehani (1985) was conducted in Jeddah by Al-Shehri (1993). He investigated the dialect of tribal B migrants in Jeddah, and to what extent they have accommodated to the urban dialect of the city. Unlike the situation in Makkah (Al-Jehani, 1985), Al-Shehri found that the interdentals' level of accommodation (discussed above) to the urban variants is very low. It was proposed that the realisations of the interdentals have become identity markers of each group, and therefore the Bs are projecting their identity by maintaining their variants.

Al-Ghamdi (2014) investigated the dialect of Ghamdi migrants in the city of Makkah. The Ghamdi dialect resembles the south-western group. The linguistic variables examined were the interdentals and the diphthongs (ay) and (aw). Al-Ghamdi found that these migrants are moving towards monophthongs (the Makkah variants), but they maintain their fricative variants of the interdentals. As proposed by Al-Shehri (1993), Al-Ghamdi suggests that the interdentals function as identity markers for these migrants and are hence maintained.

Al-Qahtani (2015) investigated the distinctive realisation of the classical Arabic /dˤ/ as [ɮˤ] in the Qahtani dialect of Tihamah (south-western regions). In most dialects of the south-western

---

[11] The Arabs are originally from Egypt and the Levant, while the non-Arabs are immigrants from other Asian countries, such as Pakistan, Afghanistan, Uzbekistan, etc.

regions and Najd, this sound is realised as [ðˤ], while in Urban Hijazi dialect it is realised as [dˤ]. Al-Qahtani found that there is a change in progress from [ʐˤ] towards [ðˤ] led by the younger women.

2.6 Sociolinguistics of Najd

Sociolinguistically, Najd is one of the most underresearched areas in the Arab world. There is an abundance of traditional descriptive studies of Najdi Arabic (e.g., Abboud, 1964, 1978, 1979; Al-Sweel, 1987, 1990; Ingham, 1994); but to my knowledge, there have only been two sociolinguistic studies of Najdi Arabic (Al-Essa, 2008; Al-Rojaie, 2013), which will be reviewed below. However, both studies tackle the Qassimi variety of Najd, which Ingham (1994) classified as *Mixed Northern-Central Najdi*; and as established in 1.5.4, it is a different variety from the variety examined in this thesis (Central Najdi).

Al-Essa (2008) examined the dialect contact between Najdi and Hijazi Arabic, which are considered two separate regional varieties. She examined the speech of minority Najdi (Qassimi) migrants in Jeddah (Hijaz). The quantitative sociolinguistic approach was used to study the co-relation between ten linguistic variables and the social factors of gender, age and level of contact with the Hijazi community. She found that "diffusion of the Urban Hijazi features is higher among high contact speakers who are engaged in frequent and intimate interaction with members of the Hijazi community. Socially salient phonological and morphophonemic variables are acquired by speakers who had an early input through an urban Hijazi caregiver" (Al-Essa, 2008, p. ii). The linguistic outcome of this contact was "summarized in three processes: (i) the leveling out of regional markers, (ii) simplification, and (iii) an increase in the morphophonemic regularity of the verb" (ibid., p. ii). Data from older Najdi speakers in her study shows that the salient features of Najdi Arabic in that contact situation (affrication of /t/ and /d/ to [ts] and [dz], respectively) are still in use. Al-

Essa, as a Najdi speaker who moved with her family to Jeddah, used her familiarity both with the dialects and the community to deliver a well-grounded analysis of the situation.

Al-Rojaie (2013) examined regional leveling of the affricate [ts] (< /k/) in Qassimi dialect. The affrication of /k/ in Qassimi dialect can be found in a phonological environment (adjacency to front vowels) and in the 2SF objective/possessive suffixed pronoun (realised as [-ki] in Classical Arabic and [-ts] in Qassimi dialect). It was found that "affrication is significantly favoured in the phonological context of front vowels, particularly the high front ones. Whereas suffix-based affrication is categorically used as [-ts], stem affrication is strongly correlated with the age, educational level, and gender of the speaker. In particular, older uneducated speakers of both sexes tend to maintain the use of the local variant [ts], whereas younger and middle-aged educated speakers, particularly women, increasingly shift toward the use of the supralocal variant [k]" (p. 43). This is due, he suggests, to the outward diffusion of the supralocal variant, associated with the dialect of major cities (Riyadh in this case), at the expense of the marked variant.

2.7 The Current Study and Original Contribution.

To the best of my knowledge, the current study is the first sociolinguistic investigation of Central Najd, particularly the B and S dialects. The original contribution of this study can be summarised in two points. First, after the drastic socio-political and economic changes in the region, little new information has been produced on the current sociolinguistic situation. Second, the only reference to the dialectal variation in Najd is found in the work of Ingham (1994), in which the linguistic differences between the B and S dialects of Central Najd are very limited.

2.7.1 The linguistic outcome of the massive urbanisation

Urbanisation has altered the socioeconomics of most countries in the Arab world, but Saudi Arabia was the most affected due to the oil boom. Miller (2007) points out that "[l]ittle is known about the linguistic outcomes of this massive urbanization process…For a number of cities, we find relatively old dialect descriptions, which usually fail to account for variation and change." (p. 2). The present study will begin to fill this gap, as it will investigate the dialect contact between the original inhabitants of Najd, who shifted from either a nomadic or rural lifestyle to an urban one.

The premise of this research is that Bs are expected to show convergence to the Ss, though the reverse remains possible. Cadora (1992, p. 1) suggests that in case of B-S contact, a B dialect goes through the following five stages before becoming fully urban:

Bedouin → Bedouin-Rural → Rural → Rural-Urban → Urban

Abu-Haidar (2006b) adds that "[t]his is a gradual process that takes place when nomadic tribes settle in sedentary rural or urban areas and level their speech patterns in the direction of the host communities' dialects. Linguistic leveling may be complete or partial, depending on how far the settlers accommodate to their sedentary environment" (p. 269). In our case however, the two groups have moved to the city at almost the same time, i.e., nomadic to urban (Bs) and rural to urban (Ss). Furthermore, the Ss do not view themselves as the original population of the city or that Bs are migrants, and obviously nor do the Bs. In the literature, Arabic dialect studies have concentrated on stable B or S dialects (Abu-Haidar, 2006b), but they "have only seldom provided us with information about the direction of ongoing development processes" (Palva 1994:459). In the present study, we predict that there is a change in progress towards the S dialect. We aim to answer the following research questions:

- What is the linguistic outcome of the massive urbanisation in Najd?

- Is there any dialect accommodation? If yes, who is converging to whom?

- What are the independent social variables governing this convergence?

2.7.2 Newly discovered linguistic features

Two of the variables that were quantitatively examined in this research were discovered by the researcher, who is a native speaker of Najdi Arabic (see section 3.2). These variables are the merger of /u/ and /i/ (Chapter 4) and the 3SM suffixed pronoun (Chapter 5). The other variable, the imperfect prefixes (Chapter 6), was explicitly mentioned in Ingham (1994). Another group of linguistic variables, some of which were discovered by the researcher, will be discussed qualitatively in Chapter 8. It can be stated, therefore, that this study establishes new distinctive linguistic features of the Central B and S dialects of Najd.

Chapter Three: Methodology and Data Collection

In this chapter, I intend to present the method of data collection (the sociolinguistic interview), social factors (age, tribe, education and contact), sampling and data analysis.

3.1 Introduction

Variationist sociolinguistics, which emerged as a field following Labov's foundational study of Martha's Vineyard in 1961, seeks to examine language variation based on one or more social factors within a speech community, as opposed to regional dialectology which examines dialectal differences across geographical areas. In a typical variationist study, a number of speakers (ideally a representative sample of a community), who are usually grouped by age, gender, social class, education, etc., are interviewed (the most common method for data elicitation), and the outcome is transcribed, quantified and statistically analysed (Milroy and Gordon, 2003; Tagliamonte, 2011).

As this relatively new research design is adopted across the world (specifically in the Arab world), slight changes to the methods of data collection have been made. For example, elicitation methods based on read speech, e.g., the use of word lists, tend to fail in Arabic-speaking communities as they trigger the use of Standard Arabic rather than casual/careful style within dialectal Arabic (Al-Wer, 2013).

3.2 The Position and Background of the Researcher

I am a Najdi B from Ajmi tribe, who was born and raised in Al-Kharj city (from which the sample is drawn). I am from the second generation of Bs, i.e., those who were born in the city to parents who lived in the desert for the first half of their lives and then moved to the city in their 20s-30s (first generation). I did not live in a B neighbourhood, but rather there were Ss

and Saudis from the south-western regions, whom I interacted with in and outside school. That is, I did not live in an exclusively B district, and I am familiar with other dialects of Najd and Saudi Arabia.

Being a member of the Najdi community helped me to approach and recruit participants, as well as to elicit the vernacular (to be defined in section 3.3.1.3) during the interview. Also, I was able to uncover the linguistic variables and predict the social factors that might be responsible for linguistic change, if there is any. More references to how being a native Najdi helped facilitate this research will be provided when needed throughout this chapter. Milroy (1980, p.80) points out that "the closer the field worker is matched to subjects in terms of various social attributes, the more successful he or she is likely to be."

3.3 The Sociolinguistic Interview

The sociolinguistic interview, developed by Labov in his study of New York City's Lower East Side (1966), remains central to variationist sociolinguistics. It is typically a one-on-one interview during which the interviewee is encouraged to produce stretches of speech, rather than short questions and answers. "Face-to-face interviews are the only means of obtaining the volume and quality of recorded speech that is needed for quantitative analysis" (Labov, 1984, p.32). There has been some criticism of the method, and several modifications have been proposed since it was first adopted. In this section, I will discuss the sociolinguistic interview as a data collection tool and then proceed to the structure, design and procedures of the interview in this research.

3.3.1 Review of the sociolinguistic interview

3.3.1.1 Suggested length of the interview

Due to the open-ended nature of the sociolinguistic interview, one of the controversial questions is how long the interview should last. Labov (1984) suggests that it should last for

1-2 hours. The reason for this time frame, as also asserted by Douglas-Cowie (1978), is that speakers tend to be more relaxed after the first hour, and therefore the style of their speech during that time is likely to resemble the 'vernacular' (to be defined below in section 3.3.1.3), which is of interest to variationists (see below for further discussion of this point). This is somewhat at odds with the findings of Schilling-Estes (1998), who suggests that speakers' style may fluctuate because of various reasons, regardless of how much time has elapsed.

Overall, considering the main objective of the interview, which is acquiring sufficient good-quality data, the length of the interview may vary according to the frequency of the variable(s) examined and how relaxed the interviewee is. Milroy and Gordon (2003) suggests that "useful phonological data can often be obtained in a relatively short time – perhaps as short as 20 to 30 minutes" (p. 58). If the interviewee is relaxed after the first 10 minutes (due to various factors, e.g., knowing the interviewer, the dialect and appearance of the interviewer being similar to his/hers, the place of the interview), it can be assumed that 30-40 minutes, excluding the first 10 minutes, should be enough to yield sufficient data, depending on the frequency of the variable.

3.3.1.2 Suggested structure and design of the interview

The sociolinguistic interview is a semi-constructed event in which interviewees are expected to respond to open-ended questions, with the goal of obtaining sufficient speech for quantitative analysis. The researcher prepares a list of questions prior to the interview. There is no exact format for these questions, but a set of rules (such as those found in Labov, 1984) must be taken into consideration when preparing the questions: (a) they should not be written in a bookish style or lexicon, but rather they should be short and colloquial; and (b) they should be designed to encourage an extended narrative style of speech, rather than short answers.

In order for the sociolinguistic interview to be successful, there must be careful preparation and planning. The goal of the interviewer is to engage in free conversation with the subject, and it is the interviewer's responsibility to manage and keep the conversation going. Therefore, it is essential for the interviewer to have a list of topics prepared. (Milroy and Gordon, 2003, p.58). Labov suggested the interview technique of *conversational networks*, which helps the interviewer simulate an unstructured conversation and allows him/her to navigate around the questions with a natural transition between topics (for more information see Labov, 1984, p. 35).

The structure of the Labovian sociolinguistic interview is based on his model of *attention to speech* (Labov, 1972). The first part of the interview, i.e. open-ended questions encouraging speakers to produce narrative style speech, is designed to elicit a conversational style, which is then divided into casual and careful style. The additional tasks of reading passage, word lists and minimal pairs are then introduced to speakers, with the objective of gradually shifting the attention to the language (i.e. the tested variables) and helping to build a formality continuum for comparison.

Criticism of the attention to speech model revolved around two major issues. The first issue concerns the operationalisation of distinguishing careful from casual style in the body of the interview. Labov proposed the examination of channel cues (Labov, 1966), e.g., change in volume and pitch in the interviewee's speech, but this idea was rejected early on (see Wolfram, 1969). Labov also designed the decision tree (Labov, 2001, p. 94), which has similarly not been widely adopted (Becker, 2013). The second issue concerns the more formal tasks of the reading passage, word lists and minimal pairs, mainly questioning whether conversational and reading styles are comparable types of linguistic behaviour (Milroy, 1980). Another problem with these tasks is the assumption that all participants are literate

(Baugh, 2001, p. 110). Moreover, in Arabic-speaking communities these tasks would trigger the use of Standard Arabic due to the diglossic situation (Al-Wer, 2013).

An alternative to the attention to speech model has focused on the role of *addressee* in stylistic variation. In their attempt to capture the properties of *African-American Vernacular English*, Labov et al. (1968) resorted to a new approach in which speakers were once recorded in a group (interacting with other speakers of the same variety) and once with the fieldworker (one-on-one). The style of speech was found to be different according to the addressee. Later, an addressee-based model was proposed by Bell (1984), named *audience design*, which he describes as follows:

> "It assumes that persons respond mainly to other persons, that speakers take most account of hearers in designing their talk. The *speaker* is first person, primary participant at the moment of speech, qualitatively apart from other interlocutors. The first person's characteristics account for speech differences between speakers. However, speakers design their style for their audience. Differences within the speech of a single speaker are accountable as the influence of the second person and some third persons, who together compose the audience to a speaker's utterances. (p. 159).

The audience design model is based on a study of New Zealand radio newscasters, who were found to style-shift when reading the same news items on different stations, i.e. with a different audience (Bell, 1977). Following a similar model, Rickford and McNair-Knox (1994) tested the stylistic variation in the speech of an African-American teenager, who was interviewed first by an African-American and then by a European-American interviewer. The data from the two interviews showed that the interviewee altered her style according to the addressee (interviewer).

73

3.3.1.3 The vernacular and the observer's paradox

The ultimate goal of the sociolinguistic interview is to capture the most spontaneous form of speech, in which the least attention is paid to the language. Labov (1984) refers to this style of speech as the *vernacular*, the variety acquired in pre-adolescence years. Eckert, on the other hand, refers to it as "the language of locally based communities" (2000, p. 17), a definition asserting this variety's status as distinct from the supra-local standard variety (Milroy and Gordon, 2003, p. 50). The pursuit of the vernacular stems from the belief that it is the most systematic form of speech, because it is not influenced by a higher status variety.

The major obstacle to capturing the vernacular ethically is that "any systematic observation of a speaker defines a formal context where more than minimum attention is paid to speech … [however] face-to-face interviews are the only means of obtaining the volume and quality of recorded speech that is needed for quantitative analysis." (Labov, 1984, p.28). Labov refers to this phenomenon as the *observer's paradox* because the researcher is interested in the kind of speech produced when the subject is not being observed. The effect of the observer can be reduced by improving the fieldworker–participant relationship and implementing various practical techniques that have been introduced as the field evolves. For example, Eckert (2000), in her study of a high school in Detroit, spent two years as an unofficial presence in the school, improving her long-term relationship with the students. When interviewing them, she stated that they were sufficiently comfortable with her that they discussed sensitive topics, even introduced her to their friends. As for practical techniques, Labov (1984) suggested introducing topics, like childhood rhymes/customs and near-death experience, that might shift the attention away from speech. Another technique, again suggested by Labov, is for the fieldworker to portray himself/herself as the learner in order to reduce the power differential and thus help make the speaker relaxed. In Arabic sociolinguistic studies, Bassiouney (2009) lists how some linguists attempted to overcome the observer's paradox: Haeri (1997), an

Iranian woman studying language variation in Cairo, asserted her religious identity (Muslim) which she shared with her informants; Holes (1984, 1986), while studying the Shiites and Sunnis of Bahrain, told his participants that he is interested in social customs, rather than language; Dahir (1999), as an outsider studying variation in Damascus, recruited an insider interviewer who was trusted by the community; and Jabeur (1987), in his study of Rades, Tunisia, introduced himself as friend of a friend, rather than creating a new relationship with his interviewees.

3.3.2 The interview in this research

3.3.2.1 Objective

On how to obtain useful data in sociolinguistics, Milroy and Gordon (2003) state that "[t]here are a variety of approaches available to sociolinguists. In choosing from among these – as with all decisions in the study design – the investigator is guided by the aims of the research. What constitutes "good data" depends on the research objectives, as do the methods for collecting such data. Decisions about data collection are crucial because patterns of language use are sensitive to various contextual factors" (p.49). Therefore, due to the structure of the community and the linguistic situation in Najd, the interview design in this research focuses on eliciting spontaneous speech through free conversations, ignoring Labovian contextual styles. Our aim is that the spontaneous speech in the interview will hopefully represent the version of B/S dialect interviewees use with acquaintances and people (not outsiders) they meet for the first time informally. That is, it is the least formal variety we can obtain which might be expected to exhibit variation/change. If captured, the vernacular on the other hand might not exhibit variation because the two groups have been in contact for only five decades and the style used with family and close friends might be completely B or S. Following the

addressee/audience design model (Bell, 1984), a suggested formality continuum for Bs and Ss, based on my informal observations, is laid out in Figures 3.1 and 3.2.

Most Formal			Vernacular
Standard Arabic	Najdi S Dialect	B dialect Variety 2	B dialect Variety 1
Used in formal contexts	Used with Najdi strangers (B or S) in all informal encounters	Used with B acquaintances and S close friends	Used with family and B close friends
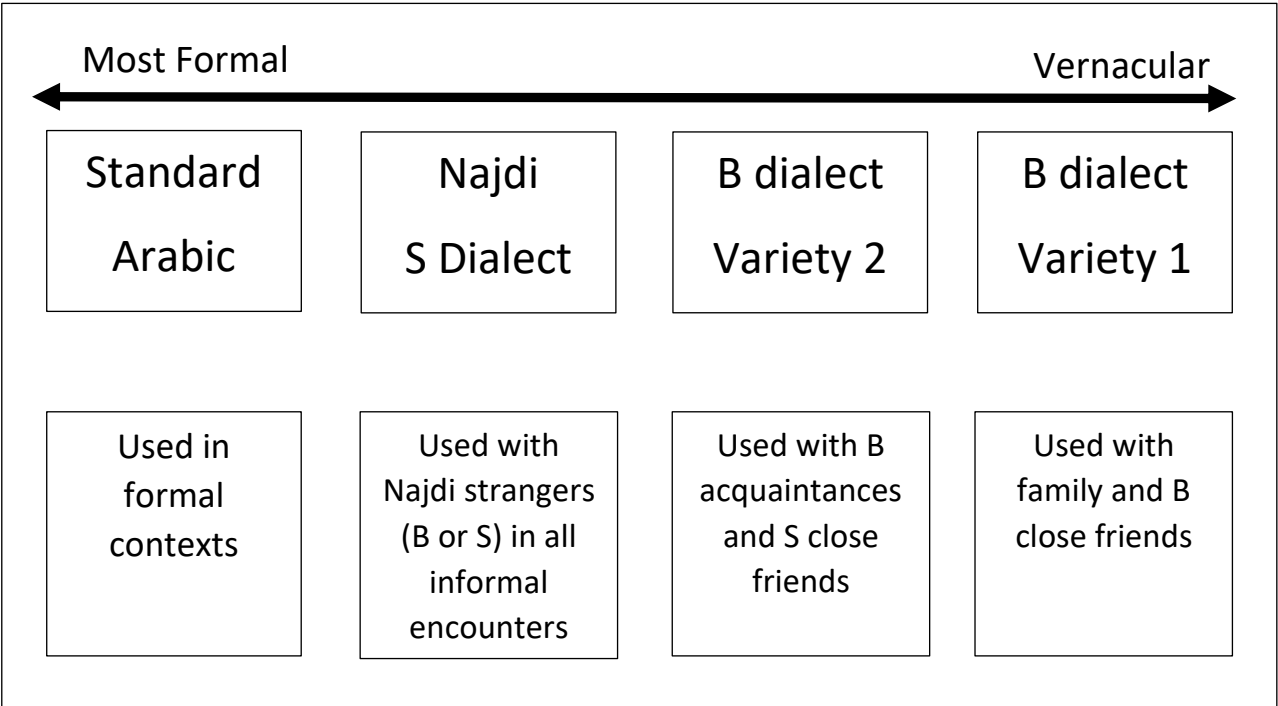
Figure 3.1: *Suggested formality continuum for Najdi B speakers.*

Starting with the Bs' formality continuum (Figure 3.1), the most formal variety is Standard Arabic, a higher variety that is triggered by not only the addressee but by situation and topics as well. It is not the objective of this study to elicit this variety as shifting to it is considered by Arabic linguists as code-switching rather than style-shifting. It is included here, however, to emphasise that there is an extreme level of formality to which Najdi speakers may 'code-switch' in particular situations/contexts. The next formal variety for the Bs is Najdi S dialect. This variety is used by Bs when they are addressing a Najdi stranger (B or S) in almost all informal occasions. Variety 2 is an informal variety used by Bs when they are addressing acquaintances and close S friends. Variety 1 represents the vernacular, and it is expected to show no changes.

It is the hypothesis of this research that Bs' effort to approximate their speech to the S dialect is what causes them to produce Variety 2. In other words, Bs converge to the S variety but their production is inconsistent, fluctuating between B and S. It is hypothesised, at least in our sample, that no B is able to completely shift to the S dialect. When interviewed by the researcher (B) in a very relaxed informal setting (see section 3.3.2.4), Bs are expected to exhibit S features because of the formality of the interview and the observer's paradox; however, the rate of convergence is expected to increase, but not code-switch to Standard Arabic, when they are interviewed by a S speaker (see Chapter 7).

Most Formal		Vernacular
Standard Arabic	**Najdi S dialect**	**Local S Dialect**
Used in formal contexts	Used with Najdi strangers (B or S) in all informal encounters	Used with family and close S friends
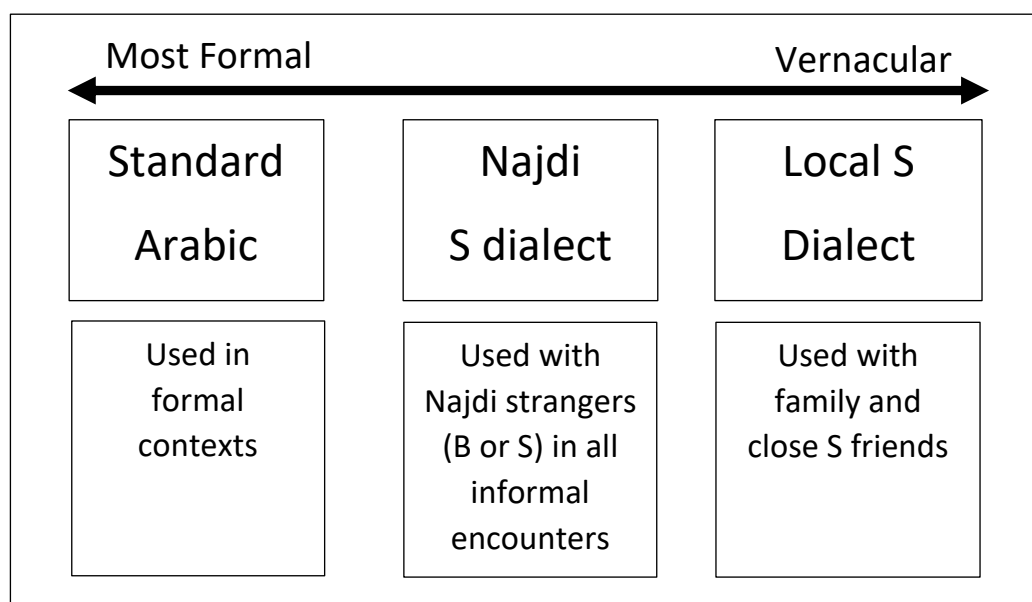
Figure 3.2: *Suggested formality continuum for Najdi S speakers.*

As for the formality continuum for Ss (Figure 3.2), it is hypothesised that Ss, unlike Bs, have two styles (beside Standard Arabic). The local S dialect, representing the vernacular, is used with family and close S friends who share the same local features, while Najdi S style is used elsewhere. That is, each local S dialect has marked features and lexical items that are avoided by its speakers when they are addressing people with whom they do not share these features. It should be noted, however, that as far as the investigated variables are concerned, local S

dialects and Najdi S variety are the same. Therefore, the goal of this study when interviewing Ss, is to have them relaxed enough to avoid using Standard Arabic, and consequently the outcome is expected to be consistent. In other words, Ss are hypothesised to show no convergence to the B dialect or other dialects.

3.3.2.2 Structure and Design

Since the length of the interview is determined by the goal of the research (Milroy and Gordon, 2003), the interview in this research lasted on average 40-60 minutes, which is more than enough to elicit a sufficient number of tokens for the examined phonological and morpho-phonological variables. The minimum number of tokens for statistical analysis is 30 tokens per variable per speaker (Tagliamonte, 2011, p. 136). Given the average length of interview in this study, the number of tokens yielded ranges between 35 and 70 per variable per speaker in most cases.

Since the objective of the interview in this study is to elicit spontaneous speech, a list of topics has been prepared in advance, taking into consideration the following:

- Formal topics that might trigger Standard Arabic are avoided.
- The topics are designed to shift attention away from speech and encourage the production of a more casual style, e.g., childhood memories (Labov, 1984).
- When introducing the topics to the interviewee, the researcher does not use a bookish formal style (Labov, 1984). Instead, the following request template '*can you talk to me/us about…?*' is used.

Below is the list of proposed topics used for the interview in this study:

- School: do you remember your first day at school? What do you remember?
- Work: first job, first day at work, first supervisor.

- Blue-collar/retail jobs for Saudis, and how attitudes have changed towards them.

- Marriage: marriage situation in Saudi Arabia:

  - Expenses

  - Traditions

  - Balancing religion and globalization

  - Difficulties facing potential couples.

- Hobbies: previous/current hobbies, future hobbies.

- Memorable events: occasions that have stuck in your memory (either good or bad), and what makes them unforgettable.

- First things (first car, first phone, first friend, etc.)

- Careers for women.

- Any topics the researcher may find the interviewee is interested in.

- Additional broad topics: cooking, desert picnics, hooliganism, parenting, finance, business, health, immigration, TV programmes and shows.

The second part of the interview is a picture task, designed specifically to elicit more tokens of the infrequent *guttural effect* variable (see Chapter 8, section 8.3.2). Although the picture task has been adopted by many sociolinguists, not least those investigating Arabic dialects, in the present study it has yielded a mixture of Standard Arabic and spoken Najdi tokens for most interviewees. More discussion of the outcome of this task is provided in section (8.3.2).

3.3.2.3 Procedures / Overcoming the observer's paradox

In this section, the recruitment technique and physical setting of the interview will be discussed in term of how they are designed to help overcome the observer's paradox. Next, the procedures of the interview will be presented.

When recruiting speakers, I asked for help from relatives (father, brother and cousin) and close friends. The strict use of *friend of a friend* technique, rather than the *snowball* method, has played a major role in reducing the tension and formality of the interview. In other words, because recruited speakers knew that the researcher was a relative of their close friend, they tended to trust him with recording their speech and were more comfortable chatting with him before, during and after the interview. Besides, in the Najdi community, and most parts of Saudi Arabia, meeting a friend of a friend for the first time and chatting with him/her is not something out of the ordinary; rather, it is a desirable event where they are excited to meet new faces (but not complete strangers) and share stories about work and common interests with them.

The place of the interview was also a friendly welcoming environment that helped encourage speakers to engage in chatting and expressing themselves freely. The speakers were asked to choose the place of the interview. When some speakers preferred or strongly suggested a particular place, their preference was always granted. If not, I suggested several places and let them choose without any pressure. The places where the interview was conducted were: the researcher's house, the interviewee's house, the moderator's house,[12] or a public place (usually a coffee shop). Due to the design of houses in Saudi Arabia, they are the optimal place to conduct the interview. In each house/flat, there is always a separate guest room by the entrance where guests, usually males, always sit and dine. Such rooms are usually quiet and do not cause any disturbance to the family of the host. It should be noted that all rooms, where the interviews were conducted, were carpeted and fully furnished which did not create any echoes in the recordings.

---

[12] 'The moderator' here refers to the direct friend of the interviewee. In this case the moderator was my father, brother, cousin or close friend.

Another practice that is known to help reduce the observer's paradox and facilitate the production of casual speech is changing the dynamics of the interview from one-on-one to a group event (Milroy and Gordon, 2003). For example, two interviewers have been used in Wolfram and Schilling-Estes (1996) to help (a) reduce the tension of a one-on-one encounter, (b) keep the conversation going and (c) give fieldworkers the chance to take notes without stopping the conversation. The problem with this setting is that some interviewees might feel overwhelmed and outnumbered by what they perceive as two strangers. Another setting is to have two, or more, interviewees (Labov, 1986), which encourages the interviewees to speak casually as they feel that they outnumber the interviewer. There are three problems with this setting if it is intended to gather data from all interviewees: (a) one speaker might dominate the conversation resulting in less recorded data for the other(s), (b) the length of the interview needs to be extended to elicit enough data for all speakers, causing the interview to be too long, and finally (c) the quality of recorded speech might be affected by overlaps during the conversations.

In this study neither of the previous settings have been used, but rather an adaptation of Gumperz's technique (1982) has been used. While Gumperz conducted the interviews with a family member present, in this study the moderator was present in most interviews, making the situation optimal for the interviewee (i.e., he is chatting with his friend in the presence of a third party, who is also not a complete stranger). In the few cases where the moderator was not present, I had already established contact with the interviewee on another occasion. To elaborate, for example I met some of my brother's friends at an informal dinner and then approached them to find out whether they were willing to be interviewed. They agreed in every case, and an appointment was made at a more convenient time/day for them. When we met afterwards, without the presence of the moderator, the interviewees were completely relaxed because the relationship had already been established.

81

The presence of the moderator was beneficial in reducing observer's paradox effects because: (a) there was a familiar face during the interview, (b) most topics were discussed fondly (i.e., past humorous or embarrassing stories, in which the interviewee and the moderator were involved, were shared with the interviewer), and (c) the conversation was free-flowing and there was rarely a shortage of data. It should be noted that although the role of the moderator in the interview was extremely important, they were reminded prior to meeting the interviewee to keep their talking to a minimum but still engage in the conversation when needed.

Before conducting the interview, the researcher exchanged friendly chat with the interviewee, usually about their relationship to the moderator. During that time coffee and tea were offered by the host, whether the researcher or the interviewee. Before recording, it was emphasised that the setting is informal and the use of spoken dialect, as opposed to Standard Arabic, is encouraged. The recording device, Olympus LS-14, is equipped with two directional and one omni-directional microphones, all of which are capable of recording high-quality audio without the need to use a head-worn microphone. All efforts have been made to keep the microphone distance constant (about one meter away from the interviewee). During the first 3-5 minutes of the interview, ethnographic information–age, tribe, occupation, education– was collected. After that they orally answered a short questionnaire designed to measure their level of contact with the other group (see section 3.5.5.3).[13] The choice of topics was left to the speaker, but they mostly requested a topic to start the chat. At this stage, individual differences between interviewees played a major role in determining the direction of the conversation: some speakers started with one topic and proceeded through the entire interview without using the topics list; most went over the list but chatted for about 5-7 minutes about each one; and very few went over the topics quickly providing short remarks about each one.

---

[13] The questionnaire was in the form of short questions asked by the interviewer (in Najdi Arabic).

As interviewer, I kept my talk to a minimum, except with the last group where I had to talk more to encourage them to talk.

The interviews went smoothly with no problems. Some speakers, however, were less talkative, producing short interviews and therefore barely sufficient data. Others were the sons of mixed marriage, e.g., B father and S mother, and therefore excluded (see 3.6.1.3 below).[14] Another minor issue is that four S speakers, two old and two young, switched to Standard Arabic during some or most parts of the interview, despite all efforts made to reduce the chance. The two older speakers were competent in Standard Arabic and overtly commented about how frequently they use it, even with friends. The other younger speakers preferred to talk about religious topics, which triggered Standard Arabic. More discussion of this issue is provided in section 8.5.

3.4 The Linguistic Factors

Linguistic factors often play a role in determining the direction of linguistic change, mostly along with social factors. For instance, a group of speakers belonging to a particular age group (a social factor) may be found to use variant X more frequently when it is preceded by a particular sound (a linguistic factor). In this study, however, no linguistic factors were detected for any of the three variables. After the data was transcribed, the intended tokens have been examined closely, and there was no sign of the influence of such factors. To elaborate, the S tokens produced by Bs have been grouped in one document, using *data filter* in Excel, and possible linguistic factors, e.g., preceding/following sound, stress, word class,

---

[14] I was able to detect particular linguistic patterns which suggested they had a S background, and with extreme politeness I asked them whether their mother is from the same group. The reason why we always ask about the mother is because the last name, which indicates tribe and hence B or S, is always of the father. Three speakers, who were the sons of mixed marriage, were excluded.

were found to be irrelevant. Besides, on numerous occasions, speakers have used the same form once with the B variant and once with the S during the interview.

Linguistic factors are not to be confused with the *envelope of variation* (Milroy and Gordon, 2003, p.180), which is the only context in which a variable occurs. The envelope of variation of each variable is provided in the chapter designated for each.

3.5 The Social Factors Under Study

The social factors which are suggested to be explanatory for sociolinguistic variation in the Najdi community are age, tribe, education and level of contact, each of which will be discussed below. In the next section, however, I shall explain why the social factor of gender has been excluded in this study.

3.5.1 Exclusion of gender in this study

Gender has been excluded for both practical and theoretical reasons. Practically speaking, it is religiously and culturally unacceptable in the Najdi community for a man to interview a woman. Also, most women would not be comfortable if their voice is recorded and would be listened to by a man, except for blood relatives. The Najdi community is considered the most conservative in Saudi Arabia, to the extent that some women may refuse to have their voice recorded even if they are interviewed by a female interviewer and assured that no men will listen to it (Al-Essa, 2008). Theoretically, due to gender segregation in almost all social spheres in Saudi Arabia, men and women form separate communities exhibiting rather different sociolinguistic practices. I believe the sociolinguistics of women in Najd must be further investigated separately by a female sociolinguist as it is almost impossible for a male researcher to gain access to that community.

3.5.2 Age

One of the chief social factors by which a sample is grouped and expected to exhibit different linguistic behaviour is age. This factor is relevant because "establishing that a pattern of variation represents a change in progress typically requires the consideration of speakers of different generations." (Milroy and Gordon, 2003, p.35). Age is believed to enable sociolinguists to monitor the development of a linguistic change by providing easy access to each stage via recording the speech of individuals who were born in that particular stage/generation (Labov, 1963). This principle is referred to as the *apparent-time hypothesis*, as opposed to *real-time* in which a linguistic change is studied in real time. It is indubitable that real-time analysis is more reliable than apparent-time, but the latter is obviously adopted to overcome the problems of making recordings across an extended period of time and the availability of recordings of the target variety. In other words, instead of recording data over a long period of time, or comparing new data to old recordings (which is usually of low quality), synchronic observations of the speech of different generations can help draw diachronic inferences on the development of a linguistic change.[15] Labov (1972), however, suggests that there should be "at least one measurement at some contrasting point in real time" (p.275). This measurement can be old recordings if available, or earlier descriptions of the target features/dialect.

An age-related issue one should be aware of is *age-grading* differences, in which the use of some forms is likely to be associated with a stage of life rather than indicating a linguistic change. Although such differences have been reported in a few studies (see Chambers, 2009 p. 201), they are not common in the sociolinguistic literature (Milroy and Gordon, 2003, p.35).

---

[15] See Milroy and Gordon (2003, p. 35)

Age has not usually been used as an explanatory factor by itself, but rather its interaction with other social factors is what interests variationists. In Amman, Jordan, the speech of the older generation of the two ethnicities that make up the population of the city, Palestinian and Jordanian, tends to show more local features than the speech of younger speakers, who tend to converge towards the standard dialect (Al-Wer, 2002b). Another, broader, example is the interaction between age and education, especially in the Arab world. In some Arabic variationist studies (e.g. Jabeur, 1987; Sadiq, 2016), age corresponds to the level of education, i.e., older speakers are mostly uneducated. Education, therefore, entails exposure to Standard Arabic and in most cases more contact with other dialects. In studies where age does not interact with other social factors, e.g., Al-Essa (2008), it was found to be irrelevant to linguistic change.

In the current study, age is employed as a social factor because it provides a mirror to the dialects before and after they were in contact. However, it is likely to inter-correlate with other social factors. For example, a B speaker in his 70s, who has probably spent half of his life in the desert and the other in the city, is less likely to exhibit any changes in his dialect. However, if he moved to the city earlier, had access to education and lived in a S neighbourhood, he is more likely to partially converge to the S dialect.

3.5.3 Tribe

Tribal affiliation is one of the semi-fixed social variables in the Arab world, beside ethnicity and religion (Bassiouney, 2009). It is, however, established that some countries/communities are more tribal than others. In most Gulf countries, as well as Iraq and Jordan, tribal affiliation is more dominant than in other countries; and B communities in all parts of the Arab world are strictly tribal. It has been noted by Arabists that most B tribes have a distinctive dialect of their own, and several studies have been devoted to examining the dialect of a particular tribe

(e.g., Johnstone, 1964, 1967b; Al-Hazmy, 1975; Ingham, 1979, 1982b, 1986; Al-Qahtani, 2015).

It should be reiterated here that, in this study, the factor of tribe is only applicable to Bs – not because Ss are not affiliated with tribes of their own, but rather because the concept of tribal affiliation is more valued and magnified in B communities.

All B tribes examined in the current study retain some linguistic features exclusive to that tribe only (see section 8.3.6), but they are not the focus of this research. Tribe as a factor, here, is expected to explain linguistic variation and patterns between B tribes across the examined linguistic variables, which they all share. That is, all B tribes are expected to use /u/ (instead of /i/), but if there is a convergence pattern on /i/, will members of tribe X show more convergence than tribe Y? I believe the answer is: extremely likely, due to historical, geographical and identity-related reasons, all of which are discussed below.

3.5.3.1 History of tribes

As stated in section (Chapter 1), current B tribes can be easily traced back to their ancestral tribes, due to Arabs' interest in genealogy. As these tribes branched into smaller groups, each one started to form a new clan and separate itself from the bigger tribe, though still acknowledging their affiliation. For example, although the B tribes of Ajman and Qahtan are originally from the Yemenite tribe of Hamadan, they have developed a separate identity, including the dialect. It is therefore crucial to examine the genealogy of B tribes as it might help track linguistic changes (as we will see in Chapter 6).

Another crucial cornerstone in the history of a tribe is the time when it moved to the region in which it has been in contact with another group, and therefore its dialect might be affected. In this study the groups in focus are B tribes that are originally from the south-western regions, and the time when they moved to Najd is essential in accounting for discrepancies in the rate

of convergence, if there is any. The Ajman tribe for example are considered late migrants to Najd compared to the Dawasir; therefore, if there is convergence to the S dialect, the former is expected to show less convergence than the latter.

3.5.3.2 Geographical location of tribes

Although B tribes share a set of linguistic features, they are surprisingly known to inhabit different regions. Assigning a location for a nomadic group seems to defy the core principle of the lifestyle; however, Arab genealogists noted that B tribes tend to move and settle within an area, to which they use the term *manāzil* (literally 'houses'). The original *manāzil* of each B and WB tribe have been recorded (Al-Hugail, 1967; Oppenheim, 2004) and the approximate radius is given in Figure 3.3. It should be noted that *manāzil* in this context is strictly used to refer to the period before the urbanisation and not the city/town in which they settled after the oil boom. It is believed here that because the tribes were to some extent geographically distant from each other, their linguistic behaviour might vary.
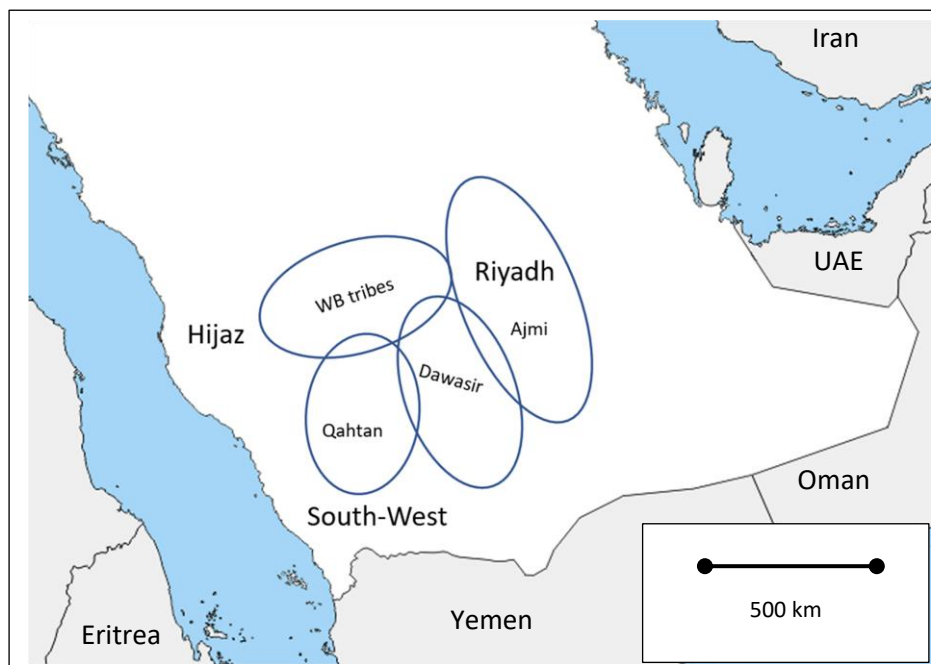
Figure 3.3: *The manāzil of B and WB tribes in Najd before the oil boom (Al-Hugail, 1967; Oppenheim, 2004).*

3.5.3.3 Tribal identity

In his *Muqaddima* (1377), Ibn Khaldun, a leading Arab historian, describes the social structure of the B communities and asserts that tribal solidarity built on blood ties is the main factor of identity (Retsö, 2006, p. 130). In tribal communities/countries, mostly in the Arabian Peninsula, "tribal affiliation constituted a crucial factor in the allocation of political, economic, and social power among existing Arab tribes" (Husam ad-Din, 1990; Smith 1990: cited in Farghal, 2007, p. 583). In Najd, as well as in other tribal communities in Saudi Arabia, individuals affiliate themselves with their tribe rather their original place of residence, or sometimes country.[16]

Tribal solidarity is displayed in several ways, some of which involve the dialect. Each of the B tribes examined here has a distinctive dialect, mainly manifested in phonological/morpho-phonological features and a set of lexical items (see section 8.3.6). Although B tribes share the examined linguistic variables, it is hypothesised that members of each tribe will exhibit similar behaviour towards any linguistic change. In other words, a linguistic feature might be adopted by some tribe members and then spread across the group, or get rejected, depending on the status of those members within the tribe. In either scenario, a tribe is likely to exhibit similar linguistic behaviour towards any linguistic change.

To sum up, we are not sure whether tribe as factor will show differences in the convergence rate, if there is any. It is, however, highly likely that differences will be detected, since each tribe forms a close-knit group of its own.

---

[16] Some groups of B tribes have migrated from Saudi Arabia to other Gulf countries, and they affiliate themselves with the tribe before country.

3.5.4 Education

Education as a social factor in most parts of the Arab world works as a 'proxy' factor, mainly by altering the individuals' social network and thus increasing contact with speakers of the target features (Al-Wer, 2002a). For example, in Cairo, education has been used by Haeri (1991) to determine the socio-economic class of speakers. She used a social-class index of four indicators: (1) parents' occupation, (2) speaker's education, (3) speaker's neighbourhood, and (4) speaker's occupation. However, the education of the parents of the speaker plays a role in determining the parents' occupation and neighbourhood (Sadiq, 2016). Another example is the study of rural speakers in Rades by Jabeur (1987), where it is found that the factor of social integration, the most significant factor, corelates with age, education level, and percentage of lifetime stay in Rades, all of which are determined one way or another by the level of education. In other words, education is accessed by younger speakers (age), who then migrate to other cities during school years and possibly settle in the same city after being employed (percentage of lifetime stay in Rades).

In basic education years in Najd, school can be a place of contact with other groups depending on the neighbourhood of the school. In some areas, for example, the majority of residents are Bs, all of whom attend the same school. In that case, the pupils usually build relationships with members of the same group inside the school and interact with them outside school as well, forming a dense multiplex social network. In other schools, attended by relatively even number of both groups, B and S, linguistic change is then determined by other factors like the speaker's social network and attitudes.

It is higher education that can significantly change the linguistic behaviour of Najdi individuals because it "involves significant alterations to individuals' socialisation patterns. It involves leaving one's home town, changes in familial links, expansion in social contacts,

interaction with speakers of other dialects, exposure to different social values, shifting of one's loyalties and attachments to various social groups, changes in priorities and ambitions." (Al-Wer, 2002a, p.43). This is true for all Bs and Ss in the suburbs of Riyadh who decide to pursue their university education. Until 2008, there were only two major universities in Riyadh Province (located in Riyadh city) which Najdis of the second generation attended (i.e., those who were born in the city to parents who used to live in the desert, B, or villages, S, before moving to the city). After 2008, the government commenced huge investment in higher education by building universities and affiliated colleges in almost every city/town of Saudi Arabia. Put differently, the second generation of Bs (aged between 30-55), who lived in the suburbs of Riyadh, are expected to show different linguistic behaviours according to whether they have a university degree or not.

3.5.5 Social network and the factor of 'contact with the other group'

3.5.5.1 Review of the social network approach

Developed by the anthropologist Elizabeth Bott (1971) and pioneered by Milroy (1980) in her Belfast study, the social network approach to linguistic change suggests that speakers with more closed networks, i.e., with a large number of local ties and contacts, tend to exhibit fewer signs of linguistic change than those with more open networks, i.e., fewer local ties and more contact with speakers of other dialects (Milroy and Gordon, 2003, p. 118). That is, if the ties of a social network are dense (most members know each other) or dense-multiplex (members interact in more than one social context, e.g., work and club), then members of such networks are more likely to adhere to the linguistic norms of the group resulting in language maintenance; and conversely members of open networks (with weak ties) are more susceptible to linguistic change (Milroy and Gordon, 2003, p. 118).

Although the notion of social network had been introduced earlier by Blom and Gumperz (1972), Milroy (1980) was the first to operationalise it by using what she called the Network Strength Scale. The scale "assessed speakers' network characteristics with reference to various relationships within the neighbourhood of kin, work, and friendship that had emerged in the course of the fieldwork as significant to participants. Speakers scored one point for each of the following conditions they satisfied:

  - were members of a high-density, territorially based group (e.g., a bingo or card-playing group, a gang or a football team, or football supporters' club)

  - had kinship ties with more than two households in the neighbourhood

  - worked in the same place as at least two others from the neighbourhood

  - worked in the same place as at least two others of the same sex from the neighbourhood

  - associated voluntarily with workmates in leisure hours." (Milroy and Gordon, 2003, p.121)

It should be noted that in the social network approach, the focus is placed on first-order ties, i.e., direct contacts, as opposed to second-order ties, i.e., indirect contacts.

In Arabic sociolinguistics, the social network approach was adopted as early as 1987 by Jabeur. In his analysis of rural speakers in Rades, he devised an adaptation of the Network Strength Scale to measure the social integration of rural migrants in the urban community. Each speaker is given a score of 0-3 in five bands, with a total range of 0-15; and the higher the score the more integrated a rural speaker is into the urban community. The five bands are:

  - Social contacts with rural district and length of visits

  - Social contacts with (rural) relatives in the community of Rades

  - Social contacts with neighbours from the same rural district

  - Social contacts at work or at school

  - Social contacts outside work (excluding neighbours and relatives) (Jabeur, 1987, p.98)

Al-Essa (2008), studying the dialect of Najdi Qassimi migrants in Jeddah (Hijaz), adopted the approach but focused on measuring the level of contact of Najdi migrants with the host Hijazi community. The contact index consisted of four criteria, and each speaker is given a score of 1 for each criterion he/she fulfils or 0 if not. The total score is then calculated, and the levels of contact are assigned as the following: 0-1 indicate 'low contact' and 2-4 'mid-high contact'. The four criteria are:

- Formal relationships at school and work, or marketplace.

- Participation in the neighbourhood affairs.

- Close friendships with Hijazi locals

- Kinship and mixed marriage with Hijazis in the family. (Al-Essa, p.69)

When comparing the two scales devised by Jabeur and Al-Essa, it is noted that Jabeur placed some weight on 'frequency of visits to the original community'. Although Al-Essa is also studying migrants who live in a different community in which they are considered a minority, she excluded this criterion probably because of the long distance between the new settlement and the hometown, approximately 900 km. The situation in Najd, specifically in the suburbs, is very different for the two target groups, and therefore the contact index in this study is slightly different, as is discussed in the following section.

3.5.5.2 The social network factor in this study

The structure of social networks in Najd, especially in the suburbs, is built on either lifestyle (B or S), or tribe. In B communities, which are tribal, social network is mostly dense and/or multiplex. It can be as closed as including members of the bigger family and relatives only, less closed involving members of the same tribes, even less closed involving Bs of other tribes, or open involving members of other groups (mainly S). In some cities, less commonly in Riyadh, some neighbourhoods are entirely inhabited by a single B tribe, sometimes even

named after it, forming a social network that fulfils almost all their everyday social needs. Similarly, this is true for the S groups, although their network usually consists of Ss from other tribes and not entirely from the same tribe, due to the tribal structure in S communities (see section 1.3.4).

In the current study, it is expected that Bs who have open networks, hence interacting with Ss, are likely to exhibit linguistic change in their speech more than those who interact with only Bs, whether relatives, Bs from the same tribe, or Bs from other tribes. It should be noted that in this study a B's social network is classified as open only if the first-order network ties are both Bs and Ss, or mostly Ss. This hypothesis is not only based on social network's enforcement of linguistic norms but also on contact with the other group, which then results in language accommodation (Giles, 1973; Trudgill, 1986). The social network approach overlaps with contact/accommodation as lack of contact favours lack of change (Trudgill, 2002, p.709). More discussion is provided in Chapter 8.

Since in Najd, and most parts of Saudi Arabia, working days for both schools and most government sectors end at about 1-2 pm, and due to the nature of weather being hot during the day and pleasant in the evening, people have more time to spend with friends even on working days. Therefore, the first-order ties are usually colleagues and close friends, with whom individuals interact most of the day. In some cases, colleagues are close friends, meeting after school/work. In most cases where they are not, colleagues are considered members of one's social network due not only to the amount of time spent with them at work, but because most Najdis are usually involved with their colleagues, arranging regular gathering one or more times each week.

One characteristic of the Najdi community that must be taken into consideration when examining the social network is that both groups are not considered migrants per se, but rather

people who have shifted lifestyle. Bs dwelt in the deserts of Najd while Ss lived in villages that expanded into the larger city in which they live. Yet, neither group is considered to be migrants, nor do they consider themselves as such. Therefore, the frequency of visits to the hometown is not relevant in this study[17].

3.5.5.3 The contact index in this study

In this study, the social networks of Al-Kharj will be analysed using an adaptation of the Network Strength Scale (Milroy, 1980). The aim is to examine the frequency of contact with the other group which might result in change through language accommodation (Giles, 1973; Trudgill, 1986). Milroy and Gordon (2003, p. 117) state that there is no canonically correct method to analyse social network, and based on the discussion above, it can be stated that the structure of social networks in Najd is different from the other social networks that have been examined. When analysing the social networks of Najd, we will focus on first-order ties in and outside school/work. As stated in the previous section, Najdis spend most of their time with their colleagues (with whom the relationship is more than just school/work) or with their friends outside school/work, if they are different. The following contact index was devised to measure the social contact with members of the other group inside and outside school/work:[18]

- Social contacts at school/work:

    - Who do you spend time with at school/work?

    - What is his/their tribe(s)?

---

[17] Even though some tribes, such as Dawasir, have a hometown in which they are the majority, they are not considered migrants in other cities of Najd because they tend to huddle in neighbourhoods where others from the same tribe have settled. Furthermore, some of them have settled in these cities for quite a long period of time to the extent that all members of their bigger family and relatives have settled in the new town and consequently they have all lost contact with their original place.

[18] It should be noted that other questions were included in the index, but they turned out to be irrelevant or not applicable. For example, the question 'who lives in your neighbourhood? And do you interact with them?' yielded useless results because the neighbours were either the same school/outside-school friends for most young speakers, relatives and family members with whom they interact most of the day, or members of the other group with which they do not interact.

- Is/are he/they the same friend(s) outside school/work? (If yes, the second group of questions is omitted.)

- Social contacts outside school/work:

    - Is it one or more person?

    - What is his/their tribe(s)?

    - How many hours do you spend with him/them?

A speaker's contact with the other group is considered high if he interacts with Ss more than Bs in and outside school/work. However, given the nature of social networks in Najd, especially that of the Bs (who are the focus of this research), the network of two thirds of the sample scored low, as the majority of their social contacts (in and outside school/work) were either relatives, Bs from the same tribe, or Bs from other tribes.

3.6 Sampling and the Sample

One of the key factors to the success of any variationist study is selecting a representative sample of the target community. *Random sampling*, pioneered in sociolinguistics by Labov (1966), was designed to overcome the problem of representativeness by giving speakers an equal and unbiased opportunity to represent a group (Milroy and Gordon, 2003, p. 25). However, this method was found to be biased in other ways, which caused some commentators to question its main advantage (i.e., representativeness). For example, Labov used telephone directories, which excludes all speakers under voting age, those who are not registered and of course those who do not have phones. Therefore, random sampling has generally been abandoned in favour of *judgment sampling*, in which the researcher selectively constructs a representative sample based on a set of predetermined criteria (Milroy and Gordon, 2003, p. 30). These criteria are not fixed, but rather vary according to the objectives of the research and the social structure of the target community. For example, Romaine

(1978), who was investigating the language of working-class speakers in Edinburgh, drew her sample from schools in areas that were ranked low in a census based on demographic characteristics such as housing, parents' education, health and employment.

3.6.1 Representativeness criteria in the current study

In order to for the sample to be representative of Najdi B or S dialect, the following criteria were set, based on my observations and knowledge of the Najdi community. All speakers in the sample met all three criteria.

3.6.1.1 Being a B or S

One of the fundamental questions to the current study is how can we determine if someone is B or S. Miller (2007) clarifies that "[r]eference to 'origin', 'lineage' and 'family' is a corner-stone of self-affiliation discourses…. Many urban dwellers categorize themselves by referring to a tribal – regional – family affiliation, rather than a contemporary place of residence" (p. 6) Previously in section 1.3.4, it was established that tribal affiliation is essential to answering this question since Najdi tribes can be grouped by lifestyle, B or S, as recorded by genealogists. Although Najdis nowadays do not pursue a B or S (farmer) lifestyle, it is what lifestyle their ancestors had led that is of interest here. Put differently, a Najdi from the Dawasir tribe, for example, who currently lives in the city, is classified as B because his ancestors (of the same tribe) have been known to lead a B lifestyle. The same applies to S tribes, meaning that Ss do not pursue a S farmer lifestyle, but they are classified as such because their ancestors had led that type of lifestyle.

This means that a speaker is classified as B if his last name, equal to the name of his tribe, indicates one of the tribes that were known to have led a B lifestyle; and the same applies to Ss. "Linguistically, the categories of bedouin, rural and urban dialects are still used by most linguists on the basis of the presence/absence of a set of features rather than by reference to a

geographical region or lifestyle. Categorizing a dialect X as a bedouin/ bedouinized dialect does not mean that the speakers pursue a nomadic bedouin way of life but that they display in their speech a number of features associated with bedouin dialects." (Miller, 2007, p. 6)

3.6.1.2 Born and raised in Najd

Although Najdis, especially Bs, affiliate themselves with the tribe rather than their current place of residence, which helps preserve their dialect, it is essential here to maintain that the recruited speakers were born and raised in Najd. The main reason for this is to avoid any influence of other, non-Najdi dialects. If a B lives, or lived part of his life, outside Najd, he might either preserve his dialect as a result of tribal affiliation, or he will converge to the host dialect. In either case, the outcome of both scenarios is not the focus of this research.

Those who were born and raised in Najd but whose ancestors are originally from another region in Saudi Arabia cannot be representative of Najd. This is because the speech of these individuals can: (a) resemble to a great extent the Najdi dialect, (b) their original dialect, or (c) something in between, depending on their social network and their connection to their place of origin. For instance, if someone from the south-western regions of Saudi Arabia lives in Najd, but his social network is confined to those from the same region, and he frequently visits his hometown, his speech is more likely to resemble his original dialect. Therefore, such speakers are excluded because the effect of their original dialect is difficult to measure. In addition, migrants from other Saudi regions have moved to Najd recently, and in most cases they maintain an attachment to their place of origin.

3.6.1.3 Not son of mixed marriage

As mentioned earlier in this chapter, during the interview the researcher detected differences in the speech pattern of some speakers and asked them, politely, whether their mother was from the same group or not. Their answer was always 'no, she is not from the same group',

and as a result they were excluded. The reason they were excluded was that their dialect was a consistent mix of B and S features. For example, all Bs had traces of /u/ as in *kull* 'all', and no one was found to have completely lost it. However, a speaker whose last name indicates a B tribe (after the father's) was found to use /i/ as in *kill* consistently; when I enquired about his mother, he informed me he was in fact the son of a mixed marriage. Conversely, one S speaker used *kull*, which is a B variant and no S produced any token of this sort; his mother was found to be B.

To sum up, the data of such speakers was excluded because they behaved radically differently from those who are born to two parents from the same group. If, hypothetically, more of these speakers had been interviewed and their data was included, we might misguidedly have arrived at the conclusion that some Bs have fully converged on /i/ and that some Ss partially converge to the B dialect.

3.6.2 The sample city: Al-Kharj

Riyadh province, representing central Najd, encompasses around 18 towns/cities, one of which is Al-Kharj, a medium-sized city located about 100 km south of Riyadh (Figure 3.4). The sample has been drawn from Al-Kharj for both theoretical and practical reasons.

Before the oil boom, Al-Kharj was basically an oasis called Yamamah, nowadays an old neighbourhood in the city. The population of the city increased steadily after the foundation of the Military Industries Corporation in 1953, along with other military bases. The city (i.e., the corporation) attracted Ss who lived in nearby villages and Bs who lived in the Najdi desert because the corporation offered all types of jobs that matched all the educational levels/skills of the local population, ranging from engineers and technicians to mechanics and drivers. Because Al-Kharj's population is associated with the foundation of the corporation, and those who took these jobs were local Ss and Bs, it is not surprising that the majority of the

99

population is comprised of these two groups. For that reason, Al-Kharj is the optimal city for

this research because:[19]

- Neither of the two groups is considered a minority in the city, while in other cities like

    Ad-Dilam or Hawtat-bani-Tamim the majority are Ss.

- The expected direction of convergence is either B > S or S > B, compared to Riyadh

    where there are other Saudi dialects (e.g., Hijazi, South-western) and other Arabic

    dialects (e.g., Egyptian, Syrian, Jordanian).
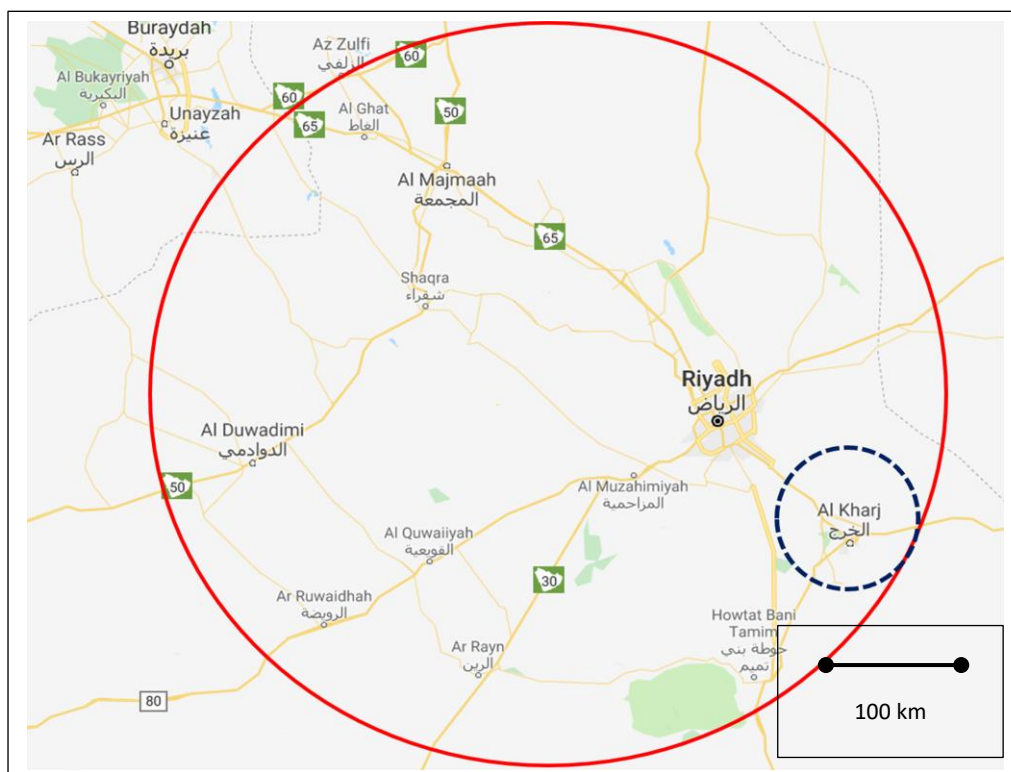
Figure 3.4: *A map featuring most of Riyadh Province and Al-Kharj city (from Google Maps).*

Practically, it was easier and more time saving for me to recruit speakers from Al-Kharj

because friends and close friends of the researcher's relatives, who comprise the sample, are

based in Al-Kharj. Also, because of the nature of the relationship between the moderators and

speakers (friends and close friends), the researcher was able to obtain ethnographic

---

[19] All of the information in this paragraph is based on the commentary of old interviewees who have been working in the Military Industries Corporation for more than three decades, and have lived through these stages of the city.

information about the speakers before meeting them and thereby to help balance the sample according to social factors. The last practical reason is that the lifestyle in Al-Kharj is slow-paced, compared to cities like Riyadh, and therefore people in Al-Kharj are usually free and willing to meet other people.

3.6.3 The sample

The sample in this study is comprised of 23 B speakers, who are the focus of this study, 13 S speakers, who could be considered the control group, and 3 WBs (Bs from the mid-western parts of the Arabian Peninsula), who are included for the sake of comparison with the Bs. To elaborate, the Bs are the group that is expected to show change (convergence), and thus their number is larger than the other groups. The Ss, on the other hand, are expected to show no variation, but due to scarcity of the literature on Najd sociolinguistics, the linguistic and sociolinguistic traits of the dialect must be established. Therefore, only 13 S speakers were interviewed. Lastly, the WB tribes, constituting a different group, have been observed to behave differently from the B group across the examined variables, i.e., they behave similarly to the Ss in one variable; and since they are marginal to the study, only 3 speakers, one from each tribe, have been interviewed.

For the B group, all efforts have been made to balance the sample across the social factors of tribe, age, contact and education. However, while it is easy to do so for the factors of age and tribe, it was difficult or almost impossible to achieve for factors like contact and education. The tribe and age of a B speaker is determined before recruitment, and the researcher was able to adjust the numbers of speakers per slot during the recruitment process. For the factor of education, it is almost impossible to find young speakers with a low level of education because basic education in Saudi Arabia is compulsory. As for the factor of contact, it is

difficult to predict the level of speaker's contact in advance because it is measured after meeting the participant via filling in the contact index (see section 3.5.5.3).

As for the Ss, there were no factors, other than age, that were expected to affect the outcome. Later after data analysis, it was found that some S speakers switched to Standard Arabic during the interview, hence the factor of style (formal and informal) was added. Formal style indicates those speakers who switched to Standard Arabic, while informal style indicates those who used the spoken dialect throughout the entire interview.

Table 3.1: *Distribution of the sample by social factors*

Group	Social factors	Levels	Number of speakers	Total for each group	Total number of speakers
B	Tribe	Ajmi	8	23	39
		Dosari	9		
		Qahtani	6		
	Age	Old	11		
		Young	12		
	Education	High	6		
		Low	17		
	Contact	High	5		
		Low	18		
WB	Tribe	Otaibi	1	3	
		Sihali	1		
		Subai	1		
S	Age	Old	6	13	
		Young	7		
	Style	Formal	4		
		Informal	9		

3.7 Data Analysis

The quantitative approach in variationist sociolinguistics, pioneered by Labov in the 1970s, is based on: (a) quantifying all the occurrences as well as the zero variant of a variable,

Principle of Accountability (Labov1972), and (b) analysing the data with the help of available suitable statistical tools. The main goal of this approach is to help us understand what series of factors (social or linguistic) are responsible for the behaviour of a dependent variable. The core principles of the quantitative approach have not changed over the past 40 years, but the statistical tools and models have evolved and developed significantly (Tagliamonte, 2011).

In the current study, the quantitative approach is adopted, with the goal of determining which group (B or S) is converging towards the other and what social factors are conditioning this convergence. The data for each group is analysed separately because the social factors for each group are entirely different. The premise of quantitative analysis here is that if a group is producing the variant(s) associated with the other, they are converging to that group; and by using statistical tools we can determine which sub-group, e.g., particular age group, is leading the change. For instance, the merger variable has two variants (/u/ and /i/). The null hypothesis is that Bs use /u/ while Ss use /i/ all the time. If for example some or all Bs are found to produce both variants, as expected, the statistical analysis of data will probably show which sub-group of Bs (e.g., tribe, age group) is leading the convergence on /i/.

The emergence of intermediate forms between the variants is expected, and they were indeed recorded during the data transcription. However, the frequency of these forms was low, and therefore they were excluded.

In the next sections, I shall discuss data transcription and statistical analysis.

3.7.1 Data transcription

The data in this study was transcribed auditorily. Because of (a) my intuition as a native speaker, (b) the nature of the phonological contrast between the variants and (c) the high quality of the audio, distinguishing whether the variants were B or S was quite an easy task.

Since we are interested in three variables only, transcribing the entire interview would be time consuming and pointless. Therefore, only the word in which one of the variables occurred is transcribed. A separate Excel spreadsheet, listing all occurrences of three variables chronologically, was created for each speaker. After that, a separate spreadsheet for each variable was created for each group, making a total of 6 spreadsheets. Each spreadsheet is organised in the following way:

- The columns indicate the speaker's number, levels of the social factors for that speaker, the value of the dependent variable (B or S) and the word in which the token occurred.

- The rows indicate every occurrence of the variable (see sample in Table 3.2).

The recorded speech (approximately 30 hours) yielded a sufficient number of tokens to fit into the statistical model of choice (discussed in the next section); the number of tokens per spreadsheet is provided in Table 3.3. However, the fitness of the model is subjected to the B/S token ratio in each data set. That is, in some data sets the number of tokens of the target feature might be extremely low to the extent that a statistical model cannot be fitted.

Table 3.2: *Sample of the spreadsheet for the merger for Bs (DV = dependent variable)*

Group	Participant	Tribe	Age	Edu.	Contact	DV	*Word*	Gloss
B	speaker1	A	O	Low	L	B	*k-u-llaha*	'all of it'
B	speaker1	A	O	Low	L	B	*ya:x-u-ðˤ*	'he takes'
B	speaker1	A	O	Low	L	B	*k-u-llin*	'all'
B	speaker2	A	Y	High	L	B	*yarg-u-d*	'he sleeps'
B	speaker2	A	Y	High	L	B	*g-u-mna*	'we stood up'
B	speaker2	A	Y	High	L	S	*k-i-llaha*	'all of it'
B	speaker3	D	Y	High	L	B	*ʔadx-u-l*	'I enter'
B	speaker3	D	Y	High	L	B	*ʔagʕ-u-d*	'I sit'
B	speaker3	D	Y	High	L	S	*tadx-i-l*	'she enters'

Table 3.3: *Number of tokens for each data set (spreadsheet)*

The linguistic variable	B group	S group
[u] / [i]	1264	506
-ih / -ah	1068	572
ya- / yi-	761	414
Total number of tokens	**4585**	

3.7.2 Statistical analysis

3.7.2.1 Evolution of statistical tools in variationist sociolinguistics

In the field of language variation and change, statistical tools are used to assess the relationship between a dependent variable and a list of one or more independent variables. The dependent variable represents the target linguistic feature; while the independent variables represent the social/linguistic factors (also referred to as *predictors* or *explanatory variables*) constraining the production of the dependent variable. The statistical tool used for this purpose is logistic regression; and the software program that had been used by variationists is *Varbrul* (Cedergren and Sankoff 1974: cited in Tagliamonte, 2011, p.121). The toolkit used was *Goldvarb* (Sankoff et al., 2005).

Goldvarb dominated variationist sociolinguistics for over 30 years, but in the early 2000s, it received criticism as to whether it is the most appropriate statistical tool for the field (Tagliamonte, 2011, p.129). A better statistical tool, *generalised linear mixed-effects modelling* (GLMM), was introduced because it overcomes the main drawback of the previous model, discussed in the next section.

3.7.2.2 Fixed vs. mixed logistic regression

Goldvarb uses standard logistic regression which analyses the influence of multiple (fixed) factors on the production of a binary linguistic variable (Tagliamonte, 2013). For example, if we are examining the production of /i/ or /u/ in the speech of B speakers grouped by age and education, *Goldvarb* will enable us to examine how the fixed factors of age and education can influence the likelihood of the speaker producing either /i/ or /u/. This model is referred to as *fixed logistic regression*, because it includes fixed factors only. However, this poses a major problem with such models when applied to linguistic data. "As it is usually run, without a factor group for speaker, GoldVarb [fixed logistic regression] necessarily ignores the grouping and treats each token as if it were an independent observation. This leads the program to overestimate – potentially drastically – the significance of external effects, those of social factors like gender and age." (Johnson, 2009, p. 363). Also, Tagliamonte (2011) points out that fixed models can "create inappropriate aggregations, especially with unbalanced numbers of tokens across individuals (as is typically the case)" (p. 130). To put it differently, the crucial problem of fixed logistic regression in variationist sociolinguistics is its exclusion of the random effect of individuals and items (words). For instance, if an individual produced tokens at a higher rate than his/her sub-group, this may alter the data distribution for that sub-group, resulting in less or exaggerated significance. Items have a random effect as well, because some words are more frequent than others, and some speakers may use some words more frequently than other speakers.

Mixed logistic regression (or *mixed-effect logistic regression*) has been introduced to the field as it integrates both fixed and random factors, hence the name 'mixed' (Tagliamonte, 2011, p.130). Mixed-effects models offer a flexible and powerful tool for analysing grouped data (e.g., longitudinal data, repeated measures); the flexibility of such models can handle within-group correlation and balanced/unbalanced data, all of which are often present in grouped

data (Pinheiro and Bates, 2000, p. vii). There are several statistical packages that offer Mixed logistic regression, but *lme4* package (Bates, Maechler, Bolker, and Walker 2015) for *R* (R Core Team 2018) is the most commonly used, at least in variationist studies.

It should be noted that there are two types of mixed logistic regression: linear mixed models (LMM) and generalised linear mixed models (GLMMs). The former is suitable for modelling continuous responses while the latter is applicable to binary responses. As responses in the present study are binary, GLMM is the model of choice. The function that performs GLMM in *lme4* package is *glmer*.

3.7.2.3 Using and coding glmer in R

Applying *glmer* requires fixed and random factors. Random effects can be fitted in the model by either random-intercept or random-slope. Random intercept is used with factors that are between-speakers (i.e., factors that are different across speakers). That is, for a social factor that has two levels, speakers are assigned one level only, which is repeated in the data set with all occurrences of a variant produced by that speaker. A speaker cannot be `young` in one occurrence and `old` in the other. By contrast, factors like `style` vary within speakers, meaning that a speaker can produce the same (or different) words once with a `careful` style and once with a `casual` one. In this case a random slope is needed.[20] In this study, only random intercept is used because all social factors are between-speakers.

The formula code for *glmer*, used in this study, is as follows:

```
Model.X <- glmer (dependent variable ~ factor.1 + factor.2 +
factor.3 +   (1 | random factor.A) + (1 | random factor.B),
data = Y, family =binomial)21
```

---

[20] See Barr, Levy, Scheepers & Tily (2013) and Sadiq (2016, p.65) for more details.
[21] Based on Bates, Maechler, Bolker, and Walker (2015)

`Model.X` represents the name of the model and it can be filled by any name. `data = Y` indicates the name of the data set in R. `(1 | random factor.A)` represents a random intercept. `family = binomial` indicates the regression family that is used when the dependent variable (response) is binary.

3.7.2.4 Model fitness

Depending on data distribution, the number of levels for each factor and many other mathematical issues, a model may or may not converge (i.e., yield results). When it fits (converges), a measure called *Somers2 Dxy* is applied to the model in order to test the 'goodness of fit' (Baayen, 2008, p. 305: cited in Tagliamonte, 2011, p. 149). After applying this test to a model, the value under *C* indicates the level of fitness, with higher being better (see Example 3.1 below). The C value in the example (`0.959882`) indicates a high level of fit because it is close to 1.

Example 3.1: (taken from the present study)

```
somers2(probs, as.numeric(Bs.u_i$DV)-1)
```

C	Dxy	n	Missing
0.959882	0.919764	1264.000000	0.000000

3.7.2.5 How to interpret results from glmer

After applying *glmer*, the results are retrieved by requesting a summary of the model, using the code: `summary (Model.X)`. A sample of a summary model is provided in Example 3.2, and the following commentary on the sample is from Tagliamonte (2011, p. 148)

Example 3.2: (Adapted from the present study)

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.0473	0.8512	-3.580	0.000344 ***
Factor.1.Level.B	2.9826	0.7039	4.237	2.26e-05 ***
Factor.1.Level.C	1.6572	0.8019	2.067	0.038774 *
Factor.2.Level.B	-0.1979	0.7525	-0.263	0.792530
Factor.3.Level.B	-2.1721	0.7363	-2.950	0.003175 **

The results in the example above should be read as follows:

- The first column lists the fixed factors:levels. The names of the factors:levels in the examples are not random, but rather they are ordered to illustrate how *glmer* works. Note that `level.A` of each factor is not shown in the summary above. This is because *glmer* uses one level as baseline, giving it `0` value, and compares it to the other levels. The baseline level is chosen alphabetically.

- The second column (`Estimate`) shows the coefficient of each factor:level, estimated in log-odds. This means, if we take the value of `Factor.1.Level.B` as an example, it reads as: for every one unit change in `Factor.1.Level.B`, the log odds of `convergence` (for example) increases by 2.9826, compared to the value of `Factor.1.Level.A` which is `0` (but not shown). If the value is negative, as in `Factor.3.Level.B`, it means that the log-odds decreases.[22] In non-technical terms, and taking an example from variationist sociolinguistics, if a factor:level (let's say gender:male) has a positive significant value, we can say that male speakers are more likely to produce the target feature. Henceforth, the estimate value of coefficients will be referred to as only `estimate`.

- The third column, standard error (`Std.Error`), indicates sample fluctuation. That is, it shows how the `estimates` would fluctuate if the study was re-run with identical conditions.

- `Z value`, the third column, shows divergence in the results. This value is obtained by dividing the `estimate` by the standard error.

---

[22] For more discussion of log-odds in logistic regression, see Hosmer and Lemeshow (2000).

- The last column, `Pr(>|z|)`, shows the P value (i.e., statistical significance). The relative strength of each factor is measured in the significance code: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

In variationist sociolinguistics, we are mainly interested in "(1) statistical significance; (2) relative strength of factors, as measured here in the significance codes; and (3) constraint ranking, as measured here in the relative coefficient values [estimate]" (Tagliamonte, 2011, p. 148)

3.7.2.6 Statistical analysis in this study

In this study the social factors for Bs and Ss are different (see Table 3.1), but the random factors `speaker` and `item` apply to both. Using the code formula in section 3.7.2.3, the following *glmer* codes for each group have been generated:

- ```
  Bs.convergence.to.Ss <- glmer (dependent variable ~ tribe
  + age + education + contact + (1 | speaker) + (1 | item),
  data = Bs, family = binomial)
  ```

- ```
  Ss.convergence.to.Bs <- glmer (dependent variable ~ age +
  style +  (1 | speaker) + (1 | item), data = Ss, family =
  binomial)
  ```

It should be noted that these codes will work only if there is variation in the data set to which the model is applied. For example, the Ss code, provided above, is based on the assumption that Ss might converge to Bs.

The social factors are assigned the following levels:

- `Tribe: Ajman, Dawasir, Qahtan` (for Bs)
- `Age: young, old` (for both Bs and Ss)
- `Education: high, low` (for Bs)

- `Contact: high, low` (for Bs)
- `Style: formal, informal` (for Ss)

The levels of age, education and contact were originally three, but during experimentation

with *glmer* on all data sets, it was decided that it was necessary to reduce the levels to two in

order for the model to fit.

3.8 Conclusion

In this chapter we have reviewed the sociolinguistic interview based on two models: *attention*

to speech and *addressee/audience design*. After that the structure and procedures of the

sociolinguistic interview in the current study was discussed. The interview in this research

focuses on eliciting conversational styles only, via open-ended questions.

The sample is comprised of 23 Bs and 13 Ss. Bs can be sub-grouped by tribe, age, education

or contact, while Ss can be grouped by age or style. These sub-groups represent the social

factors in this study.

The data of both groups will be analysed using the most up-to-date statistical tool in

variationist sociolinguistics, the generalised linear mixed model (via *glmer* in R). This

statistical tool is capable of incorporating both fixed and random factors. In the current study,

as well as in most variationist studies, the random factors of *speaker* and *item* are included in

the analysis.

Chapter Four: The Merger (/u/, /i/ > /i/)

In this chapter, the vowel merger /u/, /i/ > /i/ in Najdi Arabic is subjected to a sociolinguistic analysis. It is hypothesised that the Ss have fully undergone the merger, while the Bs have retained the opposition between /u/ and /i/. However, after the heavy contact between the two groups (resulting from the massive urbanisation) and due to the scarcity of studies on this subject, the linguistic outcome of this contact has hitherto been unknown. The goal of this chapter is to decide which of the following scenarios applies to the merger in Najdi dialects:

a) The Ss have fully undergone the merger, while the Bs fully retain the opposition between /u/ and /i/.

b) The Ss have fully undergone the merger, while the Bs are converging to the S dialect and therefore are in the processes of losing the opposition.

c) Both groups have fully undergone the merger.

d) A less likely, but plausible, scenario is that the Ss have reversed the merger due to their exposure to the B dialect and more importantly to Standard Arabic, in which the opposition between the vowels is also retained.

In the first part of the chapter, an overview of phonological merger in general and the linguistic/social factors responsible for this merger in Arabic are provided. In the second part, a review of the short vowel system in modern dialects of Arabic is provided. In the last part, the data on this variable for both groups, Bs and Ss, and its statistical analysis are presented.

4.1 Phonological Merger in General: An Overview

Phonological merger can be defined as a reduction in the complexity of the sound system of a language, in which usually two distinct phonemes merge into one. They are common in

human languages, and as characterised in Labov (1994, p. 310), they are irreversible by linguistic means (Garde's Principle) and expand rapidly. The questions that raise themselves are: if the aftermath of mergers is loss of distinction, why and how do they happen in the first place? There are several attempts to identify the internal (linguistic) cause(s) of merger to help determine what makes some phonemes more susceptible than others. Labov (1994, p.328) lists the internal factors as the following:

- The functional load of the opposition.

- The number of distinctions already made along that phonetic dimension.

- The number of phonetic features on which the opposition depends.

- The discriminability of the phonetic features on which the opposition depends.

- Limitations in the range of movements that would avoid merger.

The answer to how mergers take place is provided in Trudgill and Foxcroft (1978), who provided two models for the mechanism of merger. The first model, *merger by approximation*, refers to when two phonemes gradually move to occupy one phonetic space. The other model, *merger by transfer*, refers to when words from one phonemic class are transferred to another as a result of dialect contact. Labov (1994, p. 327) categorises the first model as internally motivated while the second is externally motivated.

The merger /u/, /i/ > /i/ in Arabic is externally motivated (as we shall see below). However, there must be also internal factor(s) that cause these phonemes to be more prone to change than others. In the next section, the internal and social factors responsible for this merger are discussed.

4.2 The Merger (/u/, /i/ > /i/) in Arabic

The vowel system of Classical and Modern Standard Arabic is composed of three short vowels (/u/, /i/, /a/), three long vowels (/u:/, /i:/, /a:/), and two diphthongs (/ay/, /aw/). There

have been some changes to the vowel system of Classical Arabic, which are visible when compared to modern Arabic dialects, one of which is the monophthongisation of the diphthongs mentioned in section 1.5.5.4.3 above. Another change to the vowel system of most modern Arabic dialects is the reduction of the short vowel system (in the Najdi case /u/, /i/ > /i/). This change leaves the affected dialects with a two-short-vowel system, which is still effective because in Arabic the majority of lexical contrasts are indicated via consonants (Watson, 2002, p. 22). This merger, or reduction in general, is motivated by linguistic and sociolinguistic factors, which are discussed below.

4.2.1 Linguistic factors

The linguistic factor responsible for this merger is most likely the *functional load of the opposition*, a concept originally proposed by Martinet (1952). According to Labov (1994, p. 327) it can be defined as "the number of different lexical items that are differentiated by a phonemic opposition, and conversely, the number of homonyms that would result from its loss". Minimal pairs in which the opposing phonemes are short vowels are limited in Classical Arabic because "the major lexical contrasts in Arabic are indicated through the consonants. This is reflected in the Arabic script which is based on (mainly tri-consonantal) roots of consonants and glides, and which inserts short vowels when necessary as diacritics above and below the consonant" (Watson, 2002, p. 21).

Table 4.1: *Examples of minimal pairs distinguished by short vowels in Classical Arabic*

Word	Gloss	Word	Gloss
ʕ*ud*	'come back. Imperative'	ʕ*id*	'promise. Imperative'
q*ud*	'lead. Imperative'	q*ad*	'preverbal marker'
m*ud*da	'has been stretched'	m*ad*da	'he stretched'
m*an*	'who?'	m*in*	'from'

In Najdi Arabic, minimal pairs distinguished by short vowels have decreased (or even been completely lost) as a result of morphological changes and lexical replacement. An example of these morphological changes is the loss of the internal passive formation (1.5.5.4.5) in favour of affixation (though the form is retained by some old B speakers, according to data from the current study, see 8.3.1); and therefore, the forms *mudda* and *madda* from Table 4.1 above are realised in Najdi as *in-madd* and *madd*. As for lexical replacement, most Classical Arabic forms that are distinguished by short vowels are replaced with other forms in Najdi Arabic. For example, the Classical Arabic forms *ʕud* and *ʕid*, in Table 4.1, are replaced with *ʔirjaʕ* and *wa:ʕid* in Najdi Arabic.

4.2.2 Sociolinguistic factors

One of the accepted theories concerning the development of contemporary Arabic dialects by Ferguson (1959a) suggests that the current Arabic dialects did not develop from Classical Arabic directly, but rather there was an inter-dialectal Arabic koine from which they have evolved (discussed above in 2.2). This theory is supported by the fact that Modern Arabic dialects share a similar set of innovations, one of which is reduction of the opposition between /u/ and /i/ (Fischer & Jastrow, 1980, p. 39-48: cited in Versteegh, 1984). The emergence of the koine might be regarded as the starting point of the reduction of the opposition between /u/ and /i/ in some modern Arabic dialects, but the question 'why do some Arabic dialects retain the opposition between /u/ and /i/ in all phonological contexts?' requires further (historical) research that is beyond the scope of this thesis. In the next section, a synchronic survey of the short vowel system in modern Arabic dialects is provided.

4.3 The Merger in the Arab World

The three-short-vowel system of Classical Arabic tends to be reduced to a binary system in some modern Arabic dialects, or even in rare cases to a one-short-vowel system as in the

Jewish Arabic of Algiers and Jewish Arabic of Tripoli (Versteegh, 2001, p.87). The reduction of the short vowel system can be a result of: /u/, /i/ > /i/ as per the current variable; /u/, /i/ > /ə/; or /a/, /i/> /ə/. The dialects that seem to employ a two-short-vowel system are S dialects and dialects spoken in dense urban centres (Versteegh, 2001, p. 134). On the other hand, the majority of dialects that retained all short vowels are spoken in the Arabian Peninsula. The reduction in the short vowel system has received some attention in studies that are concerned with the history of old/new Arabic and the koineization of Arabic (Ferguson, 1959a; Fischer & Jastrow, 1980; Versteegh, 1984). The retention/reduction of Classical Arabic short vowels is also one of the linguistic features Versteegh (2001) used for the classification of modern Arabic dialects. It can be stated that this change received attention in historical and descriptive studies of Arabic, but not sociolinguistically. In the next sub-sections, we will present Versteegh's (2001) classification of modern dialects based on the reduction of short vowels.

4.3.1 Western Arabic

The dialects covered in this section are the East African (Egyptian/Sudanese) and Maghrebi (literally 'western') dialects.

4.3.1.1 East African dialects (Egyptian and Sudanese)

According to Wilmsen and Woidich (2007, p. 2) the dialects of Egypt are distinguished as follows: the dialects of the Delta, the dialect of Cairo, the Middle Egyptian dialects (from Gizeh to Asyūṭ) and the Upper Egyptian dialects (from Asyūṭ to the south). Versteegh (2001, p. 206) states that one of the traits that distinguish the Egyptian dialects from other neighbouring dialects is the preservation of all short vowels of Classical Arabic, though /u/ and /i/ are altogether elided in unstressed open syllables.

As for Arabic dialects in Sudan, they consist of one main central dialect and several other regional ones. There are also three other Arabic varieties spoken by people whose mother

tongue is not Arabic (Abu-Manga, 2009, p.375). In the present study, we will only review Khartoum Arabic as it is considered the standard spoken variety in Sudan (ibid., p. 376). In Khartoum Arabic, the opposition between /u/ and /i/ is retained in all phonological environments (Dickens, 2007, p. 560)

4.3.1.2 Maghreb dialects

The dialects of Maghreb belong to two major groups: the Sedentary pre-Hilāli dialects and the Bedouin Hilāli dialects, each of which belongs to a particular stage of Arabicisation. The first group is associated with the first phase of Arab immigration (7th-10th centuries C.E.) and the Islamic conquests, and it was spoken by those settlers in existing urban centres (Palva, 2006; Versteegh, 2001). The pre-Hilāli dialects are divided into:

- "The Eastern pre-Hilālī dialects, spoken in Libya, Tunisia and eastern Algeria; these dialects are characterised by the preservation of the three short vowels.
- The Western dialects of the pre-Hilālī group, spoken in western Algeria and Morocco; these have only two short vowels." (Versteegh, 2001, p. 211)

The other group, the Hilāli dialects, is associated with the immigration of the originally Najdi B tribes of Banu Sulaym and Banu Hilal and the southern Arabian tribe of the Ma'qil, all of whom occupied the plains of North Africa in the 11th century (Palva, 2006, p. 609). As with most B dialects around the Arab world, Hilāli dialects have retained all short vowels (Rosenhouse, 2006, p. 269).

Maghreb dialects that have reduced the short vowel system belong, then, to the Western pre-Hilāli group. Descriptive studies of the Arabic varieties spoken in urban centres of Algeria and Morocco provide more details of the binary vowel system of these dialects. The dialect of Algiers and Moroccan Arabic (spoken in most urban centres of Morocco) are reported to use

117

the short vowels /u/ and /ə/. The first vowel is retained from Classical Arabic while the latter resulted from the merger of /a/ and /i/ (Boucherit, 2006; Caubet, 2008).

4.3.2 The Levant

4.3.2.1 Syro-Lebanese dialects

This dialect group is characterised as S because of the realisation of the Classical Arabic /q/ as /ʔ/, e.g., *ʔalb* 'heart' (< *qalb*). Exceptions to this are the Jordanian dialects, which exhibit B features, one of which is the use of /g/ instead (*galb*). Versteegh (2001) distinguishes the following dialects within the Syro-Lebanese group:

- Lebanese and Central Syrian dialects, e.g., the dialect of Beirut and Damascus.

- North Syrian dialects, e.g., the dialect of Aleppo.

- Palestinian/Jordanian dialects.

The dialects of the capital cities of Beirut and Damascus employ a binary vowel system, but different from that in Algiers and Morocco. The short vowels are: /a/ and /ə/, the former being inherited from Classical Arabic while the latter is a result of the merger /u/, /i/ > /ə/. However, the opposition between /u/ and /i/ is retained in non-stressed open syllables (Lentin, 2006, p.547; Naïm, 2006, p.276). On the other hand, in the majority of Syrian dialects outside Damascus (including northern Syrian) and Palestinian/Jordanian dialects, the opposition between /u/ and /i/ is retained in all phonological contexts (Versteegh, 2001, p. 199; Behnstedt, 2009, p. 404; Al-Wer, 2007a, p. 508).

4.3.2.2 Mesopotamian dialects

Blanc (1964) reports that the dialects of Iraq can be classified based on religious affiliation: Muslims, Christians and Jews (Table 4.2). He also claims that the dialects of Muslims, including all Bs and Ss of lower Iraq, belong to one broader group while those of S Muslims

of Upper Iraq and Anatolia and non-Muslims belong to another. One of the linguistic features Blanc used for such a grouping is the realisation of the Classical Arabic /q/. Using the Classical Arabic form *qultu* 'I said', he labelled the dialects of Iraq as either a *gilit* dialect (/g/ < /q/) or a *qəltu* dialect.

Table 4.2: *The dialect groups of Iraq (Blanc, 1964, p. 6)*

| | Muslims | | non-Muslims |
	non-sedentary	sedentary	
Lower Iraq	*gilit*	*gilit*	*qəltu*
Upper Iraq	*gilit*	*qəltu*	*qəltu*
Anatolia	*gilit*	*qəltu*	*qəltu*

With regard to the short vowel system, the *qəltu* group resemble S dialects in terms of short vowels/dialect features, one of which is the merger /u/, /i/ > /ə/ (Jastrow, 1978, cited in Versteegh, 2001). As for Muslim dialects (*gilit* group), all three short vowels are retained, but the opposition between /u/ and /i/ is limited to specific phonetic environments. All occurrences of /u/ in Classical Arabic are realised in Muslim dialects as /i/ (e.g., *qultu > gilit*); however, /u/ replaces /i/ or even /a/ after labials, e.g., *busˤal < basˤal* 'onions', *ħa:muðˤ < ħa:miðˤ* 'sour' (Versteegh, 2001, p. 202). It can be summed up that *qəltu* dialects behave similarly to the other S dialects of the Levant (e.g., Beirut and Damascus), while in *gilit* dialects /u/ is retained partially.

4.3.3 The Arabian Peninsula

In this section the short vowel system in the dialects of the countries on the east to the south coast of the Arabian Peninsula and the dialects of Saudi Arabia will be covered.

4.3.3.1 The Arabian Gulf and south coast

In the dialects of Kuwait and Bahrain, all three short vowels are used, but the opposition between /u/ and /i/ is reduced to specific phonetic environments. /u/ replaces /i/ or /a/ when

sandwiched between a labial and a velar, emphatic, or liquid, e.g., *muka:n* 'place', *busˤal* 'onions', and *rubatˤ* 'he tied', but not *bitʃir* 'first-born child' or *simatʃ* 'fish'. The opposition is also retained in open syllables, but not in closed ones (Holes, 2006, p. 243; Holes, 2007, p. 610). As for the dialects of Oman and Sana'a in Yemen, all three short vowels are used, and the opposition between /u/ and /i/ is retained in all phonological contexts (Edzard, 2008, p. 479; Watson, 2009, p. 107).

4.3.3.2 Saudi Arabia

In Saudi Arabia, the reduction in the short vowel system, caused by the merger /u/, /i/ > /i/, has been noted in the cities of Hayil, Sudair, Riyadh and Hofuf, while in the remaining areas all short vowels are retained in all phonological contexts (Prochazka, 1988, p. 17; Ingham, 1971, p. 273). As per the researcher's observation, the reduction in the dialects of these cities can be extended to include the regions of each city. Hayil and Sudair represent Northern Najd; Riyadh represents S Central Najd; Hofuf represents the Eastern Province. The rest of the dialects which retained all three short vowels are those of B Central Najd, Hijaz and South-West region.

The statement by Ingham (2009, p. 123) that the dialects of Saudi Arabia form a continuum to those of the neighbouring countries is reinforced by the merger. For example, although the dialect of Hijaz is urban and S, it has retained the opposition between /u/ and /i/, as is the case in the dialects of Egypt with which they are reported to share several linguistic features (Ingham, 1971). As for the dialects of the Eastern Province, they share the loss of opposition between the vowels /u/ and /i/ with the dialects of Kuwait and Bahrain, but it is not reported whether the effect of labials reported in these countries (see previous section) exists in the Eastern dialects of Saudi Arabia. Lastly the dialects of the south-western regions of Saudi Arabia are similar to those of Yemen, in which the opposition between /u/ and /i/ is retained.

Table 4.3: *Short vowel system of modern Arabic dialects compared to Classical Arabic (Versteegh, 2001; Rosenhouse, 2006; Boucherit, 2006; Caubet, 2008; Lentin, 2006; Naïm, 2006; Behnstedt, 2009; Al-Wer, 2007a; Jastrow, 1978; Holes, 2006; Holes, 2007; Edzard, 2008; Watson, 2009; Prochazka, 1988; Ingham, 1971; Dickens, 2007)*

Part of the Arab World	Arabic Variety	Dialects		Classical Arabic	
			/u/	**/i/**	**/a/**
West		Algiers	/u/	/ə/	
		Morocco	/u/	/ə/	
		Libya	/u/	/i/	/a/
		Tunisia	/u/	/i/	/a/
		Egypt	/u/	/i/	/a/
		Khartoum	/u/	/i/	/a/
The Levant		Palestinian	/u/	/i/	/a/
	Lebanon	Beirut		/ə/	/a/
		South	/u/	/i/	/a/
	Syria	Damascus		/ə/	/a/
		B	/u/	/i/	/a/
	Iraq	Non-Muslim		/ə/	/a/
		Muslim	/u/ (partial)	/i/	/a/
The Arabian Peninsula		Kuwait	/u/ (partial)	/i/	/a/
		Bahrain	/u/ (partial)	/i/	/a/
		Oman	/u/	/i/	/a/
		Sana'a	/u/	/i/	/a/
	Saudi Arabia	Eastern		/i/	/a/
		South-western	/u/	/i/	/a/
		Hijaz	/u/	/i/	/a/
		Northern Najd		/i/	/a/
		Central Najd S		/i/	/a/
		Najd B	/u/	/i/	/a/
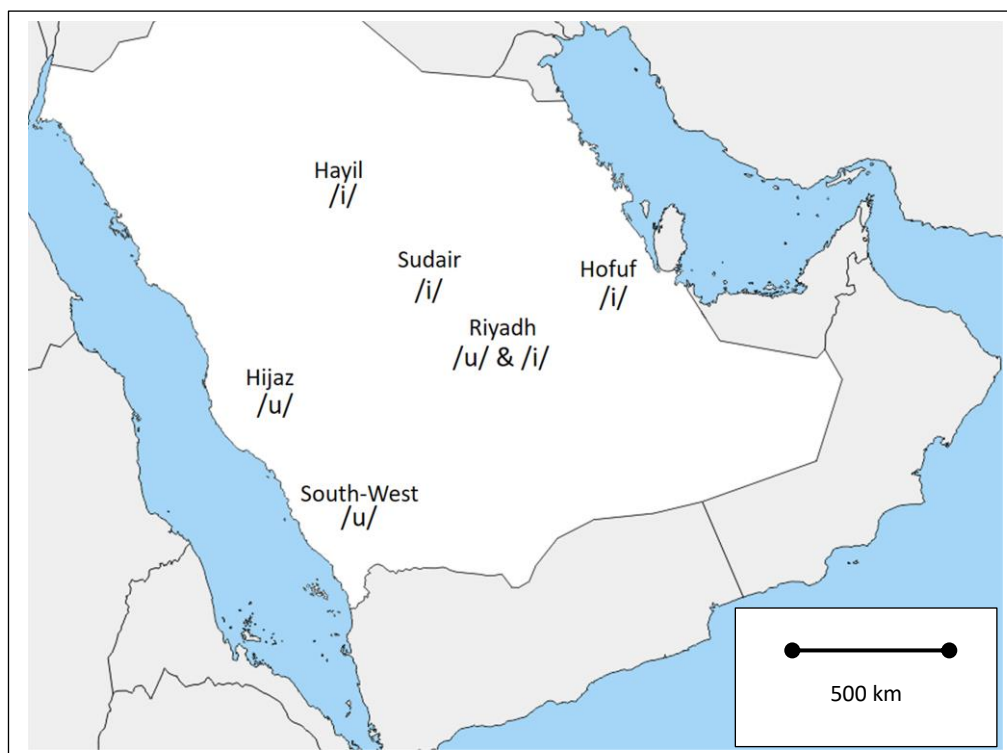
Figure 4.1: *Distribution of the merger in Saudi Arabia (Prochazka, 1988, p.17). /u/ indicates retention of the opposition, /i/ indicates the merger and /u/ & /i/ indicates that both unmerged and merged varieties are spoken there.*

4.4 The Merger in Najd: Review and Envelope of Variation

As seen in the review of this merger in the Arab world above, the reduction in the short vowel system has been associated with S dialects all over the Arab world, and the Ss of Najd are no exception. Prochazka (1988, p. 17) reported the merger to exist in several urban centres, one of which is Riyadh in Central Najd. Although he did not explicitly confine the merger to the S dialect, it is obvious that it is used by residents of urban centres who were originally S at that time. As per the present researcher's observations, the merger extends to the S dialects spoken in other cities of Najd, e.g., Hotat bani Tamim, Ad-Dilam and Al-Kharj. On the other hand, these cities, including Riyadh, are also inhabited by B tribes who are originally from the south-western regions of the Arabian Peninsula and thus retain the opposition between /u/ and /i/. Another smaller group inhabiting these cities are the WB tribes (refer to section 1.3.4),

who seem to have fully undergone the merger like the Ss. We do not know whether all WB tribes have undergone the merger. According to the researcher's own observation and the present data, however, it seems that all WB tribes have fully undergone the merger. The merger/ retention in each group (Ss, Bs, WBs) is discussed below.

4.4.1 The merger in the S dialect

Although before the oil boom the level of contact between the Ss of Najd and other dialects was low compared to other S groups across the Arab world (Ingham, 1994, p. 5), it is unknown when they lost the opposition between /u/ and /i/. It is, however, plausible that the reduction resulted from contact with other dialects over a long period of time, and that the opposition between the two vowels has been absent from the S dialects of Najd for centuries.

It should be noted that the vowel /u/ is not completely lost in the S dialect. For example, it is still used in some grammatical suffixes (Prochazka, 1988, p. 18), borrowed words from Classical Arabic and proper nouns. It is used in the object pronouns *-kum* 'you. P' and *-hum* 'them', the borrowed words of *ħubb* 'love' and *full* 'Arabian jasmine', and the proper nouns *ʕuma:n* 'Oman' and *Yanbuʕ* 'a city in Saudi Arabia'. Furthermore, Ss of Najd, especially educated ones who are exposed to Standard Arabic, are capable of using /u/ in all phonological contexts when using Standard Arabic.

4.4.2 The merger in the WB dialect

Not all B tribes that inhabit Najd are originally from the south-western parts of the Arabian Peninsula. The WB tribes treated here are small subgroups of larger tribes that inhabit the vast area stretching from Najd to Hijaz (refer to section 1.3.4). These small groups have recently migrated to some parts of Najd, notably Riyadh and Al-Kharj. Such tribes include, but are not limited to, Itban, Subae' and Suhul. The WB tribes are not the focus of this research but are included for the sake of comparison with B tribes that are originally from the south-west.

The major difference between the dialects of WB tribes and B tribes is the loss of opposition between /u/ and /i/ in WB dialect, identical to that of the Ss. There might be other linguistic differences between these two groups (WB and B), but it would require further investigation. As for similarities between them, they both led a nomadic lifestyle and their dialects are characterised as B, considering the other two variables investigated in this study (Chapter 5 and 6).

The answer to the questions 'when and why did WBs lose the opposition between the vowels?' remains unknown to the researcher. To attempt to answer the 'when?' question, the researcher observed conversations between WBs who are more than 90 years old and found no traces of /u/ (except in the cases described above in the S dialect). This means that the merger was complete at least more than a century ago. To answer the 'why?' question, it is essential to first decide whether the retention of opposition between the vowels is considered a feature of all B dialects in the Arabian Peninsula. If we consider the Arab world, not all dialects that retained the short vowels are B and not all dialects that reduced the short vowel system are S. If we consider the Arabian Peninsula, the B tribes of northern Najd, e.g., Rwalah (Prochazka 1988, p. 17), have lost the opposition, while all S dialects in the south-western regions have retained all short vowels. Therefore, the merger in WB dialects can be attributed to the place of origin, i.e., not being from the south-west (in this case), rather than the lifestyle.

To sum up, WBs share a lifestyle, values and a dialect with the Bs of Najd, but not the loss of the opposition between /u/ and /i/, which they share with the Ss.

4.4.3 The merger in the B dialect

The Najdi groups that have retained /u/ are the B tribes that are originally from the south-western regions of Saudi Arabia, namely Ajman, Qahtan and Dawasir. For the first two tribes,

/u/ is retained in the hometown of each tribe, as pointed out by Prochazcka. It is retained in Najran, a city in the south of Saudi Arabia and the hometown of the major tribe of Yaam to which the Ajman are originally affiliated, and it is retained by the Qahtan tribe whose hometown (Hasāt Qahtan) is located in the south-west region of Saudi Arabia. The last tribe, Dawasir, has been settled in Wadi Ad-Dawasir (literally translates as 'the valley of Dawasir') for decades but are originally from the southern regions of the Arabian Peninsula (Oppenheim, 2004, p. 177). Since a number of sub-groups of each tribe have been living in Najd for more than seven decades, the question arises of whether they (a) have preserved the opposition, (b) are beginning to lose it, or (c) have completely lost it. One of the goals of this chapter is to answer this question.

Table 4.4: *Examples of the merger in Najdi Arabic dialects*

Standard Arabic	B dialect	S dialect	WB dialect	Gloss
k-u-llaha	*k-u-llaha*	*k-i-llaha*	*k-i-llaha*	'all of it'
yarq-u-sˤ	*yarg-u-sˤ*	*yirg-i-sˤ*	*yarg-i-sˤ*	'he dances'
naʔx-u-ð	*na:x-u-ð*	*na:x-i-ð*	*na:x-i-ð*	'we take'
taʔk-u-l	*ta:k-u-l*	*ta:k-i-l*	*ta:k-i-l*	'she eats'
k-u-rsi	*k-u-rsi*	*k-i-rsi*	*k-i-rsi*	'chair'

4.5 Linguistic Properties of the Variable

Short vowels in Arabic are mostly part of a vocalic pattern (which alternates according to voice-active or passive-or aspect) inserted between consonants of lexical roots (see section 1.5.3.2). They are also parts of grammatical affixes or sometimes inserted between an affix and a stem. In modern Arabic dialects, short vowels in vocalic patterns as well as in grammatical affixes vary from one dialect to another. For example, the Classical Arabic simple perfective verb form *katab* 'to write', where the vocalic pattern is *a-a*, is realised as *kitab* in Najdi Arabic (Ingham, 1994, p. 23), *ktəb* in Moroccan Arabic (Caubet, 2008, p. 282),

and *katab* in Cairene and San'ani Arabic (Watson, 2002, p. 134). As for vowel variation in grammatical affixes, the Classical Arabic 3SM suffixed pronoun -*hu* 'him' is realised as -*ih/ah* in Najdi (see Chapter 5), -*uh* in Northern Najdi (see Table 1.4), -*u* in Cairene, and -*ih* in San'ani Arabic (Watson, 2002, p. 179).

In B Najdi Arabic, the environments of the short vowel /u/ are different from other dialects of Arabic, and in a few cases even Classical Arabic. For example, while the Classical Arabic imperfective verb *na-ktub* 'we write' is realised as *ni-ktub* in Hijazi, San'ani and Cairene (Ingham, 1971; Watson, 2002), it is realised as *na-ktib* in B Najdi dialect. Another example is that /u/ occurs in Hijazi Arabic next to /ħ/ as in *ruħt* 'I went' (Ingham, 1971) while in Najdi Arabic the same form is produced as *riħt*.

According to the data collected from B speakers in this study, the short vowel /u/ tends to be preceded or followed by a dorsal (/g/, /k/), an emphatic (/ðˤ/, /sˤ/, /tˤ/, /rˤ/), /χ/ or /ʕ/. It should be noted that in the S dialect, all of these occurrences of /u/ are replaced with /i/, except in the grammatical suffixes, borrowed words and proper names.

Table 4.5: *Examples of environments of /u/ in Najdi Arabic*

B dialect	S dialect	Gloss
ya:kul	*ya:kil*	'he eats'
tgusˤsˤ	*tgisˤsˤ*	'she cuts'
ta:xuð	*ta:xið*	'she takes'
nusˤsˤ	*nisˤsˤ*	'half'
ymutˤtˤ	*ymitˤtˤ*	'he stretches'
tafrˤug	*tafrˤig*	'it differs'
nafrˤuʃ	*nafrˤiʃ*	'we carpet'
yagʕud	*yagʕid*	'he sits'

4.6 Research Questions and Hypotheses

This change in the vowel system has received little attention sociolinguistically. To the best of my knowledge there is no study investigating the convergence of a three-short-vowels dialect to a two-short-vowels dialect. In addition, Najd is one of the areas that are ignored sociolinguistically. Based on the researcher's observations, the Bs are converging on the S variant while the Ss use their variant consistently. To further test this hypothesis, we aim to answer the following questions:

-Are Bs converging on the S variant, and thereby losing the opposition between /u/ and /i/, or have they already lost it altogether?

-If this change is in progress, what are the possible social factors conditioning this change?

-Have the Ss completely lost the opposition between /u/ and /i/?

-if No, what are the possible factors conditioning their use of /u/?

4.7 Results

The data presented here is obtained by interviewing 23 Bs, 13 Ss and 3WBs. The Bs, the focus group, show convergence on the S variant and the data is analysed statistically. For the other groups, only descriptive plots are presented.

4.7.1 Results for Bs

4.7.1.1 Distribution of data and descriptive statistics

For a broad perspective on the convergence rate, the percentage of Bs' use of the variant /u/ is plotted by speaker in Figure 4.2 below. Note that speakers in this figure are coded by the social factors as follows: Speaker number - Tribe (Ajmi, Dosari or Qahtani) - Age (Old or Young) - Education (High or Low) – level of contact (High or Low). Since there are only two

variants for this variable (/u/ or /i/), 65% use of /u/ in the figure below also indicates that the speaker's use of /i/ is 35%.

Preliminary observations of the data show that there is convergence towards the S variant /i/ among Bs. Our goal is to investigate (statistically) whether this convergence is conditioned by one or more social factor(s). The researcher has excluded the possibility of the effect of linguistic factor(s) on this variation (see section 3.4).
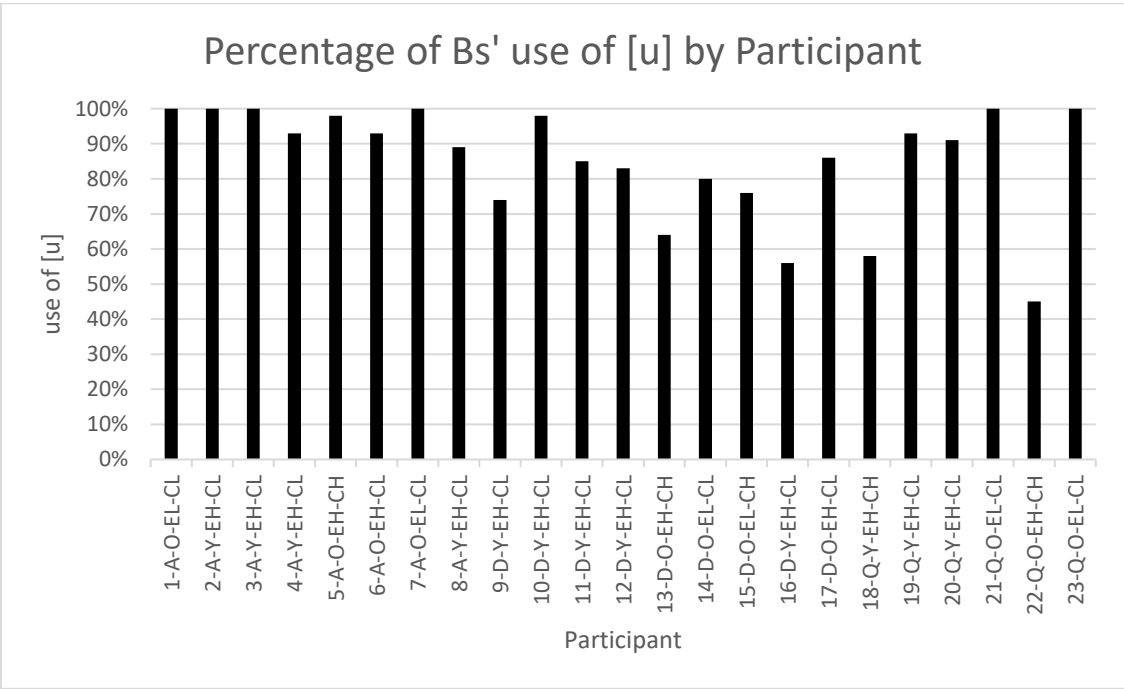
Figure 4.2: *Percentage of Bs' use of B variant [u] by Participant.*

The interviews with 23 Bs yielded 1264 tokens of the linguistic variable (/u/, /i/), with an average of 55 tokens per speaker. The distribution of the variants [u] and [i] is provided in Table 4.6 by social factors, and the distribution of each factor is additionally plotted in Figures 4.3 - 4.6.

Table 4.6: *Distribution of the variants [u] and [i] by social factor in the data for Bs*

Social Factors	Levels	Frequency of variants			Percentage	
		[u]	[i]	Total	[u]	[i]
Tribe	Ajmi	500	13	513	97%	3%
	Dosari	398	95	493	80%	20%
	Qahtani	215	43	258	83%	17%
Age	Old	564	71	635	88%	12%
	Young	551	78	629	87%	13%
Education	High	795	126	921	86%	14%
	Low	320	23	343	93%	7%
Contact	High	192	68	260	73%	27%
	Low	923	81	1004	92%	8%
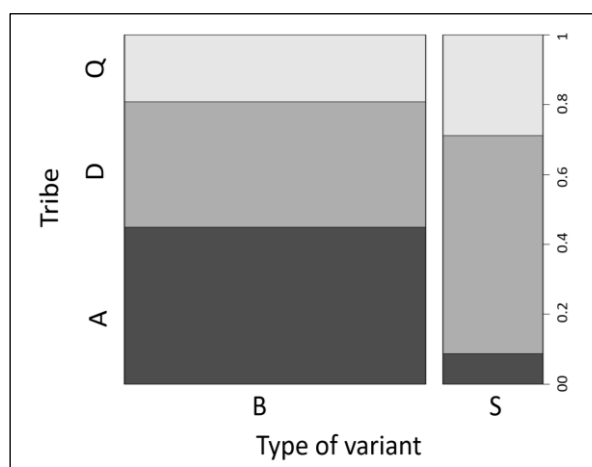
Figure 4.3: *Distribution of [u] – [i] data by the factor of tribe.*
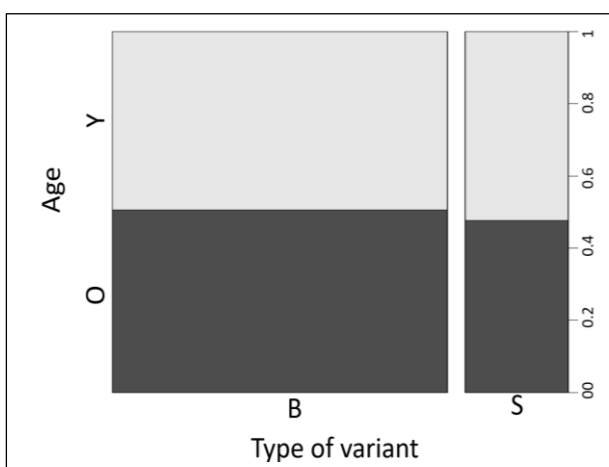
Figure 4.4: *Distribution of [u] – [i] data by the factor of Age.*
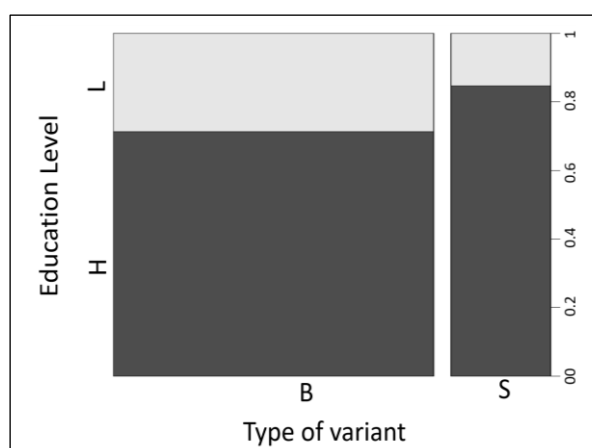
Figure 4.5: *Distribution of [u] – [i] data by the factor of education.*
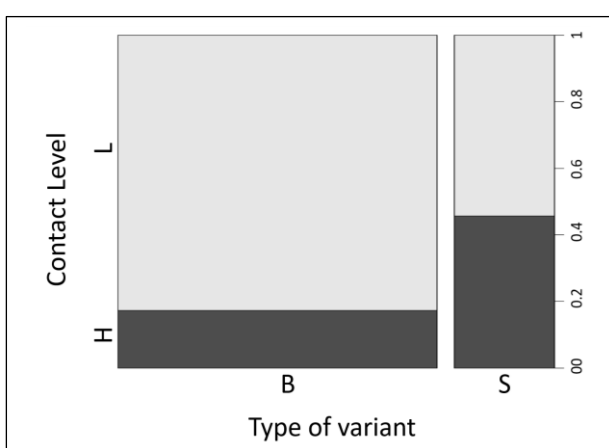
Figure 4.6: *Distribution of [u] – [i] data by the factor of contact level.*

4.7.1.2 Statistical analysis

The Bs' data set for the merger has been analysed using glmer in R. As explained in 3.7.2.2 and 3.7.2.3, glmer requires fixed and random factors. The fixed factors here are: `Tribe`, `Age`, `Education` and `Contact`; while the random factors are: `Speaker` and `Item`. The model analyses the data and provides which of the fixed factors is conditioning this variation, integrating the effect of the random factors. The glmer code used for this variable is as follows:

```
Model(Merger.in.B.dialect) <- glmer (Bs.use.of[i] ~ Tribe + Age +
Education +  Contact  + (1 | Speaker) + (1 | Item), data = Bs.(u-
i), family = binomial, control = glmerControl (optimizer="bobyqa"))
```

The model converged, and the goodness of the fit has been tested using *Somers2 Dxy* (3.7.2.4). The C value is 0.959882, which indicates a high level of fitness.

```
Fitness test of the [u]/[i] model for Bs:
somers2(probs, as.numeric(Bs.u_i$DV)-1)

         C            Dxy            n          Missing
  0.959882      0.919764   1264.000000    0.000000
```

The results of the model are provided in Tables 4.7, 4.8 and plotted in Figures 4.7- 4.10. Table 4.7 shows the results of the random factors in the model. Table 4.8, most relevant, shows the results of the fixed factors in the model. Figures 4.7 - 4.10 are plots of means and standard error.

The results of the fixed factors (Table 4.8) indicate that `Tribe` is the most significant factor. As explained in section 3.7.2.5, glmer assigns the default value of `0` to one of the levels of each factor (chosen alphabetically) and compares it with the other level(s). The level that was

assigned the default value for `Tribe` is `Ajmi`. The `Estimate` of the `Dosari` level is positive 2.9826, which indicates that Dosaris are more likely to converge on /i/ than Ajmis. The `Estimate` of `Qahtani`, on the other hand, is 1.6572, which indicates that they are more likely to converge on /i/ than Ajmis but less likely than Dosaris. The `Dosari` level is the most significant level of all factors, with a p-value of 2.26e-05. `Age`, as shown in the table was found to be non-significant. The factors of `Education` and `Contact` were found to be significant, though the latter is slightly more significant. The `Estimate` of the levels, however, is given a minus value, meaning that the respective groups are less likely to converge on /i/. In other words, people whose level of education or contact is `Low` are less likely to converge on /i/.

The discussion of the results is provided in Chapter 8.

Table 4.7: *Results of the random effects in the model (/u/ - /i/)*

Intercept	Variance	Std.Dev.	Observations	Words	Participants
Item	5.036	2.244			
Participant	1.050	1.025	1264	273	23

Table 4.8: *Results of the fixed effects in the model (/u/ - /i/)*

Factors/Level	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.0473	0.8512	-3.580	0.000344 ***
Tribe-Dosari	2.9826	0.7039	4.237	2.26e-05 ***
Tribe-Qahtani	1.6572	0.8019	2.067	0.038774 *
Age-Young	-0.1979	0.7525	-0.263	0.792530
Education-Low	-1.8457	0.8742	-2.111	0.034749 *
Contact-Low	-2.1721	0.7363	-2.950	0.003175 **

Signif. codes: 0 '*'   0.001 '**'   0.01 '*'   0.05 '.'   0.1 ' ' 1**
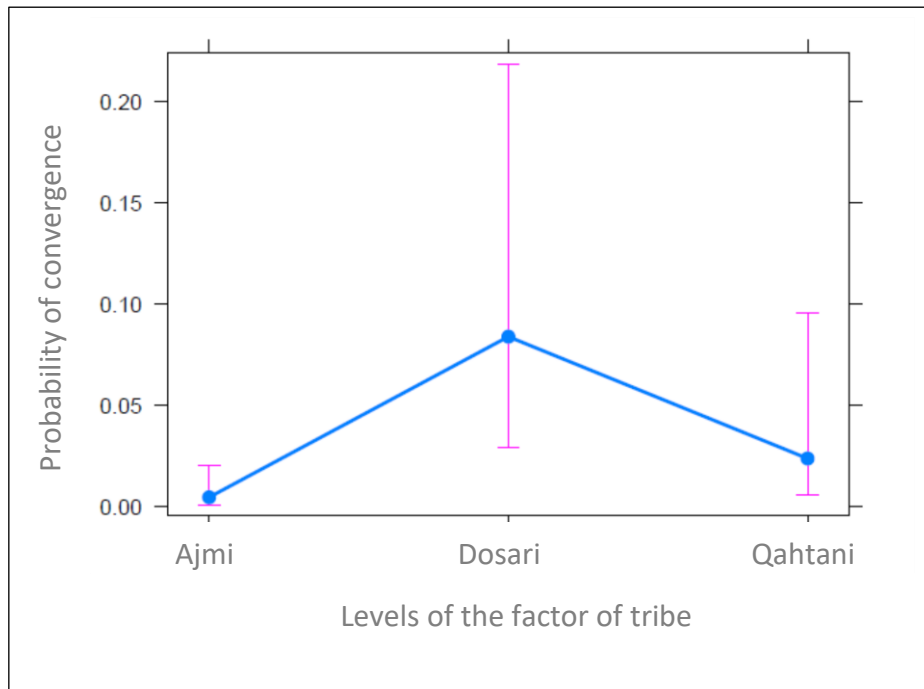
Figure 4.7: *Predicted probability of convergence on the S variant /i/ by the factor of tribe. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (Significant)*
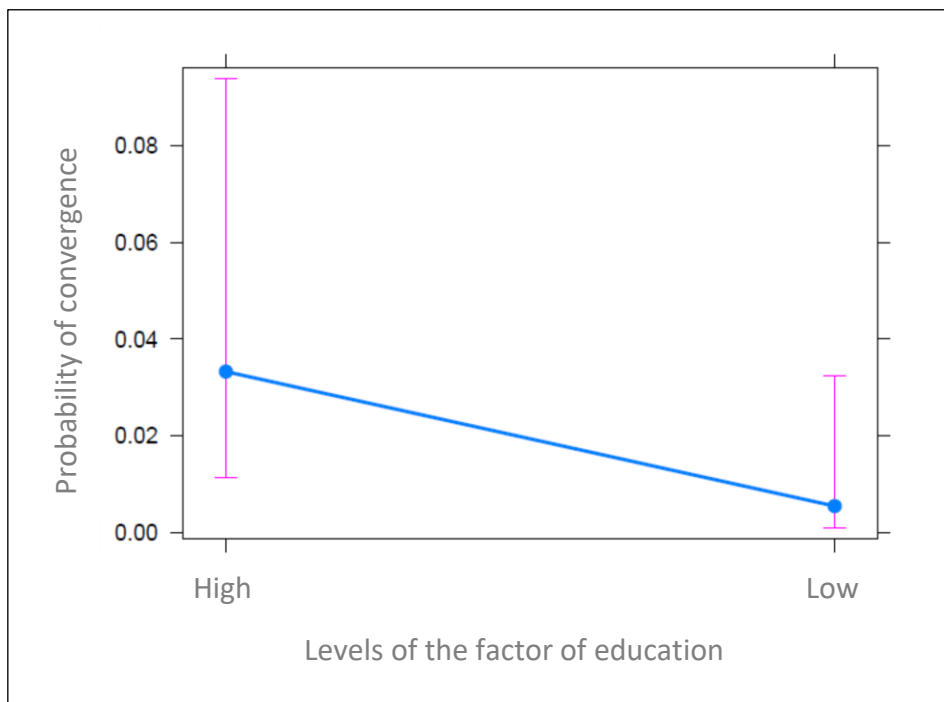
Figure 4.8: *Predicted probability of convergence on the S variant /i/ by the factor of education. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (Significant)*
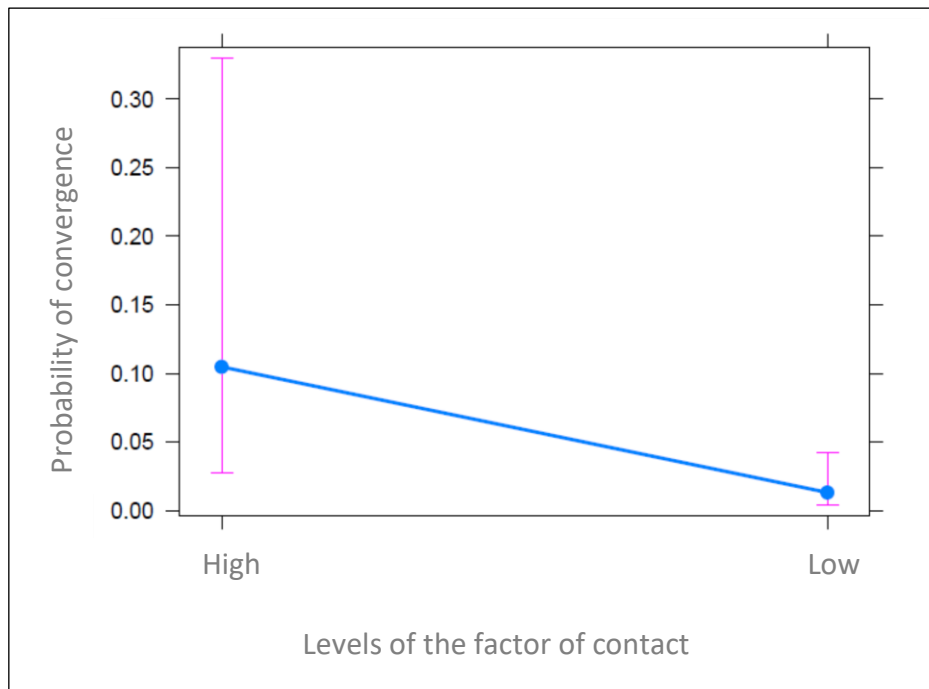
Figure 4.9: *Predicted probability of convergence on the S variant /i/ by the factor of contact. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (Significant)*
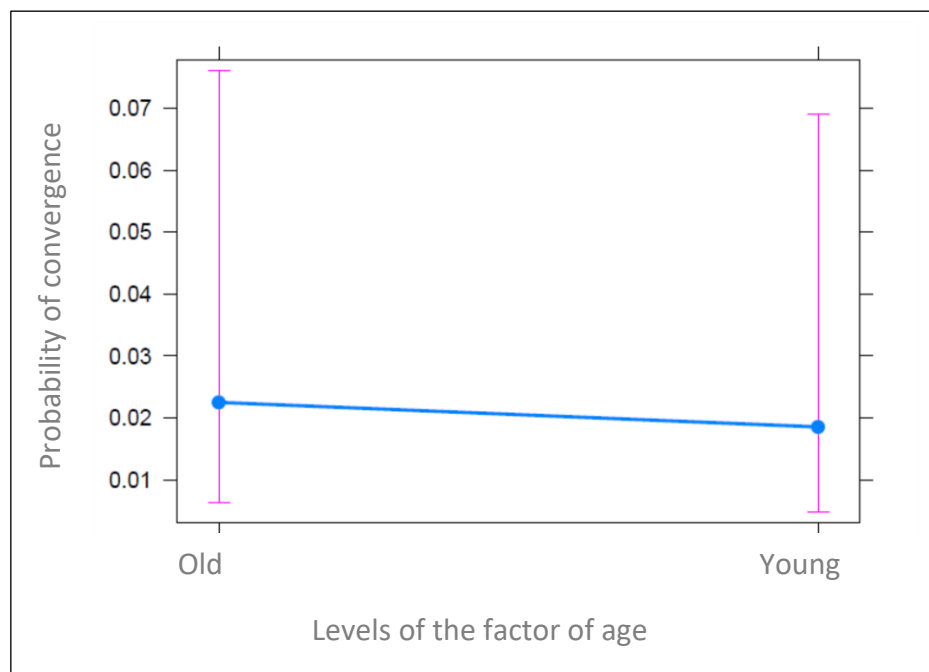
Figure 4.10: *Predicted probability of convergence on the S variant /i/ by the factor of age. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (Not significant)*

4.7.2 Results for WBs

Although the WB tribes are not the focus of this research, 3 speakers were interviewed, each from a different tribe, for the sake of comparison with Bs. The only reason WBs are distinguished from Bs in this research is because they use /i/ all the time, like the Ss. It is hypothesised that WBs will not use /u/ at all, except for the cases noted in section 4.4.1. The data for each speaker is provided in Table 4.12.

Table 4.12: *Interview data for WB speakers*

Speaker/Tribe	Age	Frequency		Percentage	
		/u/	/i/	/u/	/i/
Suba'	20	0	16	0%	100%
Suhul	30	0	41	0%	100%
Itban	55	0	31	0%	100%

The data might not be sufficient to generalise to all WBs. But based on this data and the researcher's observations of other WB speakers, it can be claimed that WB tribes have completely undergone the merger, just like the Ss. A further investigation of a larger sample of WB tribes might reveal that some speakers use /u/; however, we cannot be sure whether it is a result of exposure to Standard Arabic or that it is a residue of the variant in their dialect. This can only be demonstrated by interviewing a large sample of WB speakers who are less educated and hence less exposed to Standard Arabic. However, as stated earlier, this group is not the focus of this research.

It should be noted that except for this linguistic variable, WB dialect is characterised as B because it exhibits all B features found in the B dialect.

4.7.3 Results for Ss

One of the goals of this study is to investigate whether the B/S dialect distinction is still evident in Najd, especially after the massive urbanisation and heavy contact (with the Bs in

this case). Since the members of neither group are characterised as migrants because they all moved to the city relatively soon after the oil boom, and because neither group is known to be overtly prestigious, we are not sure if the Ss will converge on the B dialect.

13 S speakers were interviewed, and they are grouped by age and style. Age groups are old and young, as with Bs. As for style, *informal style* refers to when speakers did not use Standard Arabic while *formal style* indicates that they used it (refer to section 3.6.3). In relation to tribe, Ss do not affiliate themselves with their larger tribes (see 1.3.4) but rather to sub-groups. In other words, it is linguistically irrelevant to associate Ss with tribes, as we did with Bs, because the structure and value of tribal affiliation in the S communities is different. The interview data for Ss are provided in Table 4.13 and plotted in Figure 4.13.

The data shows that Ss do not converge on the B variant but rather used /i/ all the time, except when they switch to Standard Arabic in which /u/ is used. Those who viewed the interview as informal and therefore used spoken dialect show no variation at all. The other 4 speakers who viewed the interview as formal used Standard Arabic. As we will see in 8.5, recognising Standard Arabic forms is based on the appearance of Standard Arabic features and lexical items that are absent in Najdi Arabic. Examples of such features are medial glottal stop and the sound /q/ (>/g/)

To summarise, Ss do not use /u/ in their spoken dialect except when the situation requires using Standard Arabic (the interview in this case). That is, Ss do not use /u/ when they are speaking to family members and close friends; however, it exists in their repertoire as they use it when they switch to Standard Arabic.

Table 4.13: *Ss' use of their variant (/i/) in the interviews*

Participant	Age	Style	Ss.tokens	Percentage
1	Old	Informal	36/36	100%
2	Young	Informal	56/56	100%
3	Young	Informal	53/53	100%
4	Young	Informal	41/41	100%
5	Young	Informal	53/53	100%
6	Young	Informal	47/47	100%
7	Old	Informal	39/39	100%
8	Old	Informal	45/45	100%
9	Old	Informal	40/40	100%
10	Old	Formal	13/19	68%
11	Old	Formal	20/34	58%
12	Young	Formal	27/29	93%
13	Young	Formal	13/14	92%
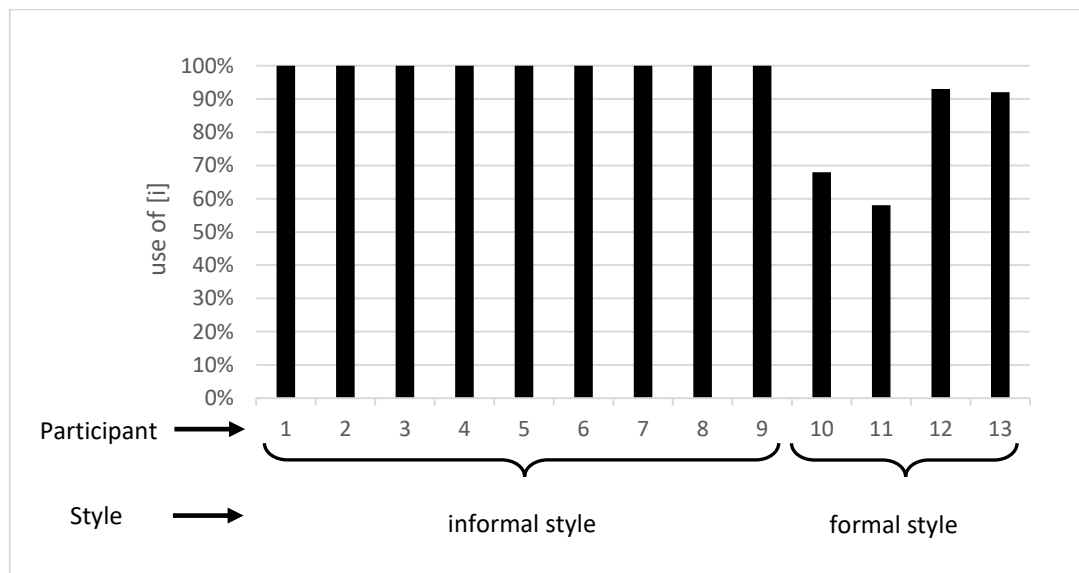
Figure 4.13: *Percentage of the Ss' use of the S variant /i/ plotted by participant and style.*

4.8 Conclusion

The review of the short vowel merger in the Arab world shows that S dialects tend to reduce the short vowel system to two vowels. The S dialect of Najd has reduced its short vowel system by undergoing the merger /u/, /i/ > /i/. The Najdi B dialect spoken by tribes that are originally from the south-western region of the Arabian Peninsula has retained all three short vowels.

We hypothesised that the Bs are converging on the S variant. To test the hypothesis, 23 Bs were interviewed and the data shows that the Bs are in the process of losing the opposition between /u/ and /i/. When statistically analysed using glmer, the data shows that the factors of tribe, education and level of contact are significant.

The Ss on the other hand, tend to use the S variant all the time except for when they switched to Standard Arabic. Such shift is triggered by several factors, but we can mainly attribute it to the perceived formality of the interview.

Additional data aiming to measure the relative sociolinguistic salience of this variable, in comparison with the other variables, will be provided in Chapter 7. The discussion of the results of this variable from this chapter is provided in Chapter 8.

Chapter Five: The 3rd Person Singular Masculine

Possessive/Object Suffixed Pronoun (*-ih/-ah*)

This chapter examines morphophonemic variation in the 3rd person singular masculine

possessive/object suffixed pronoun (henceforth, 3SM suffixed pronoun) in Najdi Arabic. Due

to a scarcity of sociolinguistic studies on Najd, both forms of the 3SM suffixed pronoun under

study (*-ih* and *-ah*) have been reported in Najd, but neither has been specifically associated

with the B or the S dialect. The researcher observed that Bs tend to use *-ih* while Ss use *-ah*.

The fundamental goal of this chapter is to establish whether the aforementioned forms are

associated with the B/S dialect distinction in Najd. The other goal is to pinpoint which of the

following scenarios resulted from the heavy contact between the two groups in the last five

decades:

- Each group uses their variant consistently (i.e., no variation)

- The Bs are converging on the S variant.

- The Ss are converging on the B variant.

The first part of the chapter will provide an overview of personal pronouns in Arabic, the

evolution of the 3SM suffixed pronoun in modern Arabic dialects and a review of the pronoun

forms in these dialects. In the second part, the data analysis and results for this variable are

presented.

5.1 Overview of Personal Pronouns in Arabic

Personal pronouns in Classical Arabic are divided into independent and suffixed pronouns.

The first group represents free-form subject pronouns, while the second represents

object/possessive pronouns that are suffixed to verbs (as object) or nouns (possessive suffix);

see Table 5.1. The personal pronouns of Arabic, both independent and suffixed, can be further broken into smaller units, each of which indicates person, number, or gender. For example, /ʔant-/ in the independent pronouns' paradigm and /-k/ in the suffixed one both indicate 2nd person; /-h/ in both paradigms indicates 3rd person; /m/ indicates masculine; /nna/ indicates feminine; /a/ indicates masculine; and /i/ indicates feminine in the 2nd person singular forms (Trager & Rice, 1954).

Table 5.1: *Personal pronouns of Classical Arabic (Trager & Rice, 1954), (S=singular, P=plural, C=common, D=Dual)*

Person	Number.gender	Independent Pronouns	Suffixed pronouns
1st person	SC	*ʔana:*	*-i:/ -ni*
	PC	*naħnu*	*-na:*
2nd person	SM	*ʔanta*	*-ka*
	SF	*ʔanti*	*-ki*
	DC	*ʔantuma:*	*-kuma:*
	PM	*ʔantum*	*-kum*
	PF	*ʔantunna*	*-kunna*
3rd person	SM	*huwa*	*-hu*
	SF	*hiya*	*-ha:*
	DC	*huma:*	*-huma:*
	PM	*hum*	*-hum*
	PF	*hunna*	*-hunna*

The suffixed pronouns in Arabic are used in both accusative/dative and genitive cases, corresponding to English pronouns *us/our*, *her/her*, *him/his*, etc. Object pronouns are suffixed to verbs to indicate both direct and indirect object, while possessive pronouns can be suffixed to nouns to indicate possession or to preposition (e.g., Classical Arabic *la-hu* 'for him') to serve other functions (Lipinski, 1997, p. 306). The forms of suffixed pronouns are identical for both cases except for the 1st person singular forms which are *-i:* for genitive and *-ni* for accusative/dative.

Table 5.2: *Examples for suffixed pronouns in Classical Arabic*

Object pronouns	Possessive pronouns
dˤarab-a-ni	*qalam-i:*
hit-he-me	pen -my
'He hit me'	'My pen'
nu-ʕtˤi:-ha:	*sayya:ratu-ha:*
we-give-her	car-her
'We give her'	'Her car'
darras-tu-hu	*baytu-hu*
taught-I-him	house-his
'I taught him'	'His house'
la-na:	*kita:bu-na:*
for-us	book-our
'For us'	'Our book'

Now we turn to the 3SM suffixed pronoun *-hu* in Classical Arabic, highlighted in Table 5.1. It is composed of /h/ which indicates 3rd person and the optional vowel /u/. This vowel, unlike the vowel /a:/ in the feminine form *-ha:*, can be dropped in pause position, e.g., *kita:bu-hØ*, or even changed to /i/ when the last vowel of the stem is /i/ or /ay/, e.g., *kita:bi-hi* 'his book', *ʕalay-hi* 'on him' (Appleyard, 2008; Trager & Rice, 1954). In the next section the evolution of this pronoun in modern Arabic dialects is discussed.

5.2 The Evolution of 3SM Suffixed Pronoun in Modern Arabic Dialects.

The pronominal system in modern Arabic dialects has changed slightly. The major changes are: (a) loss of the dual person in all dialects and (b) loss of gender marking in plural forms in most dialects, but retained in Negev and Najdi Arabic (see Blanc, 1970, p. 130; Ingham, 1982a, p. 74; cited in Vicente, 2008, p. 585). The remaining changes involve mostly vowels (e.g., change in quality, epenthesis, syncope and/or lengthening), most of which resulted from

either internal evolution in these dialects or the influence of a substratum language (Vicente, 2008, p. 584).

The 3SM suffixed pronoun in modern Arabic dialects has changed from Classical Arabic -*hV* (where *V* is mostly /u/; see the section above) to -*Vh* in most B dialects and -*V* in most S dialects. In the B form, *V* can be occupied by any short vowel of the Arabic language, and this vowel is dropped post-vocalically (i.e., when the stem ends in a vowel as in ʕasʕaː-*h* 'his stick'). In the S form, the vowel, as the entire suffix, can be /u/, /o/ or /a/ (see next section), while post-vocalically it is dropped and a weak /h/ is added, as will be explained below (Versteegh, 2001; Vicente, 2008). It should be noted that the post-vocalic form of the pronoun is irrelevant to the current study, though it will be mentioned later.

The reason for the evolution of Classical Arabic -*hV* to -*Vh* / -*V* in modern Arabic dialects is attributed to the loss of the Classical Arabic case marking system in all modern dialects. Because of case marking, all Classical Arabic forms end in a short vowel (marking a case) to which the pronoun -*hu* (-CV) is attached (Appleyard, 2008, p. 590). Therefore, modern Arabic dialects, in which case marking has been lost, employ -*Vh* (-VC) or -*V* because most stems end in a consonant, and thus a vowel is needed for syllabification; see Figure 5.1 below.

The vowel in the B form -*Vh* is originally epenthetic to repair the syllable in a post-consonantal environment. All short vowels of Arabic have been attested in this form of the pronoun (i.e., -*ah, -ih, -uh*). Although the vowel varies from one dialect to another, it is usually used consistently in a given paradigm of a dialect. For instance, a dialect that is reported to use -*uh* in an imperfective verb form is likely to use it, at least, in all verb forms of the imperfective paradigm. Turning to the S form -*V*, it is rather straightforward as the pronoun is added to all words that end in a consonant. When the word ends in a vowel, the vowel of the stem is lengthened and a weak /h/ is added, as mentioned earlier.

Before reviewing the forms of the pronoun in the modern dialects of Arabic, it is essential to point out that the B form -*Vh* is not strictly used in B dialects only, but rather it might be attested in some S dialects, too. The forms -*uh* and -*ah*, for example, are reported to be used in the S dialects of the Arabian Peninsula, as we shall see below.

Classical Arabic:

Kita:b-u -hu
↑   ↑↑
V -CV
↑
Case marking vowel

Most B dialects:

Kita:b -Vh
↑   ↑↑
C -VC

Most S dialects:
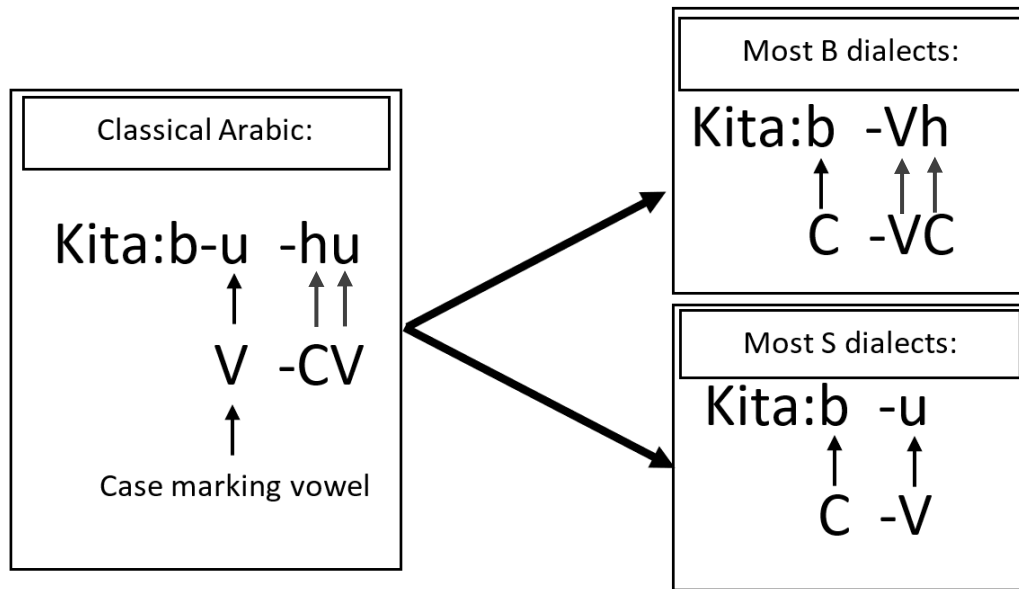
Kita:b -u
↑   ↑
C -V

Figure 5.1: *The evolution of the 3SM suffixed pronoun in modern Arabic dialects (Appleyard, 2008; Vicente, 2008). Kita:b-u-hu means 'his book'.*

Table 5.3: *Example of 3SM suffixed pronoun in modern Arabic dialects (based on Versteegh, 2001; Vicente, 2008)*

Environment	Most B dialects	Most S dialects	Gloss
Post-consonantal	*yaðˤirb-ih*	*yi-dˤrab-u*	'he hits him'
	be:t-ih	*be:t-u*	'his house'
	mrit-ih	*marat-u*	'his wife'
Post-Vocalic	*ʕasˤa:-h*	*ʕasˤ-a:(h)*	'his stick'
	ʔix-u:h	*ʔax-u:(h)*	'his brother'

5.3 Review of the 3SM Suffixed Pronoun Forms in Modern Arabic Dialects

In this section, a brief overview of the pronoun forms in modern Arabic dialects is provided. Although the grouping of dialects in relation to the pronoun forms in the previous section was

B and S, dialects in this section are listed geographically. As we did in the review of the merger in modern Arabic dialects (Chapter 4), we shall group Arabic dialects as Western, Levantine and those spoken in the Arabian Peninsula.

As stated earlier, there are two allomorphs of the pronoun: post-consonantal and post-vocalic. In Table 5.4, the forms of the 3SM pronoun in modern Arabic dialects are shown, except for Saudi Arabia and Najd which are discussed separately below.

Table 5.4: *3SM suffixed pronoun in modern Arabic dialects*[23]

Part of the Arab World	Dialect	Post-consonantal form	Post-vocalic form	Reference
West	Algiers	*-u*	*-h*	Boucherit (2006, p.63)
	Cairo	*-u*	*-h*	Woidich (2006, p.326)
	Morocco	*-u*	*-əh*	Caubet (2008, p.277)
	Tripoli	*-a*	*-h*	Pereira (2009, p.551)
	Tunis	*-u*	*-h*	Gibson (2009, p.566)
	Khartoum	*-u*	*-hu*	Dickens (2007, p. 560)
Levant	Baghdad	*-a*	*-h*	Abu-Haidar (2006a, p.226)
	Beirut	*-u*	*-Ø*	Naïm (2006, p.279)
	Damascus	*-o*	*-Ø*	Lentin (2006, p.548)
	Amman	*-o*	*-h*	Al-Wer (2007a, p.510)
	Palestine	*-oh (S)/-ah* (B)	*-h*	Shahin (2008, p.531)
Arabian Peninsula	Bahrain	*-ah*	*-h*	Holes (2006, p.247)
	Kuwait	*-ah*	*-h*	Holes (2007, p.613)
	Oman	*-uh (S)/-ah* (B)	*-h*	Edzard (2008, p.483)
	Sana'a	*-ih*	*-h*	Watson (2009, p.110)

---

[23] In Tripoli, Tunis, Baghdad and Oman, the post-vocalic form is not reported in the source given. However, the form is assumed to be *-h* according to Vicente (2008).

The forms reported in Table 5.4 cover the dialects that are spoken in the capital cities of the Arab world, mostly S. It can be stated that /h/ in the post-consonantal form is lost in S dialects of the West and the Levant but retained in Palestine and the Arabian Peninsula. As for B dialects in the West, some groups like Sinai dialects retained /h/ (De Jong, 2009), while others have lost it, e.g, *-u* in the Hassaniyya dialect spoken in Western Sahara (Taine-Cheikh, 2008, p. 243). In the Arabian Peninsula, B dialects retained /h/, similarly to S groups in the area, but the quality of the epenthetic vowel may vary, e.g., Omani *-uh* (S), *-ah* (B) (Edzard, 2008, p. 483) and the current variable in Najdi Arabic *-ah* (S), *-ih* (B).

The loss of /h/ in modern Arabic dialects of the Levant and Africa, but not in the Arabian Peninsula, is probably attributable to the Arabic koine discussed in 2.2 (Cohen, 1970; Ferguson, 1959a; Versteegh, 1984).

5.4 The 3SM Suffixed Pronoun in Saudi Arabia

In Saudi Arabia, /h/ is retained in most dialects, B and S, except for urban Hijazi which employs *-u*, like Cairene (Ingham, 1971; Isaksson, 1991). The quality of the epenthetic vowel varies across geographical and social dialects, but it is restricted to [u], [a], or [i] (but not [o]).

According to Prochazka (1988, p. 126), both forms *-ih* and *-ah* are used in the south-western dialects of Saudi Arabia, but neither of them was associated with a social group. However, Al-Azraqi (1998) reports *-ah* in the dialect of Abha (a city in the south-west) while Al-Qahtani (2015) reports *-ih* in the dialect of the south-western Qahtani tribe. In the East, Prochazka (1988) reports *-ah* in the dialect of Al-Hofuf, similar to what is found in Kuwait and Bahrain. In Northern Najd, the form *-uh* is reported in the dialects of Qasim and Hayil (Abboud, 1979; Ingham, 1994).
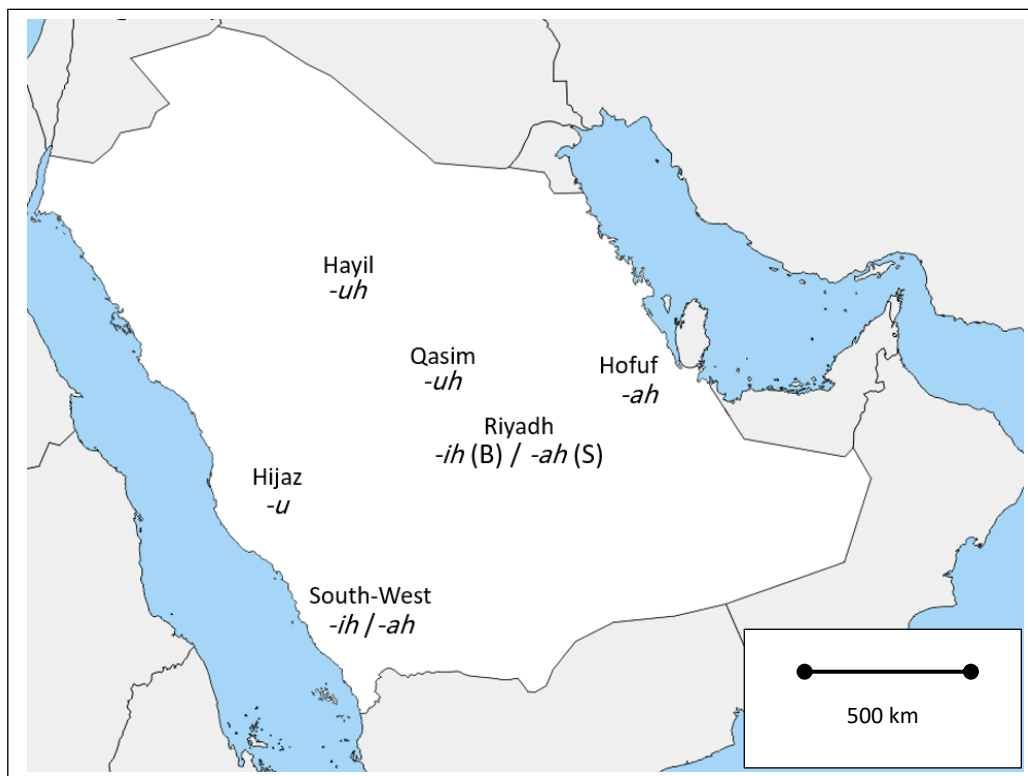
Figure 5.2: *Distribution of 3SM suffixed pronoun in Saudi Arabia (Prochazka, 1988; Ingham, 1971; Ingham, 1994).*

5.5 The 3SM Suffixed Pronoun in Najd: Envelope of Variation

In Central Najdi dialects, the forms *-ih* and *-ah* are reported by Prochazka (1988). In Ingham (1994), only *-ih* is reported as the 3SM suffixed pronoun of Najdi Arabic. Al-Essa (2008) reports both forms *-ih* and *-ah* in Riyadh. However, none of these authors has associated either form with a particular social group. The data from the current study proves that *-ih* is used in Najdi B dialect while *-ah* is used in the S dialect. The post-vocalic form for both Najdi groups is *-h*, but it is irrelevant to the study.

5.5.1 The 3SM suffixed pronoun in Najdi B and WB dialects

As observed in the data from the current study, the B dialect in Najd, spoken by B tribes that are originally from the south-western parts of Saudi Arabia, use *-ih* as the 3SM suffixed pronoun. In addition, the same form is noted in the WB dialect, spoken by the B tribes that are

originally from the mid-western parts of Saudi Arabia. The history behind the adoption of this form, i.e., the epenthetic vowel /i/, is unknown to the researcher, and to the best of my knowledge there are no attempts to explore the evolution of these forms.

The vowel in the Najdi B dialect form is always /i/ (*-ih*); however, it is produced as *-ah* when it is added to perfective verb forms and the subject is 3SM,[24]. The reason for this allomorphy is attributed to the Classical Arabic perfective 3SM subject-verb agreement marker *-a* (which is *-Ø* in Najdi Arabic). Consider the Classical Arabic forms *katab-a-hu* 'he wrote it-M' and *ra:sal-a-hu* 'he sent him a message' compared to the Najdi forms *ktib-ah* and *ra:sal-ah*, where it is likely that the use of /a/ instead of /i/ in the suffix is attributed to the Classical Arabic 3SM subject-verb agreement marker *-a*.[25]

However, it should be noted that any change to the structure of these verbs, e.g., addition of suffix, blocks the use of *-ah* in all perfective verb forms and *-ih* is used instead. For instance, if the suffix *-t* '2SM' is added to the stem before the 3SM pronoun, the epenthetic vowel stays /i/ as in *kitab-t-ih* 'you wrote it-M' (and not *kitab-t-ah*, which is the S form).

Table 5.5: *Examples for -ih and -ah in B and WB dialects*

[*-ih*]	Gloss	[*-ah*]	Gloss
sayya:rt-ih	'his car'	*ðˤrib-ah*	'he hit him'
ʕallam-t-ih	'I told him'	*ga:bal-ah*	'he met him'
l-ih	'for him'	*ʔadxal-ah*	'he caused it to enter'
minn-ih	'from him'	*istaftaħ-ah*	'he started it'
ya-ksr-ih	'he breaks it'	*garrib-ah*	'he made it closer'
zabbatˤ-t-ih	'I tidied him'	*zabbitˤ-ah*	'he tidied him'

---

[24] The Najdi Arabic perfective verb forms to which the 3SM suffix can be added are verb forms 1-4 and 9 in Table 1.2 (Chapter 1).

[25] For the perfective form (CVCVC) there is a re-syllabification rule involved when a -VC suffix is added. Ingham (1994, p. 19) reports that when a -VC suffix is added to CVCVC (e.g., kitab 'wrote'), the form is re-syllabified as CCVC-VC. Thus, when the suffix *-at* 'she' is added to the stem *kitab-* 'wrote', the form produced is *ktibat* 'she wrote'. This rule also applies to the 3SM *-ih* causing it to be produced as *ktib-ah* 'he wrote it-M' (CCVC-VC).

To sum up, in the B and WB dialects of Najd, the 3SM suffixed pronoun is realised as *-ih*, with the exception that the allomorph *-ah* is only used when the suffixed pronoun is: a) added to a perfective verb, b) the subject is 3SM and c) it is the only suffix added to the stem.

5.5.2 The 3SM suffixed pronoun in Najdi S dialect

Just like other S dialects in the Arabian Peninsula, the Najdi S dialect has retained /h/ in the suffixed pronoun. Unlike the B and WB dialects, the pronoun in the S dialect is always *-ah*, regardless of whether it is an object pronoun with a perfective or imperfective verb, a possessive pronoun, or even when added to prepositions. The history of the adoption of the epenthetic vowel /a/ by the S dialect is also unknown to the researcher. It is also not reported whether the Ss used to have more than one allomorph for the pronoun (like the Bs) and that they were eventually levelled to *-ah*.

Table 5.6: *Examples of 3SM suffixed pronoun in S and B dialects*

Najdi S dialect	Najdi B dialect	Gloss
sayya:rt-ah	*sayya:rt-ih*	'his car'
ʕallam-t-ah	*ʕallam-t-ih*	'I told him'
l-ah	*l-ih*	'for him'
minn-ah	*minn-ih*	'from him'
yi-ksr-ah	*ya-ksr-ih*	'he breaks it'
zabbatˤ-t-ah	*zabbatˤ-t-ih*	'I tidied him'

5.6 Research Questions and Hypotheses

As a native speaker of Najdi B dialect, the researcher observed the variation in *-ih/-ah* between B and S speakers. It is hypothesised that Bs always use *-ih* (except in the environment specified above in section 5.5.1) while the Ss always use *-ah* in the environments where *-ih* is used by the Bs. This hypothesis is based on another hypothesis which suggests

that unlike the merger (Chapter 4), which is salient (and therefore Bs are converging on the S variant), this variable (*-ih/-ah*) is less salient, and consequently there will be less or no convergence (more discussion of salience of the variables is provided in Chapter 7). The analysis of the interview data for both Bs and Ss will hopefully answer the following research questions:

- Is one group converging on the other group's variant?
- If yes, what are the social factors conditioning this convergence?

5.7 Results

In this section, we will present the interview data for this variable for both groups. It should be noted that only the occurrences where Bs would use *-ih* are calculated. That is, the occurrences of the variable in the environment where Bs use the allomorph *-ah* (which is identical to the S variant) are not included in the quantitative analysis of the data for either group.

5.7.1 Interview data for Bs

The production data for the Bs was collected in interviews with the 23 B speakers, grouped by tribe, age, education and contact. The interviews yielded 1068 tokens, with an average of 46 tokens per speaker.

5.7.1.1 Distribution of data and descriptive statistics

The data shows slight convergence on the S variant, possibly conditioned by one or more sociolinguistic factors. For a broader perspective on the convergence rate, the data is plotted in Figure 5.3 by speaker. Note that speakers in the figure are coded by social factors as follows: Speaker number - Tribe (Ajmi, Dosari or Qahtani) - Age (Old or Young) - Education

(High or Low) - Contact (High or Low). Since there are only two variants (-*ih* or -*ah*), 90%

use of -*ih* in the figure, for example, also indicates that the speaker's use of -*ah* is 10%.
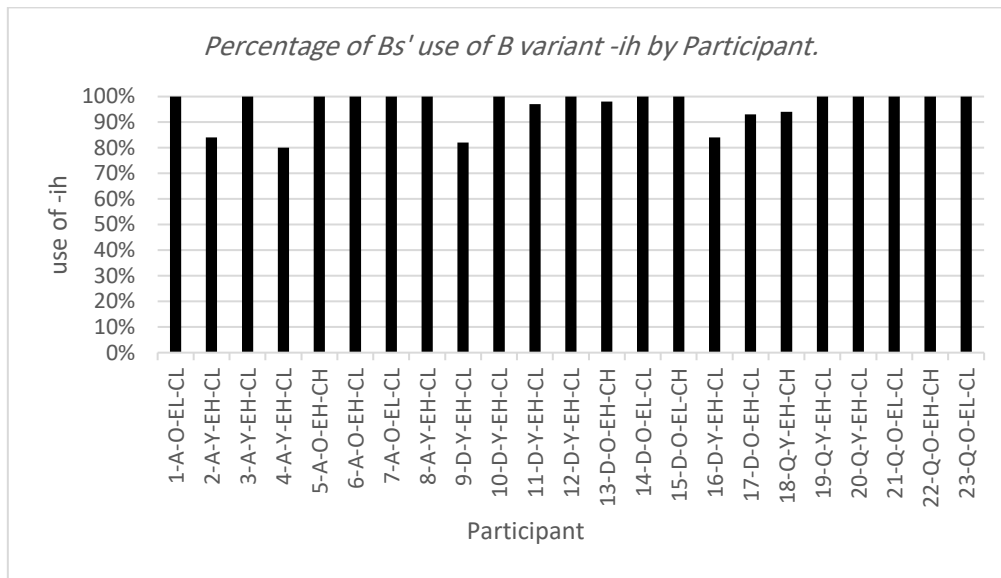
Figure 5.3: *Percentage of Bs' use of B variant -ih by Participant.*

The distribution of data by social factor is provided in Table 5.7 and plotted in Figures 5.4-

5.7. The possibility of the effect of linguistic factors has been excluded since some items were

produced with both variants by the same speakers (see section 3.4). That is, during the

interview some speakers produced the word for 'to him' once as *l-ih* and once as *l-ah*.

Table 5.7: *Distribution of the variants -ih and -ah by social factor in the interviews for Bs*

Social Factors	Levels	Frequency of variants			Percentage	
		[-ih]	[-ah]	Total	[-ih]	[-ah]
Tribe	Ajmi	525	15	540	97%	3%
	Dosari	301	15	316	95%	5%
	Qahtani	210	2	212	99%	1%
Age	Old	600	4	604	99%	1%
	Young	436	28	464	93%	7%
Education	High	704	32	736	96%	4%
	Low	332	0	332	100%	0%
Contact	High	229	3	232	99%	1%
	Low	807	29	836	96%	4%
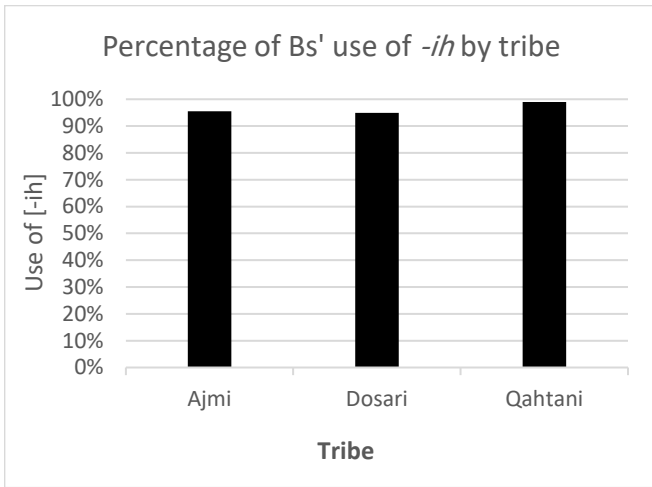
Figure 5.4: *Percentage of Bs' use of B variant -ih by the factor of tribe.*
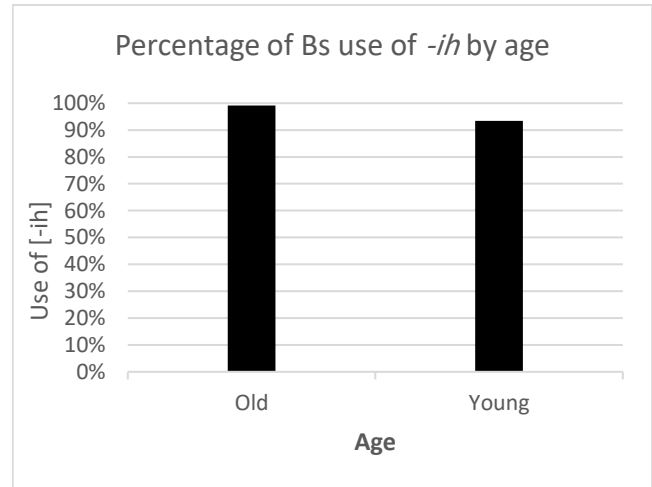
Figure 5.5: *Percentage of Bs' use of B variant -ih by the factor of age.*
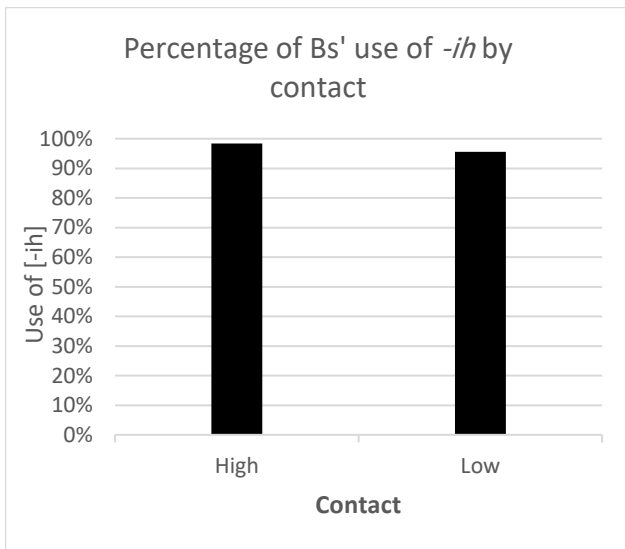
Figure 5.6: *Percentage of Bs' use of B variant -ih by the factor of contact.*
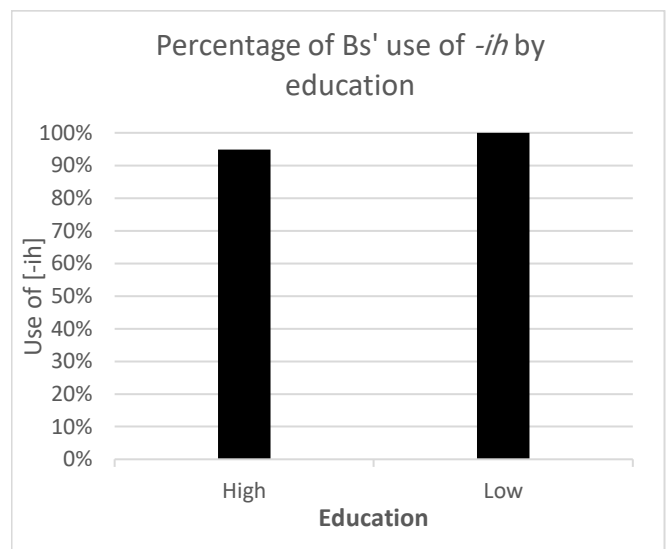
Figure 5.7: *Percentage of Bs' use of B variant -ih by the factor of education.*

Table 5.7 and Figures 5.4 - 5.7 indicate the following:

- The Ajmi and Dosari tribes show slightly more convergence than Qahtani tribe.

- Young speakers show a considerably higher rate of convergence than old speakers.

to examine these observations, the data is statistically analysed in the next section.

5.7.1.2 Statistical analysis

The Bs' data set for the 3SM suffixed pronoun has been analysed using glmer in R. As explained in 3.7.2.2, glmer requires both fixed and random factors. The fixed factors for the Bs are: `Tribe`, `Age`, `Education` and `Contact`; while the random factors are: `Speaker` and `Item`. The glmer code used for this variable was as follows:

```
Model.(Bs.use.of.ah) <- glmer ( Bs.use.of.ah ~ Tribe + Age +
Education + Contact + (1 | Speaker) + (1 | Item), data = Bs.(ih/ah),
family = "binomial", control = glmerControl (optimizer="bobyqa"))
```

However, as mentioned in 3.7.2.4, a model may or may not converge, depending on various factors. The previous model did not converge. After several attempts, the model converged, with the factor of `Education` removed, and the level of goodness of fit was high (see C value in *Somers2 Dxy* test below). The new glmer code is:

```
Model.(Bs.use.of.ah) <- glmer ( Bs.use.of.ah ~ Tribe + Age +
Contact + (1 | Speaker) + (1 | Item), data = Bs.(ih/ah), family =
"binomial", control = glmerControl (optimizer="bobyqa"))
```

Fitness test of the -ih/-ah model for Bs:

```
    somers2(probs, as.numeric(Bs.ih_ah$DV)-1)

        C              Dxy              n            Missing
   0.9638031        0.9276062      1068.0000000      0.0000000
```

151

The results of the model are provided in Tables 5.8 and 5.9 and plotted in Figures 5.7 - 5.9. Table 5.8 shows the results of the random factors in the model. Table 5.9, the most relevant, shows the results of the fixed factors in the model. Figures 5.7 - 5.9 are plots of means and standard error.

The results of the fixed factors (Table 5.9) indicate that Age is the only significant factor. As explained in 3.7.2.5, glmer assigns the default value of 0 to one of the levels of each factor (chosen alphabetically) and compares it with the other level(s). The level that was assigned the default value was Old. The Estimate of Young is (positive) 3.072, which indicates that young B speakers are more likely to converge on *-ah* than those who are old.

Therefore, it can be stated that younger Bs, regardless of their tribe or level of contact, are leading the change towards *-ah*. It is not clear whether Bs are converging on the S variant *-ah*, or they are simply levelling the paradigm of *-ih* with the perfective paradigm *-ah*. Note that although *-ih* is used in more environments than *-ah* (see 5.5.1 above), the levelling is more plausibly from *-ih* to *-ah* because the former is not phonetically acceptable with perfective verbs (e.g., *trik-ih* 'he left it-M' < *trik-ah*); however, *-ah* can be used in all environments, as is the case in S dialect.

We do not know whether the slight shift from *-ih* to *-ah* in the B dialect can be accounted for as natural drift due to redundancy (i.e., the perfective/imperfective traits of the verb are reflected in the stem) or a sociolinguistic dialect convergence. The answer to this question will be provided in Chapter 8, based on the additional data in Chapter 7.

Table 5.8: *Results of the random effects in the model (-ih/-ah)*

Intercept	Variance	Std.Dev.	Observations	Words	Participants
Item	0.8094	0.8996	1068	431	23
Participant	3.5220	1.8767			

Table 5.9: *Results of the fixed effects in the model (-ih/-ah)*

Factors/Level	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-6.450	1.777	-3.629	0.000284 ***
Age-Young	3.072	1.425	2.157	0.031023 *
Tribe-Dosari	1.210	1.245	0.972	0.331034
Tribe-Qahtani	-1.546	1.915	-0.807	0.419450
Contact-Low	-1.029	1.669	-0.616	0.537758

Signif. codes: 0 '***'   0.001 '**'   0.01 '*'   0.05 '.'   0.1 ' ' 1
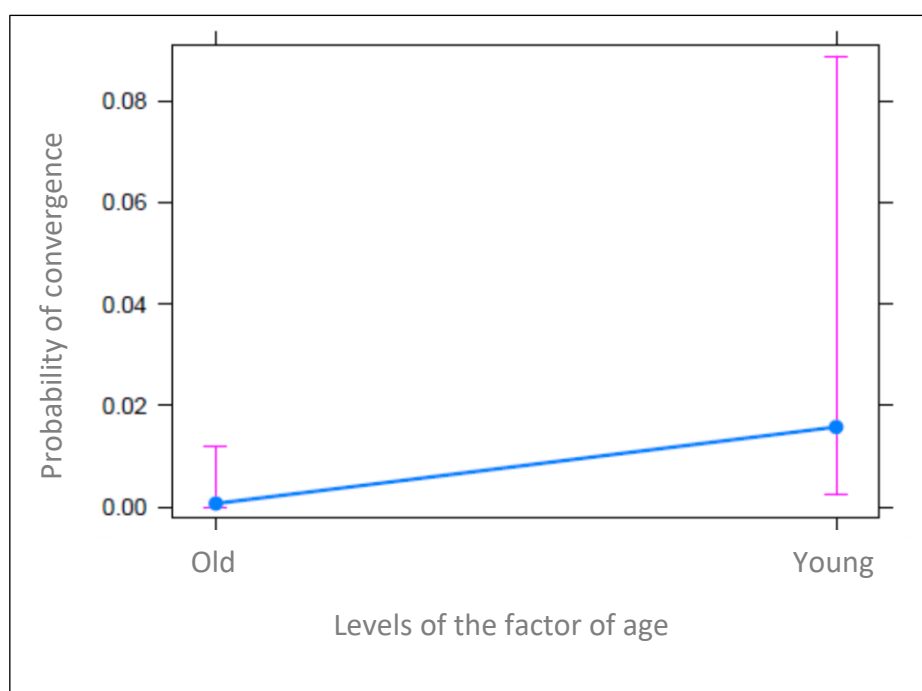
Figure 5.8: *Predicted probability of convergence on the S variant -ah by the factor of age. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (Significant)*
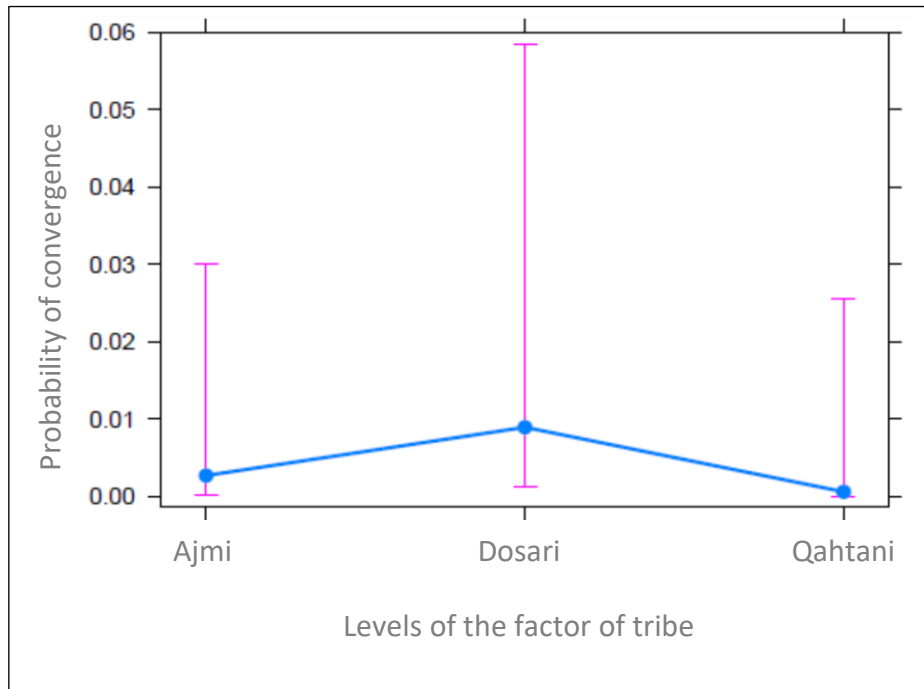
Figure 5.9: *Predicted probability of convergence on the S variant -ah by the factor of tribe. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (No significant difference)*
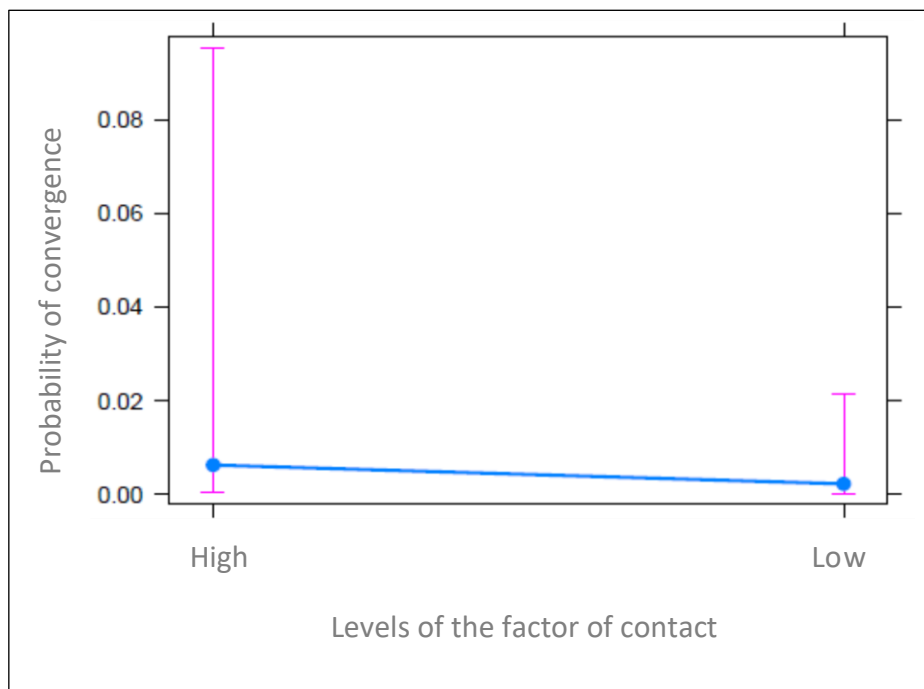
Figure 5.10: *Predicted probability of convergence on the S variant -ah by the factor of contact. The vertical bars show standard error, with 95% confidence interval, while the dots represent the mean. (No significant difference)*

5.7.2 Interview data for WBs

Similar to Bs, WBs have been observed to use *-ih* in all environments, while *-ah* is used with perfective verbs. To reiterate, this group is not the focus of the research, but 3 speakers from WB tribes (Subai', Suhul and Itban) have been interviewed for the sake of comparison. The data for WB speakers is provided in Table 5.10.

Table 5.10: *Interview data for WB speakers (-ih/-ah)*

Speaker/Tribe	Age	Frequency		Percentage	
		-ih	*-ah*	*-ih*	*-ah*
Subai'	20	54	0	100%	0%
Suhul	30	16	0	100%	0%
Itban	55	12	0	100%	0%

Although the data elicited from these speakers is not sufficient in terms of either number of speakers or tokens, we can claim, supported by the researcher's observations, that *-ih* is one of the linguistic features that WB dialect shares with the B dialect. It is not reported in the literature why WB tribes share this feature with B tribes that are originally from the south-west. We do not know whether it is a coincidence, or it is historically attributed to a period when there was heavy contact between the two groups.

5.7.3 Interview data for Ss:

The Ss are hypothesised to use *-ah* all the time (i.e., occurrences where Bs would use *-ih*), and they are less likely to converge on the B variant. This hypothesis is based on two reasons: (a) the main direction of the convergence is found to be B > S so far and b) the S dialect's 3SM suffixed pronoun paradigms (perfective and others) employ *-ah*, and its speakers are less likely to adopt a new form (i.e., *-ih*) that is not functionally useful. As for speakers who switched to Standard Arabic, as we have seen in Chapter 4, it was unknown whether they will

switch to the Standard Arabic form, which is structurally different (refer back to Figure 5.1), or if they will continue using their own spoken variant.

13 S speakers, grouped by age and style, were interviewed. For this variable, the S interviews yielded 572 tokens with an average of 44 tokens per speaker. To reiterate, these tokens are only the ones corresponding to *-ih* in the B dialect, i.e., perfective verbs were not included. The data elicited from the interviews of S speakers is provided in Table 5.11.

Table 5.11: *Ss' use of their variant -ah in the interviews*

Participant	Age	Style	S tokens	Percentage
1	Old	Informal	49/49	100%
2	Young	Informal	50/50	100%
3	Young	Informal	39/39	100%
4	Young	Informal	35/35	100%
5	Young	Informal	39/39	100%
6	Young	Informal	34/34	100%
7	Old	Informal	43/43	100%
8	Old	Informal	67/67	100%
9	Old	Informal	28/28	100%
10	Old	Formal	46/46	100%
11	Old	Formal	36/36	100%
12	Young	Formal	55/55	100%
13	Young	Formal	51/51	100%

As expected, the Ss used *-ah* consistently throughout their interviews, and no traces of other variants were found. As for the speakers who switched to Standard Arabic, they continued to use their spoken variant *-ah*, and not Standard Arabic *-hu* because *-hu* would require a different stem structure (see Figure 5.1) from that required for *-ah*. Consider the Najdi S verb *yi-gsim-ah* 'he splits it' compared to what speakers who switched to Standard Arabic used (*ya-qsim-ah*). These speakers switched to *ya-* which is the Standard Arabic form (more in

Chapter 6) and used Standard Arabic /q/ instead of /g/, but they did not use *-hu* because it would require the stem to be *ya-qsimu-hu*. The possible explanation for not using *-hu* is that it requires more attention to speech, while these speakers switched to Standard Arabic as a subconscious reaction to the formality of the interview. That is, if these speakers were requested to read a passage in Standard Arabic, they would most likely use *-hu* because they would be paying more attention to speech. More discussion of this subject is provided in Chapter 8 (section 8.5).

5.8 Conclusion

The 3SM suffixed pronoun in modern Arabic dialects can be realised as *-Vh* or *-V*. The former is generally used in B dialects and dialects of the Arabian Peninsula (both S and B), while the latter is used by most S dialects of the Levant and the West. The epenthetic vowel in the Arabian Peninsula can be /a/, /i/ or /u/. In Najd, B and WB dialects use *-ih* (and *-ah* in perfective verbs) while S dialect speakers use *-ah*.

The interview result for the Bs features frequent use of B variant *-ih* with slight convergence, led by young speakers, towards *-ah*. As for WB dialect, it is found to share the same form (*-ih*) with the B dialect. The S group on the other hand show no change at all. This is true even for speakers who switched to Standard Arabic during the interview.

Chapter Six: Imperfect Prefixes

(*ya-, ta-, na-* / *yi-, ti-, ni-*)

This chapter examines morphophonemic variation in imperfect prefixes in Najdi Arabic. The vowel in these suffixes is realised as /a/ in B dialect, giving –*ya-* (3M), *ta-* (2C, 3F) and *na-* (1PC), and as /i/ in S dialect, giving *yi-*, *ti-* and *ni-*. This variation was noted in ancient dialects of the Arabian Peninsula by early Arab grammarians, e.g., Sibawayh (Grand' Henry, 2009, p. 430), and as we will see, it has been suggested that the variation has been preserved in the contemporary Arabic dialects.

Unlike the other two variables discussed in Chapter 4 and 5, this variation was explicitly associated with the Najdi B and S dialects in the literature (Ingham, 1994). Since the variable is already established, our goal here is just to examine the presence or otherwise of convergence.

The first part of the chapter will provide an overview of verb conjugations in Arabic, discuss the history of the variable in ancient Arabic dialects and review the imperfect prefixes in contemporary Arabic dialects. In the second part, the results of the data analysis for this variable are presented.

6.1 Verb Conjugations in Arabic

Classical Arabic, as well as modern Arabic dialects, exhibits a rich subject-verb agreement system. Verbs in Arabic, either in perfective or imperfective aspect, are inflected with suffixes, prefixes, or both, which indicate the person, number and gender of the subject. Because of this rich subject-verb agreement, Arabic (as well as other languages with such

systems) allows the optional omission of the subject pronoun, as in (1) compared to (2) (Agameya, 2009, p. 356).

(1) *ðahab-t-u   ?ila   as-suːq*

went-1.S.C   to   the souq

I went to the souq.

(2) *?ana   ðahab-t-u   ?ila   as-suːq*

I   went-1.S.C   to   the souq

I went to the souq.

Verb conjugations (agreement markers) are in the form of suffixes in the perfective aspect and prefixes and suffixes in the imperfective aspect (Table 6.1). In the perfective aspect suffixes indicate person, number and gender, while in the imperfective, prefixes indicate person and gender, and suffixes indicate number. It should be noted that in Classical Arabic, as well as some modern Arabic dialects (as we shall see), the quality of vowel in these markers may vary according to verb form (Table 6.2). The verb conjugations in Table 6.1 are those of verb form I (*faʕal*), which is the basic and most frequent form in Arabic, while in Table 6.2 the verb conjugations of the other verb forms are provided (the root *fʕl,* which has the underlying meaning 'to do', is used). The vowel of the imperfect prefix is realised as either /a/ or /u/ according to the verb form to which they are attached. Additionally, in verb forms V-VIII and X, the prefixes are realised with /a/ in active voice and with /u/ in passive voice. For example, the verb forms VIII and X are realised as *ya-ftaʕil* and *ya-stafʕil* in active voice and as *yu-ftaʕal* and *yu-stafʕal* in passive voice.

As early as the pre-Islamic period, imperfect prefixes have been noted by early Arab grammarians (e.g., Al-Farra' and Sibawayh) to be pronounced with /a/ in some dialects and

with /i/ in others, and they referred to this variation as *taltala* (Nassar, 1956, p. 77-78: cited in

Ibrahim, 2008). Discussion of the history of this variation is provided in the next section.

Table 6.1: *Verb conjugations in Classical Arabic (qatal 'to kill'), (Larcher, 2009, p.639)*

	Perfect conjugations	Gloss	Imperfect conjugations	Gloss
1SC	*qatal-tu*	'I killed'	*ʔa-qtul*	'I kill'
2SM	*qatal-ta*	'you killed-SM'	*ta-qtul*	'you kill-SM'
2SF	*qatal-ti*	'you killed-SF'	*ta-qtul-i:*	'you kill-SF'
3SM	*qatal-a*	'he killed'	*ya-qtul*	'he kills'
3SF	*qatal-at*	'she killed'	*ta-qtul*	'she kills'
2DC	*qatal-tuma:*	'you killed-D'	*ta-qtul-a:*	'you kill-D'
3DC	*qatal-a:*	'they killed-D'	*ya-qtul-a:*	'they kill-D'
1PC	*qatal-na:*	'we killed'	*na-qtul*	'we kill'
2PM	*qatal-tum*	'you killed-PM'	*ta-qtul-u:*	'you kill-PM'
2PF	*qatal-tunna*	'you killed-PF'	*ta-qtul-na*	'you kill-PF'
3PM	*qatal-u:*	'they killed-M'	*ya-qtul-u:*	'they kill-M'
3PF	*qatal-na*	'they killed-F'	*ya-qtul-na*	'they kill-F'

Table 6.2: *Imperfect verb conjugations of other verb forms in Classical Arabic (Larcher, 2009, p. 641)*

Verb form	Prefixes	Imperfect verb
II	*ʔu-, yu-, tu-, nu-*	*yu-faʕʕil*
III	*ʔu-, yu-, tu-, nu-*	*yu-fa:ʕil*
IV	*ʔu-, yu-, tu-, nu-*	*yu-fʕil*
V	*ʔa-, ya-, ta-, na-*	*ya-tafaʕʕal*
VI	*ʔa-, ya-, ta-, na-*	*ya-tafa:ʕal*
VII	*ʔa-, ya-, ta-, na-*	*ya-nfaʕil*
VIII	*ʔa-, ya-, ta-, na-*	*ya-ftaʕil*
IX	*ʔa-, ya-, ta-, na-*	*ya-fʕall*
X	*ʔa-, ya-, ta-, na-*	*ya-stafʕil*

6.2 The History of Taltala in Arabic

In *Lisan Al-Arab* (1883), one of the best known Arabic dictionaries by Ibn Manẓūr, *taltala* has been defined as "*kasir* of the imperfect /t/", i.e., the production of the 2nd person imperfect prefix as *ti-* (instead of *ta-*); however, Sibawayh described the phenomenon as the production of all imperfect prefixes, not only *ta-*, with /i/ instead of /a/ (Jabri, 2007, p. 66). Sibawayh, in *Al-Kitab*, suggested that taltala is found in all ancient Arabic dialects except the dialect of Hijaz, while others (e.g., Abu Hayyan) associated this feature with the tribes of Bahra', Qayis, Tamim, and Rabi'ah (Jabri, 2007, p. 66). A more thorough review of this feature by Grand' Henry (2009) based on Rabin (1951, p. 60) states that:

> "a group of tribes who lived in a large area from Egypt to Iraq, the Quḍāʿa, had the full *taltala*. Some other tribes had a partial *taltala* only, i.e., they used *yi-, ti-, ʔi-, ni-* for the different persons of imperfect and imperative, but for the 3rd person masculine singular, they had *ya-*: Qays, Tamim, 'Asad Rabiʿa, and ʿāmmat al-ʿArab without further definition (the situation is not clear as far as Tayyi' are concerned). On the other hand, all tribes who lived in al-Hijaz, the 'Azd Sarāt, a part of the Hawàzin, and a part of the Huḏayl, in addition to the tribes from Yemen, had the imperfect preformative in *ya-* (Rabin 1951:61, 158)" (Grand' Henry, 2009, p. 430)[26]

As noted in Fischer (2006, p. 399), taltala, among several features that demonstrate that dialectal variation existed in pre-Islamic Arabic, was rejected by Arab grammarians during the standardisation of Arabic. However, there are traces of this feature in the readings of the Qur'an (Ibrahim, 2008, p. 14).[27] In other words, the vowel /a/ in the imperfect prefixes of Classical Arabic is not the 'correct' form per se; but during the standardisation of Arabic, /a/ was selected over /i/, although the latter was more common in the Arabic dialects at the time,

---

[26] This review is quite relevant to the research as it explicitly lists the dialects of the ancient Arabic tribes that inhabited the Arabian Peninsula, from which the current B and WB tribes descend, as to whether they used *taltala* or not. More on this topic will be discussed under section 6.5.

[27] The readings of the Qur'an are the way in which each reciter (of seven reciters recorded) had delivered the Qur'an to his tribe/city.

because it is in the dialect of Hijaz (the dialect of the Prophet Mohammed, peace be upon him).

The Arabic koine (discussed in section 2.2), which is believed to be an intermediate stage in the evolution of modern Arabic dialects from Classical Arabic, is said to employ taltala (/i/); and therefore, all dialects outside the Arabian Peninsula have /i/ (Ferguson, 1959a; Versteegh, 1984). To elaborate further, the Arabic (military) koine emerged in Egypt and the Levant during the Islamic conquest, exhibited /i/ and consequently all contemporary dialects in the West and Levant have /i/ in imperfect prefixes. As for dialects within the Arabian Peninsula, the contrast between /a/ and /i/ still exists, and in the following paragraph, we will tackle the possibility of whether the current contrast is linked to the ancient feature of taltala.

Versteegh (2001, p. 49) states that the pre-Islamic feature of taltala is preserved in the contemporary Arabic dialects. However, there is controversy around the evolution of this feature, i.e., how the theory of the Arabic koine and internal language change are interweaved (Ferguson, 1989, 6-7; Versteegh 1984, p. 23; Al-Sharqawi, 2017, p. 158). When tracing the dialects of the ancient tribes from which the current Najdi tribes descend, it can be claimed that the current contrast is a continuation of the ancient taltala. In section 1.3.4, we saw that the tribes of the Arabian Peninsula can be traced back to their ancestral tribes in the pre-Islamic era, thanks to the Arabs' interest in genealogy and tribal affiliation. As we shall see in section 6.5, the presence/absence of taltala in the dialects of ancient tribes, from which the current speakers (B, WB, S) descend, matches the current dialects.

6.3 The Imperfect Prefixes in Modern Arabic Dialects

The major changes to verb conjugations, as well as personal pronouns, in modern Arabic dialects are the complete loss of dual number and gender in plural forms (refer to Table 6.1)[28]. Other changes are mostly involved with the vowels of the prefixes/suffixes. In some modern Arabic dialects, verb conjugations, and more importantly prefixes, may vary according to (a) person, (b) verb form or (c) vocalic pattern of the verb:

(a) In Algiers Arabic, for example, the 3SM prefix is *yi-* while the others are *tə-* and *nə* (Boucherit, 2006, p. 65). In Moroccan Arabic, the 2SF and plural forms have no vowel, e.g., *t-kətbi* 'you write-2SF', *n-kətbu* 'we write', *t-kətbu* 'you write-PC', *y-kətbu* 'they write' (Caubet, 2008, p. 281).

(b) Just as in Classical Arabic, some dialects use different prefixes with different verb forms (refer to Table 6.2). In Tunisian Arabic (and several other dialects), the prefixes are reduced (i.e., no vowel) with verb forms II and III, e.g., *y-ʕarris* 'to marry', *y-ʃaːrik* 'to share' (Gibson, 2009, p. 568).

(c) In Amman dialect, the perfective templates of verb form I can be CaCaC (always transitive) or CiCiC (can be both transitive and intransitive). In the imperfective aspect, prefixes of the first template contain /u/, as in *yu-drus* (< *daras*) 'to study', while the prefixes of the second contain /i/, as in *yi-fham* (< *fihim*) 'to understand' (Al-Wer, 2007a, p. 514).

When reviewing the dialects of urban centres in the Arab world (Table 6.3), the forms of imperfect prefixes support Ferguson's Arabic koine theory. The vowel of the prefixes is either

---

[28] Gender in plural forms is retained in Negev and Najdi Arabic (see Blanc, 1970, p. 130; Ingham, 1982a, p. 74; cited in Vicente, 2008, p. 585).

/i/ or /ə/ in most non-Peninsular dialects (with few exceptions)[29], while in the Arabian Peninsula the contrast between /a/ and /i/ still exists and is linked to the B/S dialect distinction. The preservation of the B vowel /a/ implies that /i/ in the S dialects of the Arabian Peninsula is most likely also preserved in their ancestral dialects rather than adopted from the Arabic koine or other dialects that employ /i/.

Table 6.3: *The imperfect prefixes in modern Arabic dialects*

Part of the Arab World	Dialect	Classical Arabic			Reference
		ya-	*ta-*	*na-*	
West	Algiers	*yi-*	*tə-*	*nə-*	Boucherit (2006, p. 65)
	Cairo	*yi-*	*ti-*	*ni-*	Woidich (2006, p. 330)
	Morocco	*yə-*	*tə-*	*nə-*	Caubet (2008, p. 281)
	Tripoli	*yə-*	*tə-*	*nə-*	Pereira (2009, p. 554)
	Tunis	*yi-*	*ti-*	*ni-*	Gibson (2009, p. 568)
	Khartoum	*ya-*	*ta-*	*na-*	Dickens (2007, p. 564)
Levant	Baghdad	*yə-*	*tə-*	*nə-*	Abu-Haidar (2006a, p. 228)
	Beirut	*yə-*	*tə-*	*nə-*	Naïm (2006, p. 282)
	Damascus	*yə-*	*tə-*	*nə-*	Lentin (2006, p. 551)
	Amman	*yu-/yi-*	*tu-/ti-*	*nu-/ni-*	Al-Wer (2007a, p. 514)
	Palestine	*bi-/b(y)i-*	*bti-*	*bni-*	Shahin (2008, p. 533)
Arabian Peninsula	Bahrain	*ya-/yi-*	*ta-/ti-*	*na-/ni-*	Holes (2006, p. 253)
	Kuwait	*ya-/yi-*	*ta-/ti-*	*na-/ni-*	Holes (2007, p. 618)
	Oman	*ya-/yi-*	*ta-/ti-*	*na-/ni-*	Edzard (2008, p. 488)
	Sana'a	*yi-*	*ti-*	*ni-*	Watson (2009, p. 114)

Here are several remarks on the table (note that the remarks are from the sources cited in the table):

---

[29] In Khartoum (Dickens, 2007, p. 564), Nigerian Arabic (Owns, 1998, p. 74) and Upper Egypt (Miller, 2005, p. 921) the vowel of the imperfect prefixes is /a/.

- In Palestinian Arabic, /b/ is added to *ti-* and *ni-*, but in *yi-* the /y/ can be dropped. This form is not unique to this dialect, but in Damascus Arabic similar forms co-exist with the forms given in the table above.

- In Bahrain, Sunni Bs use /a/ with 2SM (*ta-ktib* 'you write'), 3SM (*ya-ktib* 'he writes'), 3SF (*ta-ktib* 'she writes') and 1PC (*na-ktib* 'we write'), but they use /i/ with 2SF (*ti-ktibi:n* 'you write'), 2PC (*ti-ktibu:n* 'you write') and 3PC (*yi-ktibu:n* 'they write'). The Shia S group, on the other hand, use /i/ in all prefixes.

- In Kuwaiti Arabic, the same applies as in the Sunni B dialect in Bahrain. Additionally, in verbs with the imperfect template CCaC (as opposed to CCiC), the vowel of the imperfect prefixes is always /i/, e.g., *yi-lbas* 'to wear'.

- In Oman, like Najd, the variation between /a/ and /i/ is linked to B and S dialects. Bs are reported to use /a/ while Ss use /i/.

As for B dialects outside the Arabian Peninsula, there are no traces of /a/, with very few exceptions. For example, in some B dialects of the Sinai desert /a/ is used in harmony with imperfect verbs that have /a/ in the stem (De Jong, 2009, p. 247). In the rest of B dialects (non-Peninsular), the vowel is either /i/ or /u/ in Sinai (ibid., p. 247), or /ə/ as in Hassaniyya dialect (Taine-Cheikh, 2007, p. 246).

6.4 The Imperfect Prefixes in Saudi Arabia

In Saudi Arabia, the imperfect prefixes vary according the tribes' genealogy (i.e., whether their ancestors used /a/ or /i/) and heavy contact with other dialects, rather than geographical location. Also, sometimes the prefixes vary according to the vocalic pattern of the verb, as in Amman dialect. In the review that follows, it is clear that /a/ is used by tribes that preserved it from their ancestors, and the same applies to /i/.

In the south-western region, the imperfect prefixes are reported with /u/ and /i/: the former is used with the verb pattern Cu-CCuC, e.g., *yu-ktub* 'he writes', while /i/ is used with the patterns Ci-CCiC and Ci-CCaC, e.g., *ti-ksir* 'she breaks', *ni-lbas* 'we get dressed' (Prochazka, 1988, p. 32-35). These forms, however, are probably S because /a/ is reported in the dialect of southern B tribe of Qahtan, a branch of the current Najdi B Qahtani tribe (Al-Qahtani, 2015).

In Hayil and Qassim, the genealogy of the contemporary tribes determines the existing forms. According to Abu-Hasan (2001) and al-Sharkawi (2017), Hayil and Al-Qassim are inhabited by tribes that used /a/, and therefore, the forms found in these regions are pronounced with /a/, as also reported by Prochazka (1988) and Abboud (1979). Abu-Hasan noted an exception in Al-Jiwa village in Qassim, where /i/ is used in the contemporary dialect, and attributed that to the genealogy of the inhabitants, being descendants of a tribe that exhibited taltala (/i/).

As reported by Holes (2016) there are two dialect groups in the Eastern Province: the Sunni and Shia. The first group are originally Najdi B tribes, while the second is labelled as S. The imperfect prefixes in Sunni dialect are produced with /i/ when they attach to verb stems that have /a/ (Ci-CCaC), e.g., *yi-lbas* 'he wears', and with /a/ elsewhere (Ca-CCiC, Ca-CCuC), e.g., *ta-ksir* 'she breaks', *na-rbutˤ* 'we tie'. In the Shia dialect, the vowel of the prefixes is always /i/.

Although the ancient dialect of Hijaz was claimed by Sibawayh as the only Arabic that used /a/ in imperfect prefixes (Jabri, 2007, p. 66), ironically in the contemporary dialect of Hijaz /i/ is used (Abu-Mansour, 2008). This is no surprise, since Hijaz is known for its heterogeneity (due to the holy city of Makkah) and hence the great influence of other dialects, particularly Egyptian (Ingham, 1971).
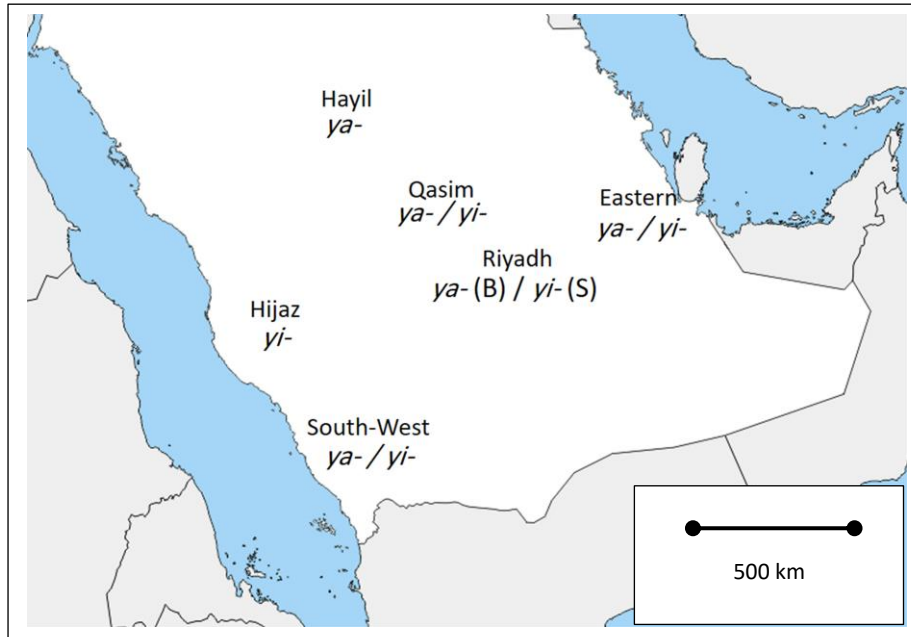
Figure 6.1: *Distribution of imperfect prefixes in Saudi Arabia (Prochazka, 1988; Al-Qahtani, 2015; Abu-Hasan, 2001; Ingham, 1994; Holes, 2016; Abu-Mansour, 2008).*

6.5 The Imperfect Prefixes in Najd

Ingham (1994) clearly states that in Najd "the Imperfective prefixes of the Active are of the form *yi-* 'he', *ti-* 'she, you' and *ni-* 'we' or *ya-, ta-* and *na-,* but following three different systems. In the dialect of Riyadh and nearby Dar'iyyah it is always *yi-, ti-* and *ni-.* In the dialect of Sudair it is *-i-* or *-a-* in dissimulative relationship with the verb stem giving *ya-msik* 'he grasps' but *yi-ṭhan* 'he grinds'. In the Northern type and in the dialects of the Central bedouins attested it is always *-a-* i.e., *ya-msik* and *ya-than*." (p. 194). In the S dialect of Sudair, which is different from the current S dialect, the alternation between /a/ and /i/ is associated with the verb pattern, as we have seen in other dialects. If the stem vowel is /i/, the prefix vowel is /a/, and vice versa. In the S dialect of Riyadh and nearby Dar'iyya (an old city which is now part of Riyadh), which can be representative of the S dialect under study, the imperfect prefix vowel is always /i/, except for the 1SC prefix, which is *ʔa-*. In Central B dialect, which Ingham defines as the dialects of fifteen tribes including Ajman, Dawasir, Qahtan (B) and Subai, Suhul, Itban (WB), the vowel of the imperfect prefixes is always /a/.

Ingham did not mention the prefixes in plural forms; however, according to the data collected for this study, they are as given in Table 6.4, which shows the full paradigm.

Table 6.4: *Imperfect prefixes in Najdi dialects, the verb ktib 'to write' is used (Ingham, 1994)*

	B and WB dialect	S dialect
1SC	ʔa-ktib	ʔa-ktib
2SM	ta-ktib	ti-ktib
2SF	ta-ktib-i:n	ti-ktib-i:n
3SM	ya-ktib	yi-ktib
3SF	ta-ktib	ti-ktib
1PC	na-ktib	ni-ktib
2PC	ta-ktib-u:n	ti-ktib-u:n
3PC	ya-ktib-u:n	yi-ktib-u:n

6.5.1 The variable in B and WB dialect

The use of /a/ in imperfect prefixes in B and WB dialects is most likely preserved from their ancient ancestral dialects. As discussed earlier in section 1.3.4, the B tribes examined in this study are originally from Yemen. Recall that Ajman and Qahtan descend from Hamadan while Dawasir descends from Azd, all of which originated from Yemen. WB tribes, on the other hand, descend from the ancient tribe of Hawazin, which also used to inhabit the region between Hijaz and Najd (Al-Hugail, 1967; Oppenheim, 2004). As mentioned earlier in section 6.2, taltala is absent in the tribes of Yemen and the tribe of Hawazin, (Rabin 1951, p. 61, 158: cited in Grand' Henry, 2009, p. 430). In other words, because the vowel of imperfect prefixes in current B and WB dialects and in the ancient dialects from which they descend is /a/, it is most likely that the modern form is preserved. The concept of linking such a feature to ancient dialects that date back more than a millennium may seem far-fetched, but in support of their adoption of the claim (e.g., Versteegh, 2001, p. 49), the data from the current study supports this hypothesis, as we will see.

6.5.2 The variable in the S dialect

The vowel of imperfect prefixes in the S dialect of Najd is always /i/, except for the 1SC (ʔa-). Unlike B tribes, it requires extensive research to trace the ancient tribes of each S tribe/family examined here because they tend to use sub-family names rather than the larger tribe (see section 1.3.4). However, according to Oppenheim (2004), the S tribes of the Arabian Peninsula were known to inhabit the central region, in which the ancient dialects were known to use taltala (Jabri, 2007). Moreover, for the Tamim tribe, an ancient tribe which still inhabit(s) the centre of the peninsula, the vowel of the prefixes is /i/ (Grand' Henry, 2009).

6.5.3 Other forms of the variable and envelope of variation

The sociolinguistic alternation between the imperfect prefix vowels /a/ and /i/ has always been associated with verb form I (imperfective template Cv-CCvC), and in this chapter we will only examine the same verb form, as it is the one exhibiting sociolinguistic variation. In this section, however, we will shed some light on the other forms of the variable, which do not show variation and thus were excluded. In the review of the prefixes in the Arabic dialects we have seen how they are affected by the structure of the stem to which they are attached. Similarly, in Najdi dialects, B and S, the vowel of the prefix can be dropped, changed or lengthened according to the stem, and each dialect may behave differently in that respect. In this section, the allomorphy of the prefixes, conditioned by the stem, is discussed, and references to dialect differences are provided. All the observations and examples below are based on the current data.

6.5.3.1 Vowel drop

In triliteral verb forms II (*y-faʕʕil*) and III (*y-faːʕil*), the vowel is dropped in both dialects, e.g., *t-ʕallim* 'she tells', *y-raːsil* 'he messaged'. In verb form V, the vowel is dropped in B dialect, but retained with changes to the stem in S. The word for 'to hang' is realised as *y-*

taʕallag in B dialect and *yi-tʕallag* (Ci-CCVCCVC) in S. For more discussion of verb forms, refer to section 1.5.3.2.

A special case of triliteral verbs are those which have /w/ or /y/ as medial consonant, also known as *hollow verbs*. The perfective form of these verbs in Classical Arabic is CV:C (e.g., *qa:l* 'he said'), while the imperfective is Ca-CV:C (e.g., *ta-mi:l* 'she bends'). In Najdi dialects the perfective form is similar to Classical Arabic, but in imperfective forms, the prefix vowel is either dropped (e.g., *y-gu:l* 'he says') or /i/ (*yi-gu:l*). Before data collection, it was hypothesised that the first form is B while the latter is S. However, it is found that both forms exist in both dialects and vary freely, though the second form is less frequent in B dialect.

In Classical Arabic, biliteral verbs have the imperfective template Ca-CVCC, e.g., *ya-mudd* 'he stretches', *ta-ħinn* 'she longs for'.[30] In Najdi dialects, the prefix vowel is either dropped, or /i/. The same hypothesis and findings for hollow verbs apply to biliteral verbs. Bs and Ss used the forms *yi-midd* /*y-midd* 'he stretches' and *ti-ħinn* / *t-ħinn* 'she longs for'.

6.5.3.2 Vowel change

In B and WB dialects, the prefix vowel in verb form I is /a/, as discussed earlier. In verb forms *yi-ffil* (IV), *yi-tifa:ʕal* (VI), *yi-ntVfaʕʕal* (XI) and *yi-ntVfa:ʕal* (XII), however, the prefix vowel is realised as /i/. Examples for these forms are: *yi-dxil* 'he causes to enter', *ni-tiðˤa:rab* 'we fight-reciprocal', *ti-ntaraggaʕ* 'it can be patched', and *ti-ntifa:ham* 'she can be negotiated with', respectively. In the S dialect the prefix vowel is also /i/, as it is verb form I.

---

[30] Biliteral verbs are those which consist of two consonants, as opposed to triliteral. Example of these verbs are *madd* 'stretched' and *radd* 'replied'.

6.5.3.3 Vowel lengthening

The prefix vowel is lengthened in two cases. The first case is when the first consonant of a triliteral verb is originally a glottal stop in Classical Arabic, e.g., *ʔkl* 'eat'. In both dialects, the prefix vowel with such verbs is a long /a/, as in B *ya:-kul* 'he eats and S *ya:-kil*. In the second case, where the first consonant is /w/ (e.g., *wzn* 'weigh', *wsˤl* 'arrive'), the vowel is lengthened, but each dialect uses different vowel. Such verbs are realised in B dialect as *ta:-zan* 'she weighs' and *ya:-sˤal*, while in S dialect they are *to:-zan* and *yo:-sˤal*. This rule resulted from the monophthongisation of the Classical Arabic diphthong /aw/, discussed in 1.5.5.4.3 and 8.3.3.

6.5.3.4 Overlap with guttural effect

Recall that in section 1.5.5.4.2, we discussed the guttural effect in initial syllables that are originally CaG in Classical Arabic (> Najdi CGa). When the first consonant of a verb is a guttural, the imperfect prefix forms an initial syllable CaG, which is then altered to CGa (e.g., e.g., *y-ħafɪr < ya-ħfɪr* 'he digs'). The guttural effect is hypothesised to be present in B dialect only, while in S dialect CiG is used, as *yi-ħfɪr*. This variable will be discussed in length in section 8.3.2, but it is mentioned here to clarify how the two variables overlap. Needless to say, tokens with gutturals are excluded in the quantification of the current variable, which is concerned with the alternation between /a/ and /i/ only.

6.6 Research Questions and Hypothesis

Based on Ingham (1994) and my own observations, the Bs tend to use /a/ while Ss use /i/ in imperfect prefixes consistently. It is hypothesised that for this variable, there is no expected convergence by either group. As for S speakers who switched to Standard Arabic (as seen in Chapter 4), they are expected to exhibit the same behaviour for this variable since the

imperfect prefixes in Classical/Modern Standard Arabic are produced with /a/. We will attempt to answer the following research questions:

- Is one group converging on the other group's variant?
- If yes, what are the social factors conditioning this convergence?

6.7 Results

6.7.1 Interview data for Bs

As mentioned in the previous chapters, the production data for Bs was collected from the interviews of 23 B speakers, grouped by age, tribe, contact and education. The interviews yielded 761 tokens, with an average of 33 tokens per speaker. Although there are three imperfect prefixes that show variation (*ya-/ta-/na-* or *yi-/ti-/ni-*), they are treated in the quantification of this variable as one (i.e., we are concerned with the vowel, not the consonant). Henceforth the 3M prefix *ya-/yi-* is used as a representative for all prefixes.

The data shows almost no convergence on the S variant, except for two tokens by two different speakers (Table 6.5). Although there is virtually no variation, the data is plotted anyway in Figure 6.2 by speaker. Note that speakers in the figure are coded by social factors as follows: Speaker number - Tribe (Ajmi, Dosari or Qahtani) - Age (Old or Young) - Education (High or Low) – Contact (High or Low). Since there are two variants for this variable (*ya-* or *yi-*), 95% use of *ya-* in the figure, for example, also indicates that the speaker's use of *yi-* is 5%.

Table 6.5: *Distribution of the variants ya- and yi- by social factor in the interviews for Bs*

| Social Factors | Levels | Frequency of variants | | | Percentage | |
		ya-	yi-	Total	ya-	yi-
Tribe	Ajmi	308	0	308	100%	0%
	Dosari	267	1	268	99.6%	0.4%
	Qahtani	184	1	185	99.4%	0.6%
Age	Old	389	1	390	99.7%	0.3%
	Young	370	1	371	99.7%	0.3%
Education	High	535	2	537	99.6%	0.4%
	Low	224	0	224	100%	0%
Contact	High	202	1	203	99.5%	0.5%
	Low	557	1	558	99.8%	0.2%

How do we account for the S tokens in the speech of these two Bs? It is worth mentioning that: one of these tokens is *yi-dx-i-l* 'to enter', which also includes the S variant /i/ instead of /u/ (discussed in Chapter 4), and the same form occurred again in the interview but with both B variants (i.e., *ya-dx-u-l*). Therefore, it can be claimed that vowel harmony is responsible for the emergence of the prefix vowel /i/ in the speech of these two speakers. In fact, vowel harmony is one of theories provided by Sibawayh to account for the feature of taltala in some dialects more than others (Jabri, 2007). Another claim is that these speakers acquired these forms as a whole unit. In this particular example, the speaker probably acquired both forms *yadxul* and *yidxil* 'to enter', and he alternates them freely. To further support this claim, when interviewing the S speakers, the researcher, as a B, used the S variant with lexical items that he does not usually use in his dialect. He used the S variant in the lexical items *yi-zʕal* 'to get mad', *ti-bʁa* 'you want' and *ti-lga* 'you find', which correspond to *y-ðˤiːg*, *ti-biː* and *t-ʕayyin* in his dialect. That is, the researcher accommodated to the S speakers by switching to entirely different lexical items that exist in his repertoire in S form only.
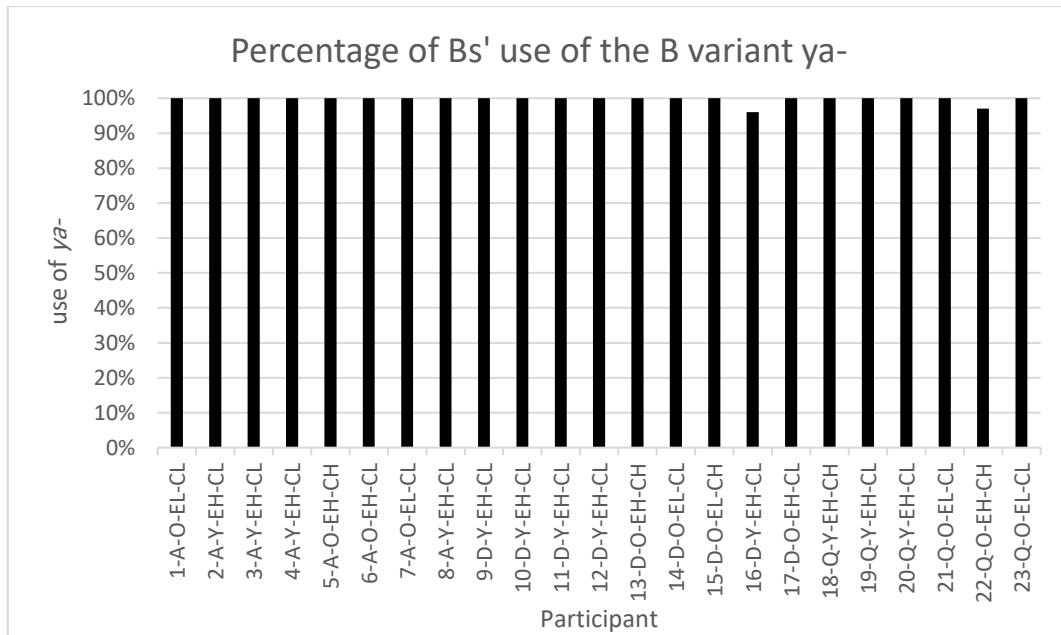
Figure 6.2: *Percentage of Bs' use of the B variant ya- by Participant.*

6.7.2 Interviews for WBs

For this variable, WBs are expected to behave like Bs, i.e., use *ya-*. As they are marginal to the study, only three speakers, each from a different WB tribe, have been interviewed. The data in Table 6.6 shows that WBs tend to use ya- all the time, though the number of tokens is less than the optimal rate of tokens per speaker.

Table 6.6: *Interview data for WB speakers (ya-/yi-)*

Speaker/Tribe	Age	Frequency		Percentage	
		ya-	*yi-*	*ya-*	*yi-*
Subai	20	13	0	100%	0%
Suhul	30	10	0	100%	0%
Itban	55	23	0	100%	0%

6.7.3 Interview Data for Ss

In this data set, 13 S speakers were interviewed, grouped by age and style. The interviews

yielded 414 tokens with an average of 31 tokens per speaker. In the quantification of the Ss'

tokens, only occurrences where Bs would have used *ya-* are calculated. Recall that Ss use *yi-*

in all verb forms, as opposed to Bs who use *ya-* with verb form I and *yi-* with others.

The Ss are expected to use *yi-* all the time. However, since imperfective verbs in Modern

Standard Arabic are produced with *ya-*, it is expected that the four S speakers who switched to

Standard Arabic (formal style) will use *ya-*, but we do not know how frequently. To reiterate,

the guide to know when those speakers switched to Standard Arabic is through the use of

particular lexical items and sounds that are not used in the spoken dialect (see section 8.5).
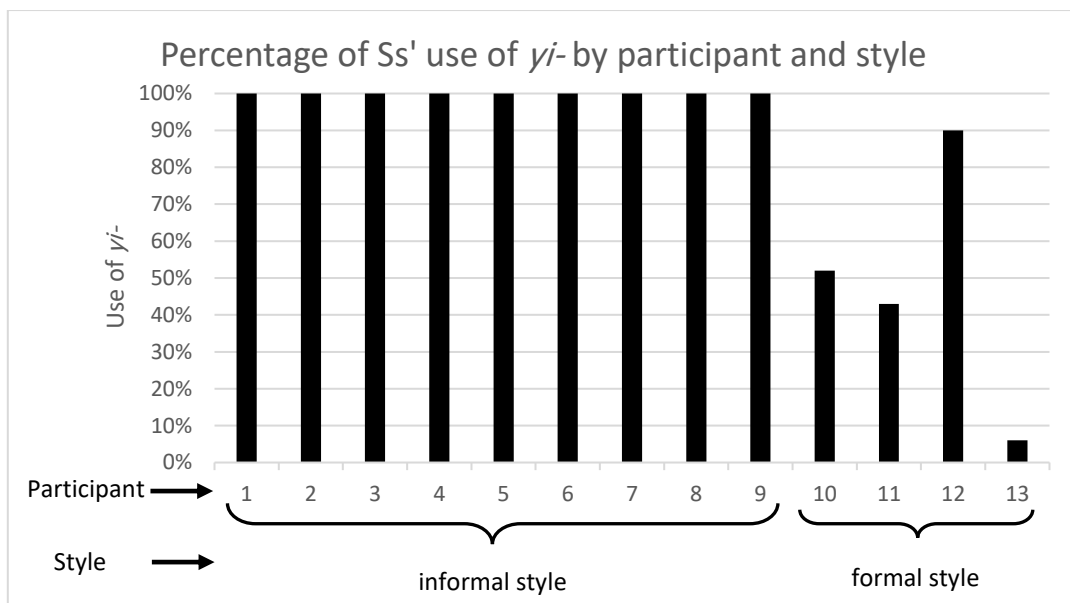
Figure 6.3: *Percentage of the Ss' use of the S variant yi- plotted by participant and style.*

Table 6.7: *Ss' use of their variant yi- in the interviews*

Participant	Age	Style	Ss.tokens	Percentage
1	Old	Informal	29/29	100%
2	Young	Informal	68/68	100%
3	Young	Informal	36/36	100%
4	Young	Informal	22/22	100%
5	Young	Informal	18/18	100%
6	Young	Informal	29/29	100%
7	Old	Informal	31/31	100%
8	Old	Informal	50/50	100%
9	Old	Informal	32/32	100%
10	Old	Formal	15/29	52%
11	Old	Formal	10/23	43%
12	Young	Formal	27/30	90%
13	Young	Formal	1/17	6%

The data shows that Ss using an informal style used their variant all the time, while those who switched to Standard Arabic used *ya-* with different rates of frequency. The two young speakers using formal style perceived the setting of the interview differently. One, who used *ya-* 3 times out of 30, thought of the interview as an informal chat, but in a few cases when discussing religious topics *ya-* was used. The other, who used *ya-* 16 times out of 17, viewed the interview as a formal discussion and kept diverging to religious topics; therefore Standard Arabic was used extensively. The other older speakers, on the other hand, altered between *ya-* and *yi-* evenly, according to the topic of conversation. More discussion of this issue is provided in section 8.5.

6.8 Conclusion

In this chapter, the presence/absence of the feature of taltala in Najd has been investigated sociolinguistically. It has been reported by Ingham (1994) that Bs use *ya-/ta-/na-* while Ss use

yi-/ti-/ni. However, since there was accommodation by Bs towards the S dialect for the previous variables, it was worth exploring whether there will be convergence here as well.

It was found that speakers of each dialect used their variant consistently and there was no convergence.[31] The reason behind this is probably attributed to the low salience of the variable (see Chapter 7).

When reviewing the imperfect prefixes across the Arab world, we find that /a/ is only used in B dialects of the Arabian Peninsula. This might indicate that /a/ is on the verge of being lost in the next decades or centuries, due to regional levelling. However, based on the history of this feature and according to the data in this study, this is not the case. First, the vowel /a/ in imperfect prefixes of the B dialects can be traced back to the ancient dialects of the tribes from which they descend. It has been preserved for more than a millennium. Second, the data in this study supports the claim that this feature is still preserved (i.e., no convergence) despite the heavy contact with the S dialect. Additional data and more discussion of this variable are provided in Chapter 7.

---

[31] An exception is when some S speakers switched to Standard Arabic and used /a/.

Chapter Seven: Salience of the Linguistic Variables

7.1 Introduction

Salience is an underdefined, elusive notion. Subjectively, it can be simplified as the level of speaker's awareness of a particular linguistic feature. Objectively, it can be defined as "the property of a linguistic item or feature that makes it in some way perceptually and cognitively prominent" (Kerswill and Williams, 2002, p. 81), or "the characteristic of being easily noticeable, prominent or conspicuous" (Siegel, 2010, p. 129). As Torbert (2004) suggests, every sociolinguist has a grasp of the notion of salience and knows which of the investigated variables are the more salient. However, the controversy lies within what makes a variable more salient than the others, what it entails, what the levels of salience are, and how it is measured.

In this chapter, we will attempt to shed some light on the role of salience in dialect contact, different models of salience and the classification of linguistic variables (based on salience). Next, the two approaches of measuring salience (*criteria-list* and *experimental*) are explored (MacLeod, 2015). The last part of the chapter deals with salience in the current study: methodology, data collection and results.

7.2 Role of Salience in Dialect Contact

'Salience' as an independent factor has been applied in studies of dialect contact (e.g., Trudgill, 1986; Nolan, Kerswill and Wright, 1991; Auer et al., 1998), language contact (e.g., Mufwene, 1991) and language learning (Bardovi-Harlig, 1987). In the aforementioned studies, linguistic features that were found to be salient were likely to be adopted or rejected by speakers of the other dialect (in case of dialect contact), favoured over other features (in

the case of language contact, including creolization and mixed languages), or acquired faster (in the case of language learning).[32]

Now we briefly discuss how salient features behaved in the studies of dialect contact above, with no reference to how their salience was determined/measured. The models of salience are discussed in the next section.

Trudgill (1986) provides evidence from the case of dialect accommodation in the city of Norwich (Trudgill, 1974) and states that "it is indeed salient features of the target variety that are adjusted to" (p. 37); he adds that features that have *extra-strong salience* are not used during the accommodation process. In Kerswill (1985) it was established that *l-vocalisation* (as in [fɪɫm] > [fɪʊm]) is salient/stigmatised in south-eastern English varieties; Nolan, Kerswill and Wright (1991) found that some speakers avoid vocalisation by hypercorrectly switching to clear [l] in the situation where there is a social pressure to speak correctly (cited in Kerswill and Williams, 2002). Auer et al. (1998) investigate "dialect accommodation by Saxonian speakers after work migration to west Germany" (p. 168) in real time. Speakers were interviewed 8 times, over a two-year period, by interviewers who used standard German. By examining 9 vocalic and 3 consonantal variables, it was found that the use of salient non-standard features decreased over time across the interviews. That is, speakers of the non-standard dialect feel the pressure to accommodate to the interviewer, who speaks standard dialect, and they adjust salient features more than the others.

7.3 Models of Salience

Before we start reviewing models of salience, the differences between *subjective* and *objective* salience (explicitly stated in Auer et al.,1998) should be clarified. Objective salience refers to the linguistic (internal) traits of a linguistic feature that contribute to making it

---

[32] See Kerswill and Williams (2002).

prominent or easily noticeable. Subjective salience, on the other hand, refers to the lay speaker's level of awareness of the linguistic variables. References to objective and subjective salience will appear in the discussion of models of salience below.

7.3.1 Trudgill (1986)

To attempt to answer the question 'why are speakers aware of some variables more than the others?', Trudgill proposes the following factors, at least one of which must be present (adapted from Trudgill, 1986, p. 11):

- There is a high-status variant of the stigmatised form, and this high-status variant is reflected in the orthography while the stigmatised variant is not.[33]
- The variable is currently involved in linguistic change.
- The variants are phonetically radically different.
- The variable is involved in the maintenance of phonological contrast.

Using Labov's taxonomy (discussed in section 7.3.4 below), Trudgill suggests that these factors can contribute to an *indicator* becoming a *marker* (i.e., salient). As stated earlier, he adds that a variable can be extra-salient and therefore behave differently from salient features. In Trudgill's model, a variable can be *not salient*, *salient* or *extra-salient*.

The major flaw of this model is that it focuses on objective internal properties of the variable. The only external factor in the model (being involved in linguistic change) is argued by Kerswill and Williams (2002) to be a result, rather than a cause, of salience. They argue that "[t]he way in which this factor can be maintained is if it is assumed that salience does not itself lead to change, and that it is the change itself that causes speakers to notice the feature involved." (Kerswill and Williams, 2002, p. 91).

---

[33] This works for e.g. Arabic or English, but not for all languages.

7.3.2 Auer et al. (1998)

Auer et al. (1998) review models of salience in Hinskens (1992), Schirmunski (1930) and Trudgill (1986), and list parameters for determining salience, distinguishing subjective criteria from objective ones (Table 7.1). They enquire whether the subjective criteria are the result of linguistic-internal traits of the features (objective).

Table 7.1: *Objective and subjective criteria for determining a variable's salience (Auer et al., 1998, p. 167)*

Objective Criteria	Subjective Criteria
Articulatory distance	Perceptual distance
Areal distribution	Usage in code-alternation
Phonemicity	Representation in lay dialect writing
Continuous vs. dichotomous	Stereotyping/mimicking
Lexicalisation	Comprehensibility

They applied the above criteria to the linguistic variables in their study of long-term accommodation by Upper Saxonian speakers, and it was found that some variables contradict the predictions made by salience. For example, the non-standard variants of (AI) and (AU) are resistant to change, despite the fact that they are salient. Conversely, non-standard forms of (A:) are lost in reading style, although the variable appeared not to be salient.

They concluded their study by answering 'no' to the earlier question of whether subjective salience is explained by objective traits of the feature. It was stated that a feature can meet all (or most) objective criteria of salience and yet does not behave like other salient variables during the accommodation process. The unpredictability of the correlation between objective salience and the behaviour of speakers (subjective) was raised earlier in their paper. It is stated that in order for a salient feature of a new dialect to be adopted, it is expected that this feature should hold some prestige in the new dialect, or that this feature is not negatively

evaluated in the accommodating dialect. This hints at including social psychological factors in the model of salience, which is suggested by Kerswill and Williams (2002).

7.3.3 Kerswill and Williams (2002)

Kerswill and Williams (2002) argue that "[d]iscussing salience in a way that divorces it from language-external factors leads to a failure to gain insights into the social patterning of linguistic features; at worst, to do so leads to circularity and labelling." (p. 107). Their model of salience (provided below) suggests that extra-linguistic factors (component 3 of the model) are "ultimately the cause of salience" (ibid., p. 109).

1- "the presence of a linguistic phenomenon whose explanation we suspect may be due to the salience of the linguistic feature or features involved. Typically, the phenomenon will be a particular pattern observed in language change, language variation, the variable behaviour of individual speakers, or the acquisition of a linguistic feature. In cases of language change and variation, the linguistic features are items being transferred from one language variety to another through diffusion; however, diffusion-type mechanisms may hold for the other types of phenomena as well.

2- language-internal explanations, such as the presence of phonological contrast, great phonetic distance, internally-defined naturalness, semantic transparency, or a particular syntactic or prosodic environment.

3- language-external cognitive, pragmatic, interactional, social psychological, and sociodemographic factors. Some have a natural link with the linguistic features being adopted (e.g., that between a syntactic feature and its pragmatic function), while others have an arbitrary relationship (e.g., the favouring of one vowel quality over another)" (ibid., p. 108)

The model is supported by a study of dialect levelling which they conducted in three urban centres in England: Milton Keynes, Reading and Hull. Their model of salience is probably the most comprehensive one. However, it is not clear how the language-external factors should be incorporated (MacLeod, 2015, p. 85)

7.3.4 Perceptual (subjective) salience

While the above models are concerned with the factors that can contribute to the salience of a linguistic feature, perceptual (subjective) salience is concerned with the different levels of speakers' awareness of a feature, and how features belonging to a particular level are expected to behave similarly. Labov (1972) categorised linguistic variables, based on the social evaluation they receive (i.e., speakers' awareness), as: *indicators*, *markers* and *stereotypes*. "*Indicators* are linguistic features which are embedded in a social matrix, showing social differentiation by age or social group, but which show no pattern of style shifting and appear to have little evaluation force." (p. 314). *Markers* "do show stylistic stratification as well as social stratification … *Stereotypes* are socially marked forms, prominently labelled by society" (ibid., p. 314).

7.4 Salience in This Study: Hypothesis and Research Questions

Salience is not usually the focus of sociolinguists when examining a case of dialect contact. However, it is usually sought, as an independent explanatory factor, to help explain why some features are favoured/disfavoured over others in production data from a dialect undergoing change. In the current study, the production data of Bs for the three variables (Chapters 4, 5 and 6) shows considerable difference in terms of their convergence to the S forms, and it is the level of salience that is expected to explain this discrepancy.

By collecting additional data, the purpose of which is to quantify the relative salience of variables (discussed in the next section), we expect that the level of salience of the linguistic

variables will agree with the production data. That is, as we saw in Chapter 4 the rate of Bs'
use of the S variant /i/ (as opposed to /u/) is higher than their use of S variants of the other
variables, i.e., -ah (Chapter 5) and yi- (Chapter 6). Therefore, it is expected that the merger
variable (/u/ - /i/) is salient. It is hypothesised that the S variant of the merger has prestige
value.

We aim to answer the following questions:

- Are any of the examined linguistic variables salient?
- If yes, do salient variables behave differently in the production data? (i.e., do salient
  variables show a higher rate of convergence?)
- What are the classifications of the linguistic variables according to Labov's
  taxonomy?

7.5 Methodology and Data Collection

MacLeod (2015) states that there are two approaches to operationalising salience: the *criteria-
list approach* and the *experimental approach*. The former represents the previous salience
models discussed above. In this approach, a set of criteria is applied to linguistic variables to
help evaluate their relative salience. In the latter, however, "the salience of a variable can be
measured by the extent to which the presence of the variable contributes a particular social
meaning to listeners." (MacLeod, 2015, p. 85). It can be stated that the criteria-list approach
incorporates both objective and subjective salience, while the experimental approach focuses
on subjective perceptual salience. Next, we shall review the experimental approach.

The experimental approach involves audio stimuli containing the target feature, and listeners
are asked to identify the dialect or social group of the speaker. Stimuli can be re-synthesised
(for example by acoustically altering the F1 and F2 of vowels, as in Graff et al., 1986),
spliced (simply replacing a variant with the target variant, as in Campbell-Kibler, 2009), or

natural (extracted from authentic speech, as in Llamas et al., 2016). The listeners are asked to record their response in the form of a scale (e.g., Graff et al., 1986) or forced choices (e.g., Llamas et al., 2016). The percentage of correct identification of a dialect through a single linguistic feature can reflect the relative salience of that feature. The higher the percentage of correct responses is, the higher the salience. An additional criterion, reflecting the level of salience in such experiments, is the speed with which speakers record their judgment (Llamas et al., 2016). Arguably, the shorter the response time is, the higher the salience. To sum up, the experimental approach implements either a simple identification task, in which only the percentage of correct responses is measured, or else a more advanced task, in which both percentage of correct responses and response time are measured.

Another, but rather different, methodology which offers insights into the relative salience of a variable is evoking short-term accommodation (in the form style-shifting) by having multiple interviewers, each of whom speaks a different variety, interview the same speaker (Llamas et al., 2009). How is short-term accommodation related to salience? As suggested by Trudgill (1986), it is usually the salient features that are adjusted during accommodation. The premise of this method (used by Llamas et al., 2009), is that if the production patterns of a speaker are analysed across interviews, each of which has been conducted by a different interviewer speaking different variety, it is expected that the rate of use of salient features will increase (converge) or decrease (diverge) according to the interlocutor's variety (Giles, 1984).

In the following sub-sections, we will review and discuss: (a) the application of the criteria-list approach to the current variables, (b) the experimental approach and (c) short-term accommodation in multiple interviews.

7.5.1 Application of the criteria-list approach to the current variables

The approach to measuring the relative salience of variables in this study is mainly experimental; however, it is worthwhile attempting to apply the criteria-list approach to the current variables in order to see which criteria they meet. The list of criteria below is adapted from Auer et al. (1998). The criteria that are difficult to incorporate (or inapplicable) are excluded.

Table 7.2: *Application of the criteria-list approach to the examined variables (criteria list is adapted from Auer et al., 1998)*

	Criteria	Merger (/u/ - /i/)	3SM pronoun (-ih / -ah)	Taltala (ya- / yi-)
	phonetic difference	✓	✓	✓
Objective	Maintenance of phonological contrast	✓	X	X
	areal distribution	✓	X	X
	usage in code-alternation	✓	X	X
Subjective	representation in the orthography	X	X	X
	stereotyping/mimicking	X	X	X
	comprehensibility	X	X	X

For the first criterion (phonetic difference), the merger is a lexical (phonological) variable, while the others are morphophonological. The phonetic difference between the variants of the merger is reflected in the phonetic traits [+back] for /u/ and [-back] for /i/. As for the other two variables, the differences between the epenthetic vowels [i] and [a] is reflected in [+high] for /i/ and [-high] for /a/. The second criterion, involvement in the maintenance of a phonological contrast, applies only to the merger because it is found in the lexical root, as opposed to the other variables. However, the words that are affected by the merger are limited

(as discussed in section 4.2.1). The third criterion, areal distribution, applies to the merger but not the other variables. As discussed in the chapters for each variable, the merger has been completed in S and WB dialects, while in B dialect it is in progress. This means that the S variant /i/ is widespread (i.e., used by Ss and WBs) compared to /u/ which is used by Bs only. For the other variables, the S variants are not widespread; they are used by Ss only. The fourth criterion (usage in code-alternation) applies to the merger only. As we have seen in the previous chapters, only the merger shows a considerable convergence on the S variant. The last three criteria do not apply to all variables. None of the variables is reflected in the orthography, is subject to stereotyping/mimicking or hinders comprehensibility.

It can therefore be stated that, based on the criteria-list approach, the merger is relatively salient as it meets four of the criteria, while the other variables meet only one.

7.5.2 The experimental approach in the current study

As discussed above, the experimental approach to salience depends on listeners' identification of a dialect or social group after listening to a word containing the target feature, produced by a model speaker. The reaction time of the responses can be also used as an indication of salience, along with the percentage of correct associations. This experimental design in sociolinguistics, developed by Llamas et al. (2016), is referred to as the *Social Category Association Test* (SCAT). More discussion of the test is provided in the next sub-section.

The SCAT was planned to be the only method by which salience would be tested in this study. The participants in the main production data of this research were asked to participate in the experiment after the interview. However, due to the following practical reasons, the number of participants in the SCAT ended up being only 10 Bs:

1. The time demand on some of the participants was already heavy, and they did not take part in the test.

2. Some of the older participants found it difficult.

3. Some participants mixed up the choices (i.e., they pressed the button assigned to the other group on the response recording tool). See section 7.5.2.1.2 below for more details of the response recording tool.

4. Others took a long period of time to record their response to most stimuli, making their responses unusable, as response time is crucial.

5. The test was run for both groups (Bs and Ss). However, the number of Ss whose data was usable was only two. Therefore, their data was excluded as it was disproportionate to that of the Bs.

Because the number of participants in the SCAT was low, a simple identification task, in which only percentage of correct association is calculated, was run via the online survey tool *Qualtrics*. This simple online survey, discussed below, is considered supplementary to the SCAT.

7.5.2.1 SCAT

7.5.2.1.1 Review

The SCAT is an adaptation of the *Implicit Association Test* (IAT), developed by Greenwald et al. (1998). IAT is a psychological test designed to measure *implicit attitude*, i.e., automatic (sub-conscious) evaluation, about various concepts (Greenwald et al., 1998, p. 1464). In this experimental design, the percentages of correct association and latency (measured in milliseconds) are calculated. It is found that longer latency is correlated with higher error rate in the association, and vice versa.

Using the same design, Llamas et al. (2016) innovated the SCAT "to examine speakers' perceptions of social category associations and of the geographical and social distributions

of key phonetic forms in the [Scottish/English] border region" (p. 7). In their words, the SCAT "rests, firstly, on the use of community consensus concerning the social categories that listeners associate with phonetic forms as a measure of the salience of those forms. Secondly, the speed with which subjects respond when making the association between a form and a social category is treated as an indicator of the association's strength, and therefore of the degree of salience of the form in question." (ibid., p. 15).

7.5.2.1.2 SCAT in the current study

SCAT is adopted in the current study and was designed using *PsychoPy* software (Peirce et al., 2019). The components of the experiment are as follows:

- *Stimuli*: following Llamas et al. (2016), the audio stimuli were all authentic recordings of Bs and Ss. Stimuli recorded by the same speaker were not played immediately after each other. Each word contains one variant only. All efforts have been made to make sure that there is no other variant (at least from the present study) in each stimulus. Since most variants in the present study (including the ones examined qualitatively) are vowels, the words were selected carefully to make sure that no more than one vowel (not variant) besides the variant is present in each word. One might argue that the stimuli may contain other elements, but to my knowledge, there are not any.

- *Visual background*: in order not to distract participants, the visual background, shown whilst the stimulus is played, was plain dark grey. The order number of the stimulus appeared in the middle of the screen. The words were not written on the screen because they are in the spoken dialect.

- *Response recording tool*: the way in which responses are recorded in such experiments is crucial since response time is considered. In Greenwald et al. (1998), a regular computer keyboard was used. The participants were trained on the keys and their

corresponding answer. Additionally, a visual reminder with labels, showing the keys and their corresponding answers, remained on the screen during each block of the experiment. In the current experiment, a more efficient tool, instead of the computer keyboard, was used. It seems to me that if the computer mouse is held with both hands, with each thumb above a button, it would be easier to record responses than on a keyboard. A visual reminder (Figure 7.1) showing the mouse buttons and the corresponding answers appeared immediately after the stimuli and lasted until the response is recorded.

- *Results*: after the experiment is finished, an Excel file of the results was automatically generated. The button pressed for each stimulus and the response time were both recorded.
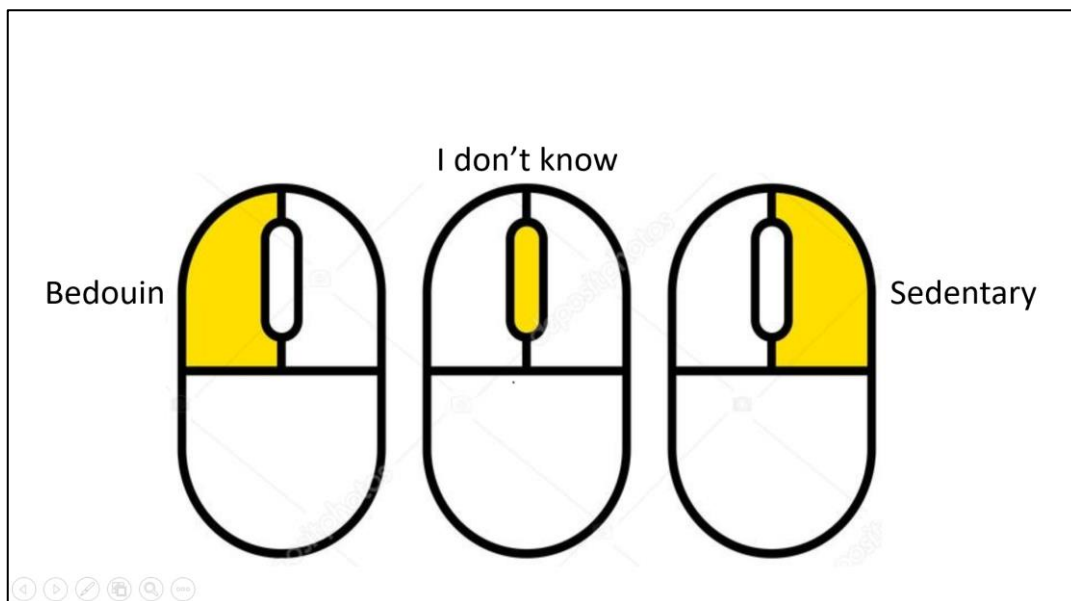
Figure 7.1: *The visual reminder used in SCAT.*

Now we turn to the design and procedures of the experiment:

1. The order number of the stimulus appears in the middle of the screen.

2. After 1 second, the stimulus is played.

3. Once the stimulus ends, the recording response time begins, along with the visual reminder.

4. Once the response is recorded, the visual reminder disappears.

5. After 1 second, the number of the next stimulus appears, and this process is repeated.

As for conducting the experiment, the procedures and the purpose of the experiment were explained to the participants, and all their questions were answered. Then they were trained how to record their responses, using the mouse. Before starting the experiment, *PsychoPy* requires the name of the participant to be entered; this is used as the name of the results file. In our case the names of the participants were coded. After that, they performed two trial questions before beginning the experiments. It was emphasised to them that they would not be able to listen to the stimuli more than once, and that once a response is recorded, it cannot be changed.

As mentioned earlier, for several practical reasons the number of participants who took part in this experiment is only 10 Bs. Therefore, the need for the following identification task arose.

7.5.2.2 Identification task

This task is supplementary to the SCAT. Using the same stimuli as the SCAT, an online survey (in the form of a questionnaire) was designed and distributed via *Qualtrics*. The number of participants in this task is 34: 14 Bs, 10 Ss, and 10 from other groups.

Since it was not possible to monitor participants while conducting the experiment, response time was not recorded. Similarly to previous studies that implemented this type of task, we are interested in the percentage of correct associations. A certainty scale, however, was also presented after each question.

Now we turn to the design of the task. First, the questionnaire was explained using simple language. Secondly, information about the participants is gathered: group (B, S or Other), age, and education. Then, the first question is presented. Each question includes an audio play button to start the stimulus, and a question asking: *is this person?* The choices were: *Bedouin*, *Sedentary*, *Other* (text entry provided), or *I don't know*. After each question, the certainty measure (0 – 10, 0 being *uncertain* and 10 *extremely certain*) was presented. A copy of the questionnaire is provided in Appendix B.

It should be noted that the questionnaire was distributed to Najdi people (B and S) and additionally to people who were born, or lived most of their lives, in Najd but are originally from other regions of Saudi Arabia.

7.5.3 Multiple interviewers

Bell's (1984) audience design model (see also section 3.3.1.2) states that speakers might change their speech style (i.e., converge) according to their addressee. Bell (1984) reviews the evidence of style-shifting in Trudgill (1981), in which Trudgill analysed his own speech in ten of the interviews he conducted, as an interviewer, in Norwich, and in Coupland (1984), in which the speech of an assistant in a travel agency is analysed. Bell (1984, p. 166-167) asserts that different rates of accommodation will be found across linguistic variables due to salience. It is stated that indicators show little or no susceptibility to style-shifting, while markers do.

In Rickford and McNair-Knox (1994) and Llamas et al. (2009), a more systematic approach to test style-shifting according to addressee is provided. The speakers are interviewed by different interviewers, and the production pattern in one of the interviews is considered the default. The default pattern is then compared to the other(s) in order to analyse style-shifting patterns (i.e., extent of accommodation). In both studies, the salience of the variables is explicitly introduced as the reason why speakers show convergence in some features but not

in others. It should be emphasised here that a complete shift is not expected, but it is the rate of use of salient features that is most likely to increase.

In the present study, Bs showed convergence to the S variant in the merger (/u/ - /i/) and the 3SM suffixed pronoun (*-ih* / *-ah*) but not in the imperfect prefixes (*ya-* / *yi-*). The interviews of the Bs was conducted by a B interviewer (myself), and yet the participant showed convergence towards the S variants. Therefore, it is worthwhile exploring whether Bs will show more convergence rate if they were interviewed by a S interviewer.

A friend of mine, an S linguist from Al-Kharj, volunteered to interview some of the speakers who had already been interviewed by me. Due to time limitations for both the interviewer and the interviewees, only three B speakers were selected to be interviewed by the S speaker. They are grouped by tribe (Ajmi, Dosari, Qahtani) since tribe was found to be the most significant factor. For the other factors, they are all young and their level of education is high.

The production pattern in the interviews conducted by the B interviewer will be considered the default pattern, and the interviews by S interviewer will be then compared to it. It is hypothesised that in the new interviews Bs will show even more convergence in variables that showed convergence in the default pattern. However, it is predicted that there will be little or no convergence among those for which there was no convergence in the default pattern.

The salience of the variables is believed to be the reason why the convergence rate in some variables will increase (as a result of short-term accommodation) and not in the others. The results from comparison of the production patterns is expected to agree with the results from the SCAT and the identification task.

7.6 Results

In this section, the results of the SCAT, the identification task and the multiple interviews are presented.

7.6.1 SCAT results

The results (association percentage and response time) for the 10 B participants who took part in the SCAT is provided in Table 7.3 and plotted in Figures 7.2 and 7.3.

It should be noted that the two measures in the experiment (association percentage and response time) are not to be analysed independently. A higher percentage of correct associations is expected to lead to a shorter response time and vice versa. The relevance of response time increases with the percentage of correct associations (Llamas et al., 2016, p.11).

Table 7.3: *Association percentage and response time from SCAT*

| Participant | Response time | | | | | |
| | The merger | | 3SM pronoun | | Imperfect prefixes | |
	[u]	[i]	*-ih*	*-ah*	ya-	yi-
1	0.509008	0.526578	1.274619	(Incorrect)	2.538044	0.875541
2	0.525529	1.174283	0.126653	1.125379	1.340732	0.092796
3	1.504473	0.974608	0.591927	1.506179	1.773098	0.175873
4	1.008198	0.722271	0.524002	(Incorrect)	1.557153	(Incorrect)
5	0.492832	1.457279	1.225025	2.389496	1.076285	0.791843
6	1.374807	0.975589	2.190096	(Incorrect)	0.260132	(Incorrect)
7	0.691908	0.809083	1.308305	1.707136	(Incorrect)	1.824555
8	0.809234	0.806515	0.876502	2.257513	0.526104	2.569939
9	0.60936	0.608827	0.891361	1.191042	(Incorrect)	0.7925
10	0.510154	0.509431	(Incorrect)	1.474543	(Incorrect)	1.790806
Average RT	**0.80355**	**0.856446**	**1.000943**	**1.66447**	**1.295936**	**1.114231**
STDEV	**0.37363**	**0.298787**	**0.59375**	**0.492454**	**0.76969**	**0.867646**
Association	**100%**	**100%**	**90%**	**83%**	**72%**	**83%**
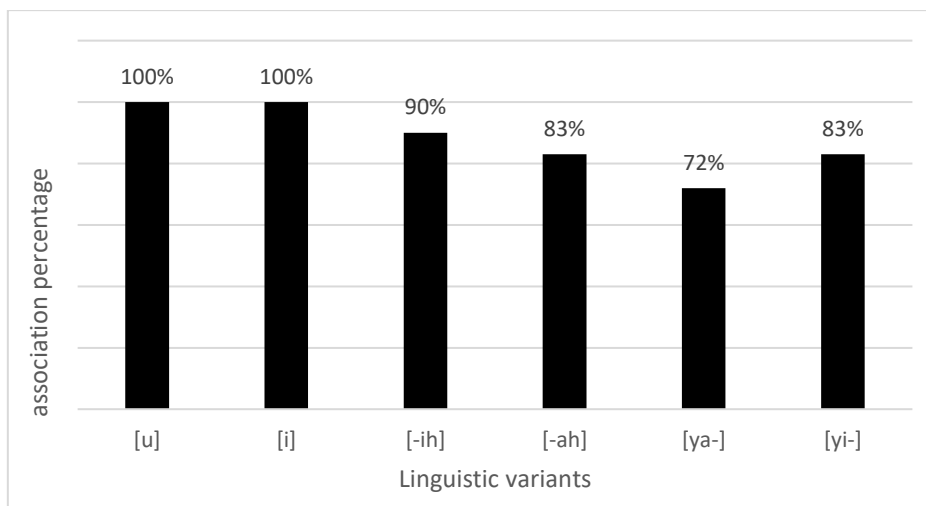
Figure 7.2: *Association percentage of linguistic variants in the SCAT.*
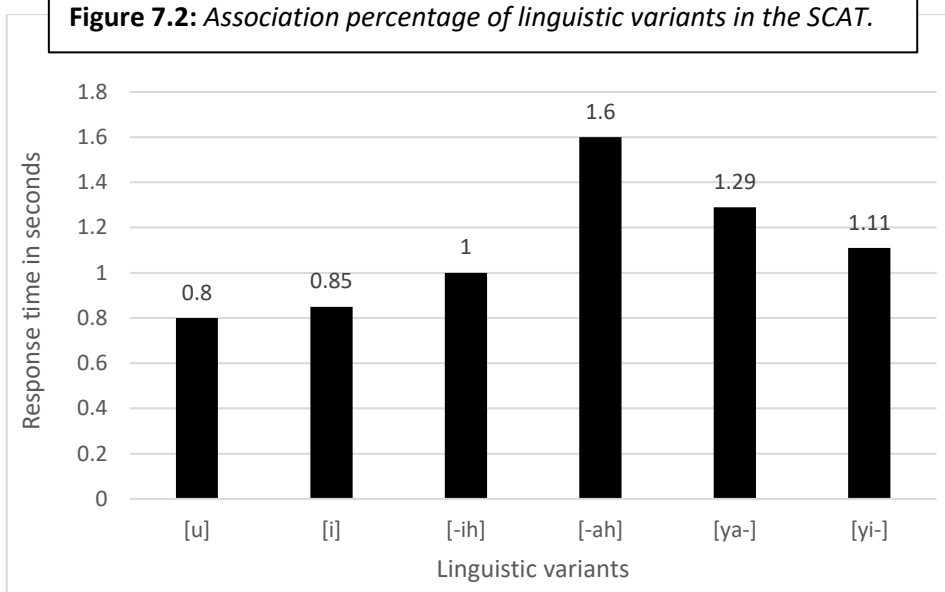
Figure 7.3: *Average response time for linguistic variants in the SCAT.*

The results show that the merger variants (/u/ - /i/) were associated correctly by all participants, while the variants of the other variables were mislabelled by some. The response time for the merger variants is the shortest, while for the other variants it is more than a second but less than two seconds for most participants. It can be stated based on these results that the merger is relatively salient compared to the other variables. This agrees with the production data as the highest rate of convergence was attested in the merger.

Now we turn to account for the difference between the association percentage and response time between variants of a single variable. The difference between /u/ and /i/ is minor, as manifested in the average response time. It can be claimed that the participants (all Bs) are as aware of /u/ as they are of its prestigious variant /i/. As for the 3SM suffixed pronoun, the gap in the two measures between the two variants (-*ih* and -*ah*) is rather to be expected. The variant -*ih* is used by Bs only and therefore it is easier to identify as B. The variant -*ah*, on the other hand, is used by both Ss (in all environments) and by Bs in perfective verbs, and thus it is highly likely that participants took more time to associate it with the S dialect and that some participants mislabelled it as B. Lastly, participants are less aware of the B variant *ya-* than *yi-*. This is probably because *ya-* is identical to the Modern Standard Arabic form, which is used by both Bs and Ss in formal settings.

7.6.2 Identification task results

As discussed above (7.5.2.2), in this task only correct/incorrect associations will be used as a criterion for relative salience. A certainty scale (0 – 10, 0 being *uncertain* and 10 *extremely certain*) was included to compensate for the lack of a response time measurement. But it is found that the former is not as reliable as the latter, as we will see.

34 participants (14 Bs, 10Ss, 10 others) responded. The forced choices for each question were *B*, *S*, *other* or *I don't know*. If for instance the stimulus represented the B dialect, all other choices were coded as *false*. If someone chose *other* and entered a name of a B tribe, it would be counted as *true*, but that did not occur in the results. The data from the certainty scale will not be used because the responses do not reflect awareness of the variants. Rather, most participants responded with 10 (extremely certain) for all stimuli, even the ones they associated incorrectly. Therefore, only correct/incorrect association will be used in the analysis of this data set.

It is expected, based on the production data and SCAT, that majority of participants will associate the variants of the merger correctly, while for the other variable there will be mistakes. The data from the identification task is plotted in Figure 7.4 for all variants. In Figures 7.5 – 7.10, the responses for each variant, showing correct/incorrect responses by the group of participants, are provided.
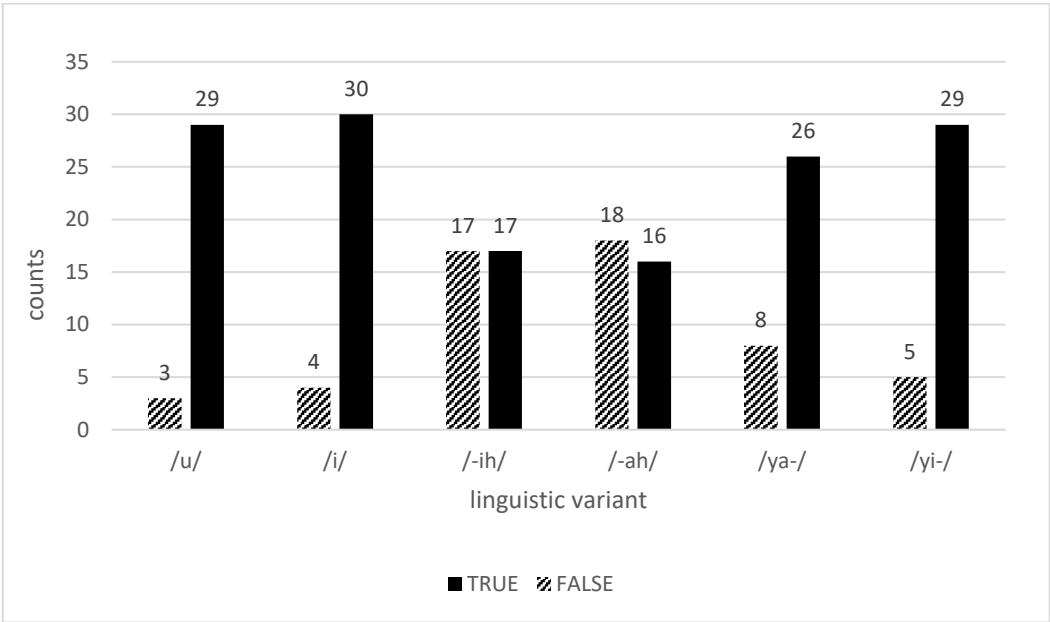
Figure 7.4: *Identification of the linguistic variants by response type (True / False). The total is 34 for each, except for /u/ (32).*

The results in Figure 7.4 agree with the production and SCAT results, except for the imperfect prefixes (*ya-/yi-*). The merger variants were identified correctly by the majority of participants while the 3SM suffixed pronoun (*-ih/-ah*) is identified correctly by 50% of the participants. This is expected and further indicates that the merger is more salient than the suffixed pronoun. Surprisingly, however, the correct association of the imperfect prefixes is relatively high, almost as high as the merger. This can be attributed to the major flaw of such tasks, which is overcome by the SCAT. In these tasks, including the current one, time is not controlled. As opposed to the SCAT, in which automatic (sub-conscious) responses are

targeted (e.g., IAT, Greenwald et al., 1998), in simple identification tasks listeners mostly record their conscious responses. In this particular task, in which listeners accessed the questionnaire online, we do not know how long they took to record their responses and how many times they played the stimuli. Furthermore, although the variants of each variable have been scattered, some participants might have modified their responses after listening to both variants. In other words, one might associate a variant with the B dialect incorrectly, and after listening to the B variant, he might change the previous response to the correct answer. It can be also claimed that when listeners record their conscious answers, they might look for other prosodic features of voice (e.g., voice quality, pitch and speech rate) to help identify dialects or social groups (Giles and Billings, 2004). Since the stimuli in this task are authentic recordings, it is possible that this is the case.

As for Figures 7.5 – 7.10, the observations that need to be highlighted are:

- No B mislabelled /i/ (Figure 7.6). As in SCAT, all Bs identified the prestigious S variant /i/ correctly.

- In the identification of the variants of the 3SM suffixed pronoun, each group identified their variant better than the other group (Figures 7.7 and 7.8). In the identification of the B variant *-ih*, 11/14 Bs identified it correctly, while 9/10 Ss mislabelled it. A similar, but reversed, pattern can be seen in the identification of the S variant *-ah*; i.e., more Ss identified it correctly than Bs.

- More participants from the *other* group identified *yi-* as S than *ya-* as B (Figures 7.9 and 7.10). This is expected since *ya-* is the Standard Arabic form.
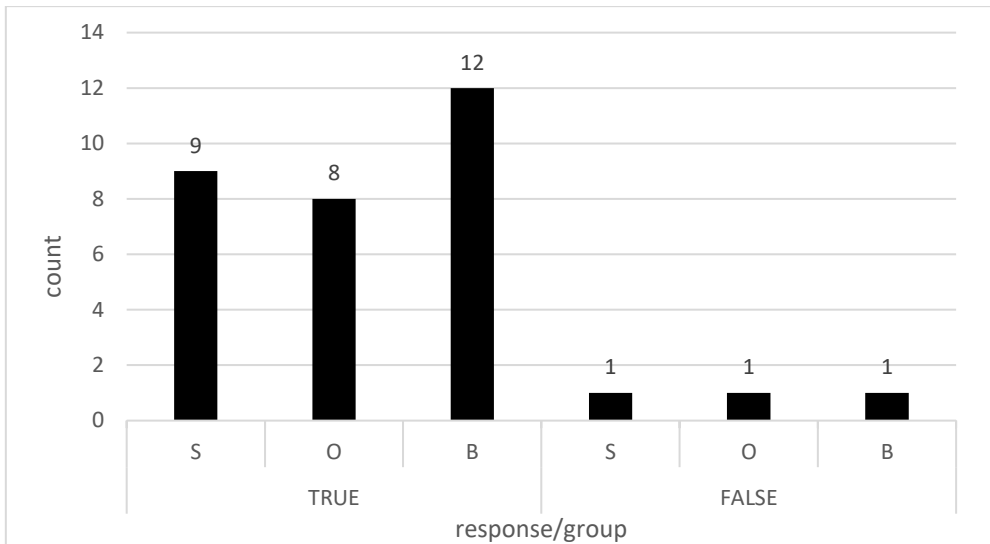
Figure 7.5: *Identification of the variant /u/ by response type and group.*
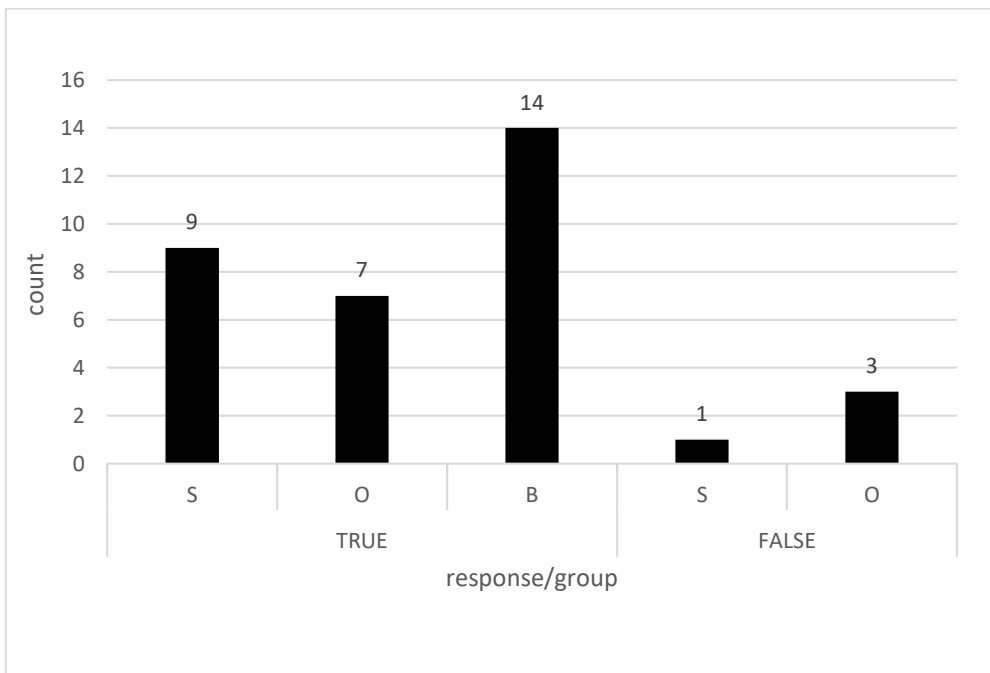
Figure 7.6: *Identification of the variant /i/ by response type and group.*
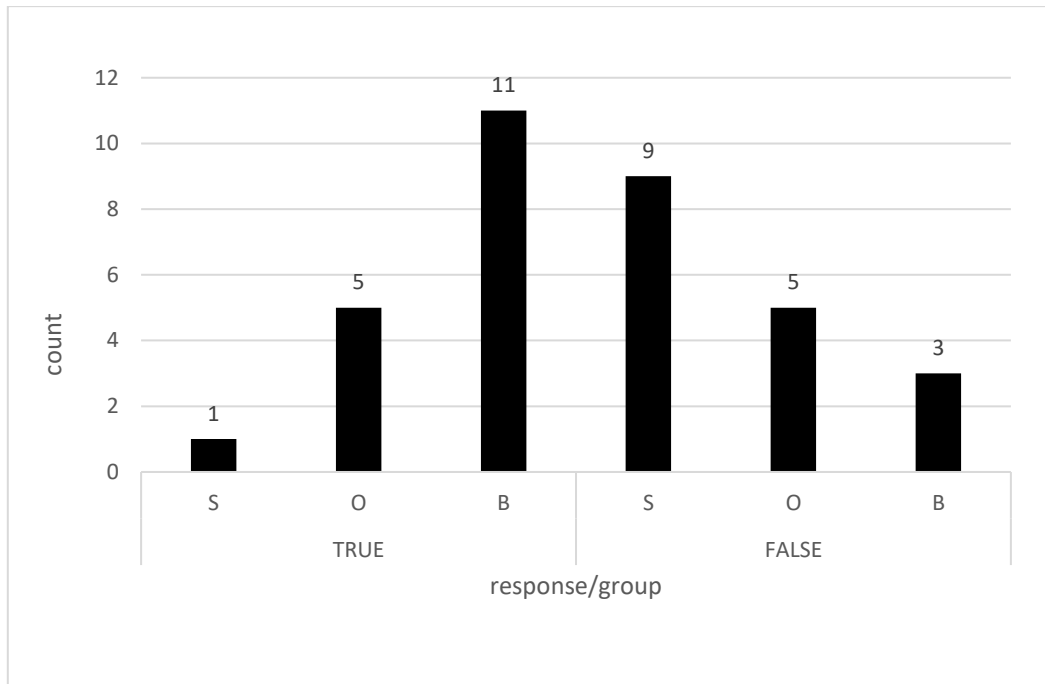
Figure 7.7: *Identification of the variant -ih by response type and group.*
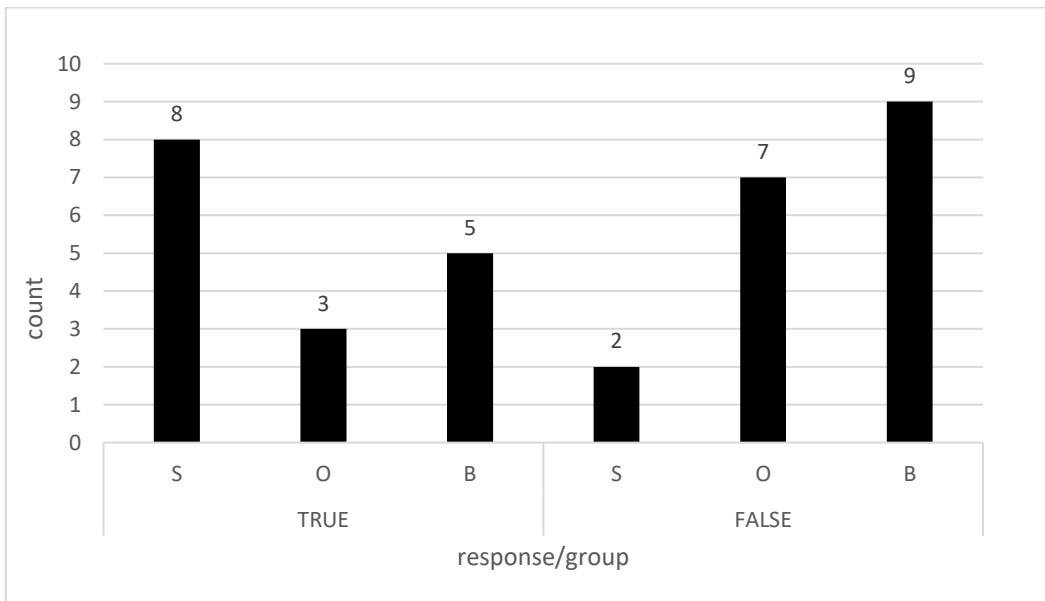
Figure 7.8: *Identification of the variant -ah by response type and group.*
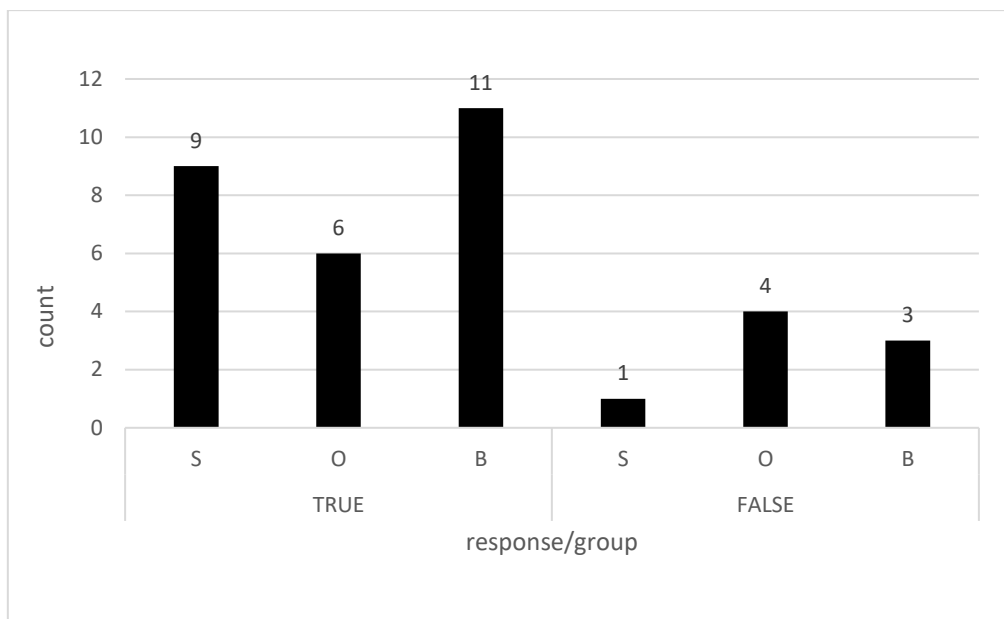
Figure 7.9: *Identification of the variant ya- by response type and group.*
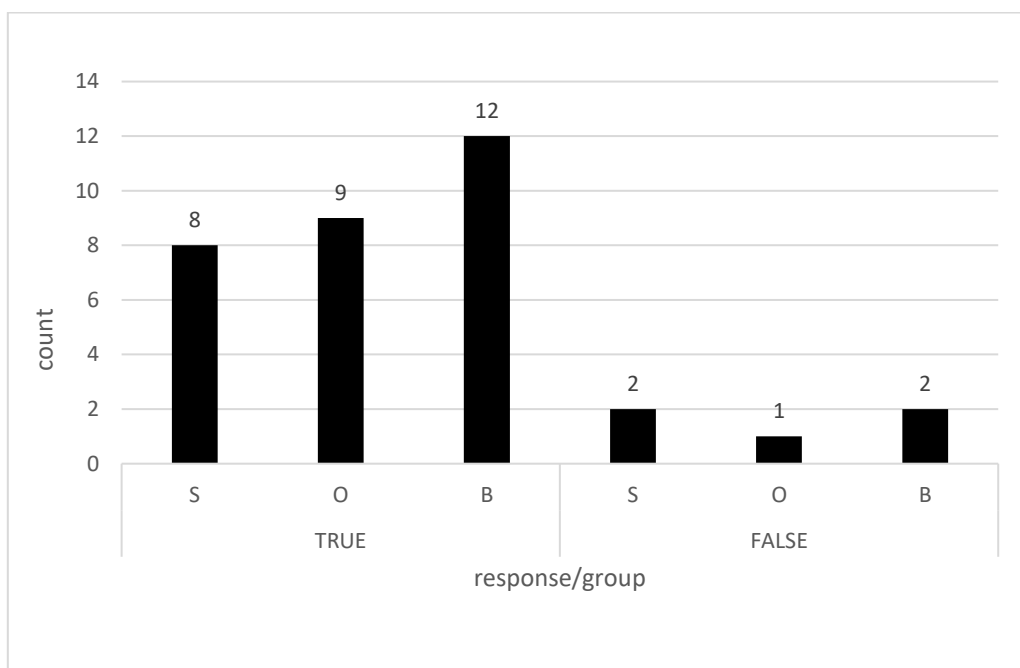
Figure 7.10: *Identification of the variant yi- by response type and group.*

7.6.3 Multiple interviews data

As seen in the production data, the Bs are converging on the S variants /i/ and *-ah* but not *yi-*.

The convergence on /i/ is, however, considerably higher than it is on *-ah*. B speakers' use of

/i/ is sociolinguistically motivated, and thus it is expected that there will be more style-shifting in the new interviews. Their use of -ah, however, is accounted for (so far) as a natural drift due to redundancy (Chapter 5); therefore, it is expected to show no style-shifting in the interviews by the S interviewer. Finally, Bs showed no convergence on yi- in the production data, and it is expected that there will be no style-shifting here either.

7.6.3.1 The merger (/u/ - /i/)

The data from both interviews for the 3 B speakers for this variable is provided in Table 7.4 and plotted in Figure 7.11.

Table 7.4: *Bs' use of /u/ with B and S interviewers*

Participant	B interviewer				S interviewer			
	Frequency		Percentage		Frequency		Percentage	
	[u]	[i]	[u]	[i]	[u]	[i]	[u]	[i]
Ajmi	25	3	89.2%	10.8%	26	17	60.5%	39.5%
Dosari	29	15	66%	34%	7	28	20%	80%
Qahtani	43	4	91%	9%	23	12	65.8%	34.2%
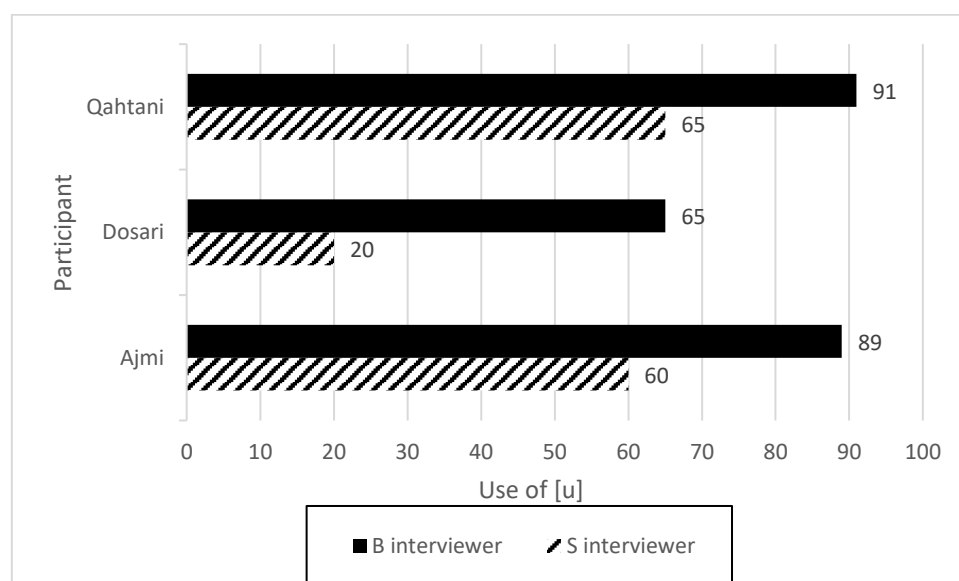
Figure 7.11: *Percentage of the use of /u/ by participant and interviewer.*

The data in Table 7.4 shows that all speakers converged on the S variant more when interviewed by a S interviewer. The Ajmi speaker converged by 30 percent more, the Dosari speaker by 45 percent more, and the Qahtani speaker by 26 percent more.

To test whether the degree of variance in the data is attributed to the interviewer effect (and not by chance), a simple Chi-square test was conducted for each speaker.

Table 7.5: *Chi-Square test results for the interviewer-effect experiment*

Speaker	Value	B interviewer	S interviewer	χ^2 (P value)
Ajmi	B (O value)	25	26	0.00858*
	B (E value)	19.7	30.3	
	S (O value)	3	17	
	S (E value)	7.8	12.2	
Dosari	B (O value)	29	7	4.70E-05*
	B (E value)	20.1	15.9	
	S (O value)	15	28	
	S (E value)	23.9	19.1	
Qahtani	B (O value)	43	23	0.00357*
	B (E value)	37.8	28.2	
	S (O value)	4	12	
	S (E value)	9.2	6.8	

The results of the test (Table 7.5) strongly suggest that the effect of the S interviewer is responsible for the increased rate of convergence on the S variant. This further supports the statement that the merger is a salient variable in Najd.

7.6.3.2 The 3SM suffixed pronoun (-ih/-ah)

As there is slight convergence led by young Bs on -*ah* in the default production pattern, it is difficult to predict whether Bs will converge more on -*ah* when interviewed by a S speaker, who uses -*ah* all the time. If this change is internally motivated, as claimed before, there

would be no increase in the Bs' rate of use when interviewed by a S speaker; but if there is an increase, this might lead us to reconsider the claim that this change is internally motivated. The data distribution by interviewers is provided in Table 7.6 and plotted in Figure 7.12.

Table 7.6: *Bs' use of -ih / -ah by interviewer, B and S*

| Participant | B interviewer | | | | S interviewer | | | |
| | Frequency | | Percentage | | Frequency | | Percentage | |
	[-ih]	[-ah]	[-ih]	[-ah]	[-ih]	[-ah]	[-ih]	[-ah]
Ajmi	59	0	100%	0%	38	0	100%	0%
Dosari	27	5	84%	16%	11	14	44%	56%
Qahtani	35	0	100%	0%	24	0	100%	0%
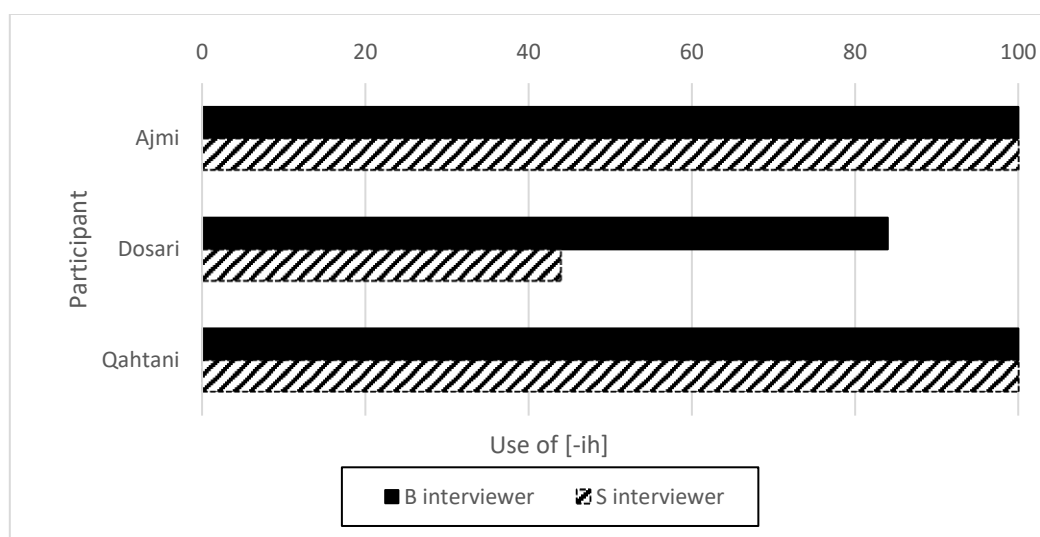
Figure 7.12: *Percentage of Bs' use of the B variant -ih plotted by participant and interviewer.*

That data reveals that speakers who did not show convergence with the B interviewer also did not converge with the S interviewer. As for the Dosari speaker, he converged significantly more on the S variant in the second interview. One possible explanation for this is that this variable has acquired *marker* status for some speakers but not for others. In other words, some Najdi Bs are aware of this variable, and they style-shift between variants according to the

interlocutor; on the other hand, other speakers are not aware of it (it is an *indicator*), and therefore do not style-shift towards the S variant.

7.6.3.3 The imperfect prefixes (ya- /yi-)

The Bs showed no convergence on the S variant *yi-*. It is expected that this variable is an indicator and there will be no style-shifting when interviewed by a S speaker. The data from both interviews is provided in Table 7.7 and plotted in Figure 7.13.

Table 7.7: *Bs' use of ya-/yi- by interviewer, B and S.*

Participant	B interviewer				S interviewer			
	Frequency		Percentage		Frequency		Percentage	
	ya-	yi-	ya-	yi-	ya-	yi-	ya-	yi-
Ajmi	26	0	100%	0%	15	0	100%	0%
Dosari	29	1	96.6%	0.4%	19	2	90.4%	9.6%
Qahtani	34	0	100%	0%	17	0	100%	0%
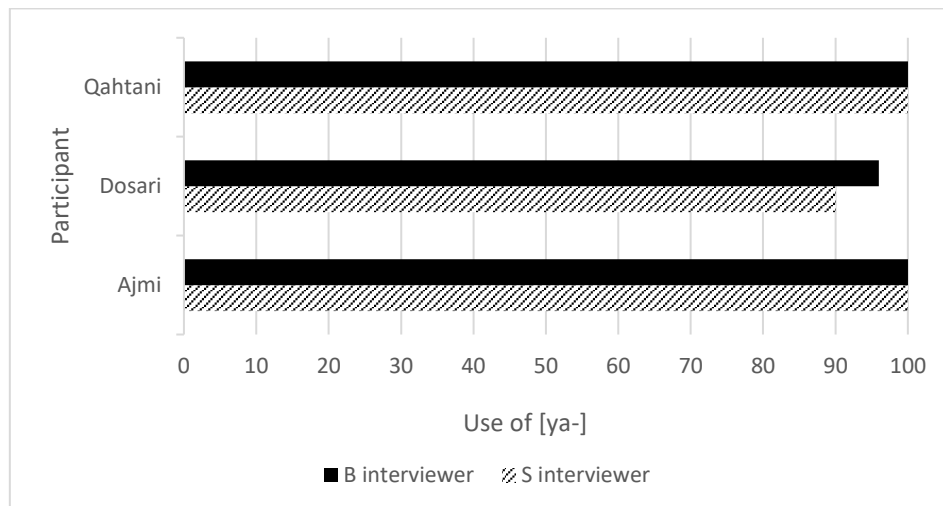
Figure 7.13: *Percentage of Bs' use of the B variant ya- plotted by participant and interviewer.*

Among the B speakers selected for the experiment, only the Dosari speaker shows convergence on the S variant with both interviewers. Although the rate of convergence increased with the S interviewer, the number of tokens is too low for us to attribute this shift

to the interviewer. As for the other B speakers, the results further support the statement that Bs are not aware of this feature, and hence they do not converge on the S variant.

7.7 Conclusion

In this chapter, the relative sociolinguistic salience of the three linguistic variables has been measured using the SCAT, an online identification task, and repeated interviews. The results from all measures show that the merger is a salient variable, which is acquiring the status of a *marker*, while the other two variables are *indicators*. Bs were able to associate the merger variants correctly in the SCAT and in the identification task, and they style-shifted to the S variant when interviewed by an S interviewer. For the other variables, there were some counterintuitive patterns; but overall, it can be stated with reasonable certainty that Bs are not aware of these variables.

Chapter Eight: Discussion

8.1 Introduction

In this chapter, we will discuss the quantitative results and their sociolinguistic relevance, qualitative remarks on the other linguistic variables (1.5.5.4), the formation of the Najdi koine and the sociolinguistics of diglossia in Najd.

8.2 Quantitative Results

In this section, we will discuss the quantitative results in relation to how they answer the research questions.

8.2.1 What is the linguistic outcome of the massive urbanisation in Najd?

Unlike the situation in Bahrain (Holes, 1986), Iraq (Abu-Haidar, 1991) and North Africa (Miller, 2004) where migration (urbanisation) caused a major shift towards the migrants' dialect, or the situation in Cairo and Damascus where the urban dialect is levelled but mainly retained (Miller, 2004), the massive urbanisation in Najd has increased the frequency of contact between Bs and Ss, neither of whom is considered the original population of the city. The linguistic outcome of this dialect contact is that B and S dialects are maintained, but there is a slight convergence towards the S dialect.

8.2.2 Is there any dialect accommodation? If yes, who is converging to whom?

The answer to this question is obviously 'yes', and it is the Bs who are converging to the S dialect. This is demonstrated by the usage rate of /i/ by B speakers (Chapter 4). As for the 3SM suffixed pronoun, we have proposed that the B convergence on -*ah* might be a natural drift, but a B speaker used -*ah* more when he was interviewed by a S speaker. On the other

hand, Bs did not show any convergence on *yi-*. Therefore, it can be stated that the result of this dialect contact is that both dialects are maintained, but Bs are showing slight convergence to the S dialect. The questions that follow this statement are: 'why are Bs converging to Ss and not vice versa?' and 'what is the mechanism of this change?'

8.2.2.1 Why do Bs converge towards S?

The direction of convergence in our case is attributed to two reasons: demography and prestige. For the first reason, Trudgill (2004, p. 85) maintains that "[i]n determining who accommodates to who – and therefore which forms are retained and which lost – demographic factors involving proportions of different dialect speakers present will be vital". As reviewed in Chapter 4, /u/ is used by Bs who are originally from the southern parts of the Arabian Peninsula while /i/ is used by both WBs and Ss. Censuses of Saudi Arabia do not show population by tribe; however, from observation I believe that, at least in the case of Al-Kharj, the proportion of speakers who use /i/ (WBs and Ss) is larger than that of those who use /u/. For the second reason, although Bs are proud of their heritage and identity, the S lifestyle is associated with modernity and education; thus, Bs who are educated and/or in contact with the Ss converge on the S variant, at least sub-consciously (Al-Sowayan, 2010, p. 345).

8.2.2.2 What is the mechanism of dialect convergence by accommodation?

According to the social psychologist Howard Giles, *dialect accommodation* is the process by which a speaker wishes to gain the receiver's approval by reducing pronunciation dissimilarities (*accent convergence*). If the speaker, however, wishes to disassociate himself/herself from the receiver, he or she may emphasise pronunciation dissimilarities (Giles, 1973, p. 90). This process is restricted to the accommodation that takes place in short-term contacts; but according to Trudgill (1986, p. 3) from the perspective of a linguist, this process can occur in the short-term as well as in the long-term. *Long-term accommodation*

refs to the situation where the effect of short-term accommodation has become permanent and habitual.

The link between accommodation and dialect convergence (i.e., *change-through-accommodation*) is presented differently in Bell (1984) and Trudgill (1986), but it is best conceptualised by Auer and Hinskens (2005) as integrating "three components into a hierarchy in which higher steps presuppose the lower ones" (p. 336). These components are presented in Figure 8.1.

Lowest level (interactional episode): short-term accommodation

↓

Middle level (the individual): long-term accommodation

↓

Highest level (speech community): language change
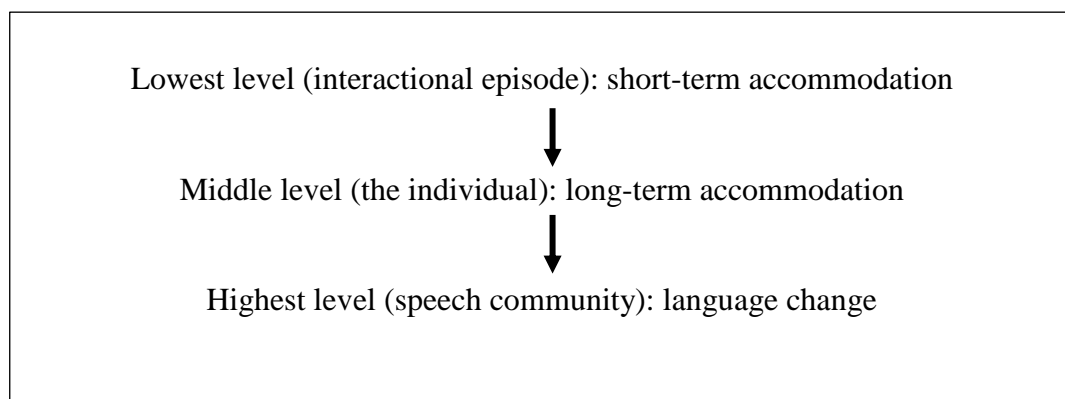
Figure 8.1: *Change-through-accommodation model by Auer and Hinskens (2005, p. 336).*

Below we will discuss the three components of this model in relation to the results relating to the merger in the present study. We will use this model to help explain the processes by which the Bs are converging on the S variant /i/.

Short-term accommodation: the Bs who use /u/ came into contact with the Ss who use /i/. In everyday personal encounters with the Ss, Bs started to accommodate to the Ss in the form of convergence. Because it is a salient feature, as we have seen in Chapter 7, Bs altered their frequency of usage of /u/ (>/i/). The Bs' short-term accommodation has been tested by means of a second interview with a S interviewer. The effect of the S interviewer was statistically significant.

Long-term accommodation: in this stage, the effect of these repeated interactional episodes becomes permanent in accommodating speakers. As Auer & Hinskens argue, "[t]his is the case when they transfer the innovation from direct interaction with the innovating speakers to situations in which these 'model speakers' fail to be the addressees." (Auer and Hinskens, 2005, p. 335). To elaborate, in the present study, the Bs showed a considerable rate of convergence on /i/ even when they were interviewed by a B interviewer (who is not the model speaker for the Bs, i.e., the model speaker is S). That is to say, the use of /i/ has become a speech habit (or has diffused in the speech) of the accommodating Bs.

Language change (dialect convergence): as the innovation spread in the community, it led to dialect convergence. In the case of Najdi B dialect, the convergence on /i/ is still in progress. Some speakers did not show any use of /i/, and no speaker has converged completely on /i/.

Another approach to language change, the *identity projection model* (Auer and Hinskens, 2005), is proposed to explain why some features are spread although there are no face-to-face interactions with speakers of the adopted feature. In the case of Najd, however, contact is the main reason behind Bs' convergence on /i/, as we will see next.

8.2.3 What are the independent social variables governing this convergence?

8.2.3.1 Social factors conditioning Bs' convergence on /i/

As we saw in Chapter 4, the factors of tribe, level of contact and education were statistically significant in determining who is leading the change.

The Dawasir tribe showed the highest rate of convergence; Qahtanis showed slightly less; and the Ajmi tribe showed the least. It cannot be determined whether this pattern holds true for all of Najd, meaning that we do not know if these tribes will show a similar pattern in other cities of Najd. If further studies prove that this pattern is found in other cities, it might be difficult to

pinpoint the reasons because historical socio-political and/or identity-related factors might come into play. If, however, this pattern is restricted to Al-Kharj, it is satisfactory to claim that (a) before and after the oil boom, the Dawasir tribe in and near Al-Kharj have been in contact with the S groups more than the other tribes, and (b) each tribe forms a separate group, mostly with a closed and/or multiplex network, and that the innovation has spread in the Dawasir's social network more than the others.

The factors of education and level of contact were also significant. The results reveal that speakers who are less educated and whose level of contact with the Ss is low are less likely to converge on /i/. However, these factors are not completely independent. As stated in 3.5.4, the factor of education in the Arab world, and in Najd, is a proxy for other factors, mainly level of contact. A Najdi individual from Al-Kharj, who has a college degree from either the same city or Riyadh, would inevitably have been in frequent contact with the Ss and other groups. For other jobs that do not require a college degree, such as government drivers or positions in the military, an individual mostly has a choice in terms of deciding whom he spends time with in the work place.

8.2.3.2 Social factors conditioning Bs' convergence on -ah

The Bs show slight convergence on *-ah*, and this change is led by the young Bs. For some speakers *-ah* was never used, while for others the rate of their *-ah* usage varied. The result from the second interviews reveals that speakers who did not show convergence with the B interviewer also did not converge with the S. On the other hand, the third speaker used *-ah* with the B interviewer, and the rate of his *-ah* usage increased in the second interview. This implies that the variable has achieved the level of *marker* for some Bs but is still an *indicator* for most. This claim is supported by the results of the SCAT and online identification task.

We cannot, however, account for why this variation is explained by age and not the factors of tribe or level of contact, for example.

8.3 Other Linguistic Variables and Differences within B Tribes

In this section, we will present our findings for the variables that are not frequent enough to be analysed quantitatively. The following qualitative remarks are based on the data and my personal observations. In the final sub-section, we will shed some light on the differences between B tribes' dialects.

8.3.1 Internal passive formation

The term 'internal passive' is used to distinguish the archaic process characteristic of Classical Arabic from affixation which is found in almost all modern Arabic dialects. In Classical Arabic the perfective active form *kataba* 'he wrote' is transformed into passive by using the internal vowels u-i-a instead of a-a-a (*kutiba* 'it was written'). In imperfective verbs (e.g., *yaktub* 'he writes') the vowels used to form passive are u-a (e.g., *yuktab* 'being written'). In modern Arabic dialects, *katab* (perfective) is transformed into passive by prefixing *in-* as in the S dialect of Najd or *it-* as in Cairene Arabic (Watson, 2002, p. 11), e.g., *in-kitab*, *it-katab*. In imperfective verbs, *-n-* (Najdi S) or *-t-* (Cairene) is infixed as in *yi-n-kitib* or *yi-t-kitib*.

Ingham (1994, p. 26) reports that the internal passive formation is fully retained in the B dialects of Najd. However, the internal vowels are different from those of Classical Arabic. The passive form of the perfective verb *sarag* 'he stole' is *sirg* 'was stolen'. In imperfective verbs, the vowel /i/ is used instead of Classical Arabic /u/ (e.g., *yaktib* 'he writes' > *yiktab* 'being written'). In the S dialect of Najd, affixation is used instead of internal alternation of vowels, as explained above.

Our findings agree with what Palva (2006, p. 611) postulated: the internal passive and other conservative features of B dialects in Najd are reduced, due to contact with the S dialects. In the data, it was noticed that the internal passive is frequently used by old B Najdi speakers but is decreasing in the speech of young Najdis. As for the S speakers, affixation is used all the time. The S speakers who switched to Standard Arabic used the internal passive with the Classical Arabic vowels u-i-a (perfective) and u-i (imperfective).

8.3.2 Guttural effect

Johnstone (1964, p. 80) and Ingham (1994, p. 19) report that in Najdi Arabic, the guttural sounds (/χ/, /ʁ/, /ħ/, /ʕ/, /h/) cause Classical Arabic syllable-initial CaG- forms (where G stands for guttural) to re-syllabify as CGa-. This rule can be found in nouns such as *ðˤha-ri* 'my back' (< *ðˤah-ri*), but it is mostly evident in imperfective verbs that begin with a guttural (e.g., *yħa-fir < yaħ-fir* 'he digs'). As we have seen in Chapter 6, imperfective verbs are prefixed with *ʔa*, *ya-*, *ta-*, or *na-* in the B dialect; thus, imperfective verbs which begin with a guttural form the initial syllable CaG. This rule has been reported in Najdi Arabic (Ingham, 1994), but it was not specified whether it is found in one dialect or both (B and S). The S dialect uses the imperfect prefix *yi-* (Chapter 6); this raises the question whether Ss will use CiG, as in *yiħ-fir* 'he digs', or re-syllabify it as *yħa-fir* (section 6.5.3.4). Furthermore, in nouns where the vowel is always /a/, will there be any variation between the groups in terms of their use of CGa or CaG?

To sum up, there are three plausible forms for this variable:

- CaG: this is the original initial syllable in Classical Arabic. It can be found in nouns and imperfective verbs (given that the dialect uses /a/ in imperfect prefixes, which is the case in Classical Arabic and Najdi B dialect).

- CGa: this is the re-syllabified form. It is expected to be found in nouns where the vowel involved is /a/ and in imperfective verbs of the B dialect where the vowel of the imperfect prefix is /a/.

- CiG: this form is expected to be found in imperfective verbs in the S dialect (where the vowel of the imperfect prefix is /i/).

Our sociolinguistic goal is to test which of these forms are used by which group. Because this variable is infrequent in conversation, an additional picture task was included to elicit more tokens. Participants were presented with GIFs (graphics interchange format) of an action, and they were asked 'what is he/she doing?'. The actions in these animated pictures (see copy in appendix A) are expected to be described with imperfective verbs that begin with a guttural. After examining the data from both conversational style (where both nouns and verbs have been elicited) and the picture task (where only imperfective verbs have been elicited), it was found that these forms co-exist in a rather complicated pattern.

Here are our findings (Table 8.1):

- Bs used CGa with both nouns and verbs in conversational style and with verbs in the picture task. In the picture task, some speakers switched occasionally between CGa and CaG. This is expected since the picture task may have caused them to pay more attention to speech; hence, they used the Standard Arabic form CaG. Bs, however, did not show any use of the S form CiG, which is as expected given that they did not show any use of *yi-* as imperfect prefix (Chapter 6).

- In conversational style, the Ss alternated between CGa and CaG in nouns, and CGa and CiG in imperfective verbs. In other words, the Ss did not employ the guttural re-syllabification consistently. In the picture task, they alternated between all three forms.

- The data for both groups indicates that the picture task triggered the use of Standard Arabic, i.e., use of CaG. The Ss do re-syllabify due to the guttural effect, but not as consistently as do the Bs. A similar pattern is detected in Kuwait (Holes, 2007, p. 618): both forms (CaG and CGa) co-exist freely in the dialects of Kuwait, which are those of sedentarised B (originally Najdi) and S.

Table 8.1: *The distribution of variants of guttural re-syllabification rule in Najdi dialects*

Dialect	B dialect						S dialect					
Style	Conversational style			Picture Task			Conversational style			Picture Task		
Form	CGa	CiG	CaG	CGa	CiG	CaG	CGa	CiG	CaG	CGa	CiG	CaG
	B	S	SA[34]	B	S	SA	B	S	SA	B	S	SA
Nouns	✓	NA	X	NA	NA	NA	✓	NA	✓	NA	NA	NA
Verbs	✓	X	X	✓	X	✓	✓	✓	X	✓	✓	✓

8.3.3 Lowering of /e:/ and /o:/

As mentioned earlier throughout the thesis, the Classical Arabic diphthongs /ay/ and /aw/ have been monophthongised to /e:/ and /o:/ (or /u:/) in most contemporary Arabic dialects. Classical Arabic *sˤayf* 'summer' and *qaws* 'bow' are, therefore, realised as *sˤe:f* and *go:s*. In Najdi Arabic, /e:/ and /o:/ are the monophthongised forms of the diphthongs, as reported by Ingham (1994). However, when followed by a short /a/, it was observed that some Bs use lower versions of these vowels. Therefore, the monophthong /e:/ is produced as /ɛ:/ (e.g., *bɛ:ðˤa* 'white' and *ʃɛ:nah* 'ugly' instead of *be:ðˤa*, and *ʃe:nah*), and /o:/ is pronounced as /a:/ in the same environment (e.g., *sa:da* 'black', and *ħa:la* 'cross-eyed, F' instead of *so:da* and *ħo:la*).

---

[34] Standard Arabic.

Another environment in which /o:/ is lowered is in verbs that begin with /w/. Ingham (1994, p. 25) reports that in verbs in which /w/ is the first consonant of a trisyllabic root it is deleted, resulting in a compensatory lengthening of the preceding vowel. Therefore, Classical Arabic *ya-wtˤa* 'step in' is realised as *ya:-tˤa*. Ingham did not comment on whether this form is used by Bs or Ss; however, I argue that such forms are only B, and they occur as a result of series of phonological processes: /aw/ > /o:/ > /a:/ (before /a/). Hence the S form of the same verb is *yo:-tˤa*.

After the interview, some speakers from both groups were presented with a picture which included scattered small images of objects that are pronounced with the target vowel /e:/ (Appendix A). The speakers were then asked to look at the picture for 15 seconds (sometimes more) and to attempt to recall as many objects as they can remember. The vowel /o:/ was not included in the picture task because it was found to be avoided in a picture task conducted in the pilot study.

It was found that lowering of /e:/ to [ɛ:] is used by old and less educated speakers. Old speakers produced the lowered form in both the interview and the picture task, while younger less educated speakers avoided it in the picture task. On the other hand, the lowered form is disappearing from the speech of educated Bs. As for /o:/, it was found to be avoided in nouns (e.g., *sa:da* 'black' < *so:da*) by educated speakers, but used in verbs (e.g., *ya:-sˤal* 'he arrives' < *yo:-sˤal*) by the majority of Bs, both educated and less educated.

8.3.4 Tanwīn

Tanwīn (also known as nunation) is one the striking conservative features of the B dialects of Najd (Ingham, 1994, p. 6; Versteegh, 2001, p. 193). In Classical Arabic it includes the addition of the sound /n/ as *-un* for indefinite nominative, *-in* for indefinite genitive and *-an*

for indefinite accusative. In Najdi Arabic this marker is clearly derived from Classical Arabic, but it has acquired a different grammatical function. Ingham (1994, p. 49) states that:

> In the dialect [Najdi Arabic] however the distinction of case has been lost and *-in* marks only indefiniteness. The conditions for the occurrence of the suffix are partly syntactic and partly semantic. In the central Najdi type the suffix occurs more rarely in utterance final position and therefore has something of the nature of a juncture element since it occurs in positions where a noun is followed by some modifying element in an appositional relationship non-pausally.

These positions, according to Ingham, are:

- When a nominal follows another nominal as in *beet-in kibi:r* (house-IND large), 'a large house'.

- When a modifying prepositional phrase follows a nominal as in *wa:ħd-in min irrabiʕ* (one-IND from the group), 'one of the group'.

- When a noun is followed by a modifying clause as in *kalimt-in ga:latha: li* (word-IND said-she-it to-me), 'a word which she said to me'.

- When an active participle is used with verbal function but not followed by an object suffix; and particularly where it is followed by a complex of *l-* and a personal pronoun suffix, as in *ana ga:yl-in l-ik* (I having said-IND to-you), 'I have said to you'.   (Adapted from Ingham, 1994, p. 49)

An interesting finding was revealed in the data. This feature is used by not only Bs but by Ss as well. The Ss, however, used it less frequently than the Bs. Why is this feature found in the speech of the Ss? Could it be that they are accommodating to the Bs? I would argue that this feature already existed in the S dialect, but the reduction process started with the Ss before the Bs. This is supported by the fact that older S speakers showed higher usage of tanwīn than young speakers.

As for the Bs, the same pattern is found: older speakers showed a higher rate of usage than younger speakers.

Our findings for this feature can be summarised as the following: tanwīn is being lost from both the B and S dialects. The reason why Bs show more usage in general is because we probably caught the change at different stages in each dialect.

8.3.5 ħili:b / ħali:b

The linguistic variation between ancient B and S dialects in the 8th century has been reported by Sibawayh in his book *Al-Kitab* (Cadora, 1989). One of the reported differences is the vowel assimilation in the B dialect: "the vowel /a/ is raised from [æ] to [i-e] before /i/ or /ī/ when the intervening consonant is a member of this group of sounds: /ʔ/, /h/, /ħ/, /ʕ/, /χ/, /ʁ/". (ibid., p. 274). As observed in the data, this type of assimilation still exists in the B dialects of Najd; however, this rule additionally applies if a guttural precedes the assimilating vowel. That is, /a/ assimilates with /i/ if the preceding or intervening sound is guttural.

Two of the lexical items in which this rule applies are *ħili:b* 'milk' (< *ħali:b*) and *ʕisˤi:r* 'juice' (<*ʕasˤi:r*). These assimilated forms of the words are stigmatised and most Bs tend to avoid them. As for other words, in which the rule applies (Table 8.2), they are used interchangeably with their non-assimilated counterparts in different social contexts. The more attention that is paid to speech, the less likely Bs will be to assimilate /a/. It can be stated, based on our data, that old uneducated speakers tend to assimilate /a/ more than younger educated speakers.

Table 8.2: *Examples of /a/ assimilation in Najdi B dialect (from the current data)*

B dialect	S dialect	Gloss
ħimi:s	*ħami:s*	'a traditional dish'
ʕiri:f	*ʕari:f*	'military rank'
ħiʃi:ʃ	*ħaʃi:ʃ*	'hashish'
ximi:s	*xami:s*	'Thursday'
ʁiri:b	*ʁari:b*	'stranger'
bixi:l	*baxi:l*	'stingy'

It should be noted that in Najdi Arabic (i.e. both B and S dialects) this vowel assimilation can be noted in several lexical items in non-guttural environments. When compared to Classical Arabic, Najdi Arabic assimilated /a/ in words like *sˤidi:g* (< *sˤadi:q*) 'friend', *giri:b* (< *qari:b*) 'close', *dʒidi:d* (< *dʒadi:d*) 'new', *kiri:m* (< *kari:m*) 'generous' and *kibi:r* (< *kabi:r*) 'big'.

To sum up, in guttural environments, old uneducated B speakers tend to assimilate the vowel /a/ with /i/, while young educated speakers tend to avoid this habit.

8.3.6 Differences between B tribes

As stated earlier in 3.5.3, there are dialectal differences between the B tribes, and there are studies each of which examines the dialect of a B tribe (e.g., Johnstone, 1964, 1967b; Al-Hazmy, 1975; Ingham, 1979, 1982b, 1986; Al-Qahtani, 2015). Although in this study we focus on the linguistic features shared by tribes which also shared a lifestyle (i.e., B), in this section, we will shed some light on the linguistic features that are distinctive to one B tribe. It must be noted that these marked features are being lost in favour of the majority or koineised forms. These features are used by older individuals and those who live in extremely closed networks of the same tribe.

The distinctive linguistic features of each tribe are few, but most of the differences indicated by lay people are lexical. Most of these differences, both morphological and lexical, have been reported by Al-Darsoni (2013).

The morphological features characteristic of a particular B tribe are *kashkashah* (Ajmi) and insertion of the 1SC independent pronoun *ana* in perfective verbs (Dawasir). The ancient Arab grammarians Sibawayh and Ibn-Jinni mention two types of *kashkashah*: replacement of the 2SF suffixed object/possessive pronoun *-ki* with *-iʃ*, as in *kita:b-iʃ* 'your book', or with *-kiʃ*, as in *ʔaʕtˤaytu-kiʃ* 'I gave you.F' (Al-Azraqi, 2007, p. 230). The former type is used in most of the contemporary spoken dialects in the south-western parts of Saudi Arabia and Ajmi tribe in Najd. It is one of the most marked features which members of this tribe, including myself, avoid when talking to other Najdis. The younger generation of Ajmi in Najd is losing this feature, given that there has been sufficient contact with other groups of Najd. The other distinctive feature is insertion of *ana* in perfective verbs by Dawasir tribe (reported by Al-Darsoni, 2013). The perfective verbs *xaðe:tha* 'I took it.F', *ʕatˤe:tik* 'I gave you' and *jibtih* 'I brought it.M' are realised as such in all Najdi dialects (B or S), while in Dawasir dialect, they are realised as *xaðe:t-ana-ha*, *ʕatˤe:t-ana-k* and *jibt-ana-h*.

As for the lexical differences, I list below some of the distinctive words across the B tribes, documented by Al-Darsoni (2009, 2013).

Table 8.3: *Examples of the lexical differences between B tribes*

Ajmi	Dawasir	Qahtani	Gloss
yaʃbaħ	*yʕa:yin*	*yxa:yil*	'to see'
ħe:dah	*ħasˤa*	*ħasˤa*	'rock'
ʃafrah	*sakki:n*	*sakki:n*	'knife'
yasmaʕ	*yasmaʕ*	*yitisˤawwax*	'to eavesdrop'
ibi:	*ubu:y*	*ubu:y*	'my father'

8.4 The Formation of Najdi Koine

Siegel (2001, p. 175) defines *koine* as follows:

> A koine is a stabilized contact variety which results from the mixing and
> subsequent levelling of features of varieties which are similar enough to be
> mutually intelligible, such as regional or social dialects. This occurs in the
> context of increased interaction or integration among speakers of these
> varieties.

The term *koineisation*, first used by Samarin (1971), refers to the processes of levelling and mixing which lead to the formation of a stabilised koine (Siegel, 1985, p. 364). According to Trudgill (1998; 2004) there are three stages involved in formation of a new dialect, each of which corresponds to a generation of the settlers:

- Stage I: rudimentary levelling (first generation, adult migrants)

- Stage II: variability and apparent levelling (second generation, first native-born speakers)

- Stage III: focusing and reallocation (subsequent generations)

The first stage is what Siegel (1985) refers to as *pre-koine*, while the second and third represent the koineisation process (Kerswill, 2013). The approximate time-span of the koineisation process of New Zealand English, for example, is fifty years (Trudgill, 2004, p. 129); however, that is not necessarily the case of other contact situations. Kerswill (2013, p. 526) maintains that koineisation is "a process with distinct but overlapping stages and a variable but finite time-span." In this section we will discuss the three stages of koineisation in relation to the formation of Najdi koine, which I argue is still in progress.

Before I proceed, it should be noted that koineisation in the case of Najd means that Bs have converged completely on the S dialect. There is no evidence in the present study that suggests that Ss have adopted a B feature, or an intermediate one. The direction of change in the three main variables and most of the other variables (section 8.3) is towards the S dialect.

8.4.1 Stage I

This stage represents the first contact between the speakers of the contributing dialects. The dialect contact between these speakers, i.e., the adult migrants and the host community, is said to result in rudimentary levelling. *Levelling* refers to the process "in which minority forms, socially marked forms and linguistically marked forms are lost" (Trudgill, 2004, p. 23). The process of rudimentary levelling is the result of language accommodation, discussed in 8.2.2.2. In other words, when speakers of mutually intelligible dialects move to a new settlement, one (or more) group accommodates to the other, hence loss of marked forms.

When the first generation of Bs and Ss moved to the city, i.e., came in contact with one another, the Bs had not accommodated to the Ss. This is reflected in the data of speakers whose age is above 50 years. The age groups in our sample have been reduced to two in order for the *glmer* model to work. The number of B speakers who are above 50 is 5, and their age is as follows: 50, 55, 60, 75 and 77. Their rate of convergence on /i/ is 2%, 24%, 0%, 0% and 0% respectively. The speaker who showed 24% convergence is a real estate agent, and hence has a high level of contact with the Ss.

Why have the first generation of Bs not accommodated to the Ss? The first reason is obviously the level of contact with the Ss. It is rooted in the heritage of the tribal community of Bs to move and settle in groups. When the Bs moved to the city, they clustered in particular neighbourhoods, isolating themselves from other groups. However, they interacted with Ss in the work place. In these interactions, the Bs either did not *need*, or did not *want*, to accommodate to the Ss. They did not need to accommodate because the use of their variant /u/ did not obstruct comprehensibility (Trudgill, 2004, p. 89). Another possibility is that Bs intentionally desired to preserve their identity by maintaining their dialect features, even if they are different from those of the Ss. Kerswill (2010, p. 323) states that personal and social

identity formation is one of the crucial community-specific variables that play a major role in new-dialect formation.

8.4.2 Stage II

This stage involves the first native-born children (second generation) in the new settlement; and according to Trudgill (1998, p. 198) it is characterised as chaotic, demonstrating extreme inter-individual as well as intra-individual variability. He adds "[u]nlike in stable situations, where children normally acquire the dialect of their peers, in diffuse dialect contact situations there will be no single peer-dialect to acquire, and the role of adults, especially perhaps parents, will be more significant than is usually the case" (ibid., p. 198). This stage is crucial to the formation of a new dialect, but it is susceptible to several factors which can delay or accelerate the process.

In the case of Al-Kharj, the Bs' rate of convergence on /i/ is relatively high, compared to the other variables; however, it is not high enough to indicate that (a) the second generation is implementing further levelling away from the speech of their parents and (b) the convergence will be complete soon. This can be attributed to two reasons.

The first and most relevant reason is lack of sufficient contact between the groups. Sufficient contact is a crucial condition in order for the accommodation to take place (Kerswill, 2010, p. 232). As proposed in the previous section, the first generation of the migrants had minimum contact with the Ss, even in the city. Therefore, it can be claimed that the speech of the second generation resembles rudimentary levelling, because they are probably the first generation who actually came in contact with the Ss, not their parents.

The second reason is *the degree of difference between the input varieties* (Kerswill, 2013, p 532). If the dialect differences are subtle, as in the case of Milton Keynes, such heterogeneity can be bypassed by the second generation "and focusing toward a new variety is accelerated"

(ibid., p. 532). In the case of Najdi dialects, the difference between /u/ and /i/ is not subtle; thus, the amount of accommodation individuals have to engage in is relatively high.

8.4.3 Stage III

This stage involves the subsequent generations. Their speech is characterised as stable due to *focussing*, defined as "the process by means of which the new variety acquires norms and stability" (Trudgill, 2004, p. 88). Unlike their parents, the speakers of this generation are presented with fewer variants to choose from; therefore, the selection of variants is more focused towards *majority forms* (i.e., forms used by larger demographic groups). According to Trudgill (2004, p. 120), majority forms are more likely to survive even if they are marked.

Now we shall briefly discuss which of the variants of each variable in the present study are the majority forms. The merger variant /u/ is only used by Bs who are originally from the south-western regions of the Arabian Peninsula, while /i/ is used by WBs and Ss. Therefore, the merger variant /i/ is the majority form. The variants of the 3SM suffixed pronoun and imperfect prefixes *-ih* and *ya-* are used by both Bs and WBs, while *-ah* and *yi-* are used by the Ss. It should be noted, however, that the S variants *-ah* and *yi-* are found in the B and WB dialects but in different morphological contexts. Thus, no variant can be claimed as the majority form because: (a) we do not know whether the number of Bs and WBs together is larger than the Ss, and (b) the S variants already exist as allomorphs in the B and WB dialects.

As discussed in the previous section, the speakers of the second generation (first-born native speakers) are categorised as the first generation who came in contact with the Ss. Consequently, the speakers of the third generation can be considered as Trudgill's second generation, whose speech exhibits further levelling. The youngest speakers in our sample (aged 17) do not exhibit more convergence on /i/ than older speakers. In fact, most of them show less convergence than speakers who are aged between 26 and 40. This can be attributed

to various social factors,[35] but in most cases, it is education (university in this case) that alters their social networks and increases contact with the Ss (discussed in 3.5.4).

Table 8.4: *Comparison between the speech of two age groups based on education*

High school		University	
Age	Convergence on /i/	Age	Convergence on /i/
17	0%	25	11%
17	0%	37	36%
17	7%	44	20%
18	26%	25	44%
17	2%	36	14%
17	15%	26	9%
17	17%	38	55%

Before I give any predictions about the formation of Najd (or Riyadh) koine, we will review what others have said about this issue. Ingham (1994) acknowledges the economic changes caused by the oil boom in Saudi Arabia and speculates that Najdi Arabic is moving towards standardisation. He postulates:

> It is possible that in the future the dialect of Riyadh will emerge as the general standard of Central Saudi Arabia and that process can be seen to be happening today. However it is also noticeable that the emerging standard avoids certain local features so that in fact Riyadh Standard speech is approaching the nature of a Koine so that if a standard does emerge, it may be rather less Najdi in nature than the dialects of the other towns of Central Najd. (p. xii)

Al-Essa (2008, p. 22) states that "[a]lthough there is no study yet to verify the emergence of a standard dialect for Najdi speakers; evidence from this study indicates that a standard Najdi dialect is being formed".

Determining whether a Najdi koine has formed or not is subjective. When I asked speakers of other dialects (e.g., Hijaz) whether they can recognise the differences between B and S

---

[35] It should be noted that mergers at /i/ seem to be common cross-linguistically. In the history of Greek, for example, there has been a series of vowel mergers at /i/ (Labov, 1994, p. 127).

dialects of Najd, they claimed that they recognise one dialect 'Najd dialect'. This is expected because speakers of other Saudi varieties use regional features, distinguishing Najdi dialect from others (e.g., loss of interdental fricatives in Hijazi dialect). The subjectivity of this issue does not apply just to lay people but to linguists as well. According to Dillard (1972, p. 302) a "[k]oine is the term for a 'common' dialect which lacks prominent features of the more conventional dialects of a language. It is the end result of dialect levelling". Therefore, the formation of Najdi koine depends on what a linguist sees as 'prominent features'.

Before I proceed, it should be noted that I am aware of the risks associated with predicting the future of language change (Mair, 2013). Just as the prophecies of Jagger (1940) and Quirk (1972) about the future of English failed (cited in Mair, 2013, p. 315-316), our speculations about Najdi koine might also fail. However, I cautiously argue that, as far as the current variables are concerned, the formation of a Najdi koine is still in progress, based on data from the present study.

There are clear indications that a koine is forming in Riyadh, and that it may eventually extend throughout Najd. Other things being equal, it would be expected to be stable after two or three generations. However, this depends on various social factors; in the present case the process is likely to be hampered by the nature of social networks and identity of Bs, in and outside the capital city. Of course, the social networks of Bs in Riyadh are more open than those in Al-Kharj; but regardless of where they live, Bs have a strong group identity by nature, hence their social network tends to be relatively closed.

Now we turn to the linguistic evidence supporting our argument. The linguistic features of the B dialect are either: (almost) lost, in progress, or completely retained. Due to the salience of the variables and the reported rate of convergence, *-ih/-ah* and *ya-/yi* are not expected to be levelled in the near future. The merger, on the other hand, is expected to be completed after

two or three generations. As for the other variables discussed above, the direction of change is as follows:

- The internal passive is almost lost in the speech of young B Najdis, in favour of affixation.

- Lowering of /e:/ and /o:/ is almost lost in the speech of young people.

- The operation of the guttural effect (CGV), which is used by both groups, appears increasingly to result in the Standard Arabic form (CaG) being favoured.

- Tanwīn is still used by both groups but decreasing.

- Finally, *ħili:b* 'milk' is almost lost in favour of *ħali:b*.

To sum up, in Najd and Riyadh nowadays, the dialect background of speakers (whether B or S) can be easily identified by central Najdi people. If the Najdi koine becomes to some extent stable (i.e., most of the B and S dialects' features are levelled), the dialectal differences will reduce but not completely disappear, making it only possible for trained linguists to recognise these differences. In Table 8.5, I provide a predicted blueprint of the linguistic features of the Najdi koine. Needless to say, the koine is expected to be complete in Riyadh before Al-Kharj and other cities.

Table 8.5: *The expected linguistic features of Najdi koine*

Linguistic variable	B dialect	Najd koine	S dialect
The merger	/u/	/i/	/i/
3SM suffixed pronoun	*-ih*	*-ih/-ah*	*-ah*
Imperfect prefixes	*ya-*	*ya-/yi-*	*yi-*
Internal passive	almost lost	affixation	affixation
Guttural effect	GCa/GaC	GaC	GCa/GvC
Lowering of /e:/ and /o:/	/ɛ:/ and /a:/	/e:/ and /o:/	/e:/ and /o:/
Tanwīn	Retained	Ø	Retained
ħili:b / ħali:b	*ħili:b*	*ħali:b*	*ħali:b*

The differences between the two versions of the koine (B and S) are mainly reflected in the 3SM suffixed pronoun and the imperfect prefixes. Although it was found that Bs are converging on *-ah*, it cannot be determined when this change will be complete. As for the imperfect prefixes, each group (as hypothesised in Chapter 6) has inherited their variant from their ancestral dialects, and the data shows no convergence on *yi-* by the Bs. Therefore, it is claimed that this variation will not be levelled in the foreseeable future.

Below we provide examples of the B and S dialects before and after the koineisation. It must be noted that the examples for 'after the koine' are not merely speculation, but somewhat similar forms have been attested in the speech of converging Bs in our data. In the examples below, the linguistic variables are underlined.

Examples:

1- (A) *tibi*   *na-dxul*   *ʕind-ih*   *walla   la.* (B before koineisation)

   (B) *tibi*   *ni-dxil*   *ʕind-ah*   *walla   la.* (S before koineisation)

   (C) *tibi*   *na-dxil*   *ʕind-ih(-ah)*   *walla   la.* (B after koineisation)

   (D) *tibi*   *ni-dxil*   *ʕind-ah*   *walla   la.* (S after koineisation)

   Want-you   we-enter   at-him   or   not
     'Do you want to enter there or not?'

2- *la:*   *ta-nsa:*   *tdʒi:b*   *li:*   *ghawah*   *bi-lħili:b* (B before koineisation)

   la: *ti-nsa:*   *tdʒi:b*   *li:*   *ghawah*   *bi-lħali:b* (S before koineisation)

   la: *ta-nsa:*   *tdʒi:b*   *li:*   *gahwah*   *bi-lħali:b* (B after koineisation)

   la: *ti-nsa:*   *tdʒi:b*   *li:*   *gahwah*   *bi-lħali:b* (S after koineisation)

   don't   you-forget   you-bring for-me   coffee   with-milk
     'Don't forget to bring me coffee with milk'

3- (A) *wa:dʒba:t-ih*   *tiktab*   *whu:*   *dʒa:ls-in*   *ya-laʕab*   (B before koineisation)
              (Internal passive)   (tanwīn)

(B) *wa:dʒba:t-ah*   *tinkitib*   *whu:*   *dʒa:ls-in*   *yi-laʕab*   (S before koineisation)
              (affixational passive)   (tanwīn)

(C) *wa:dʒba:t-ih*   *tinkitib*   *whu:*   *dʒa:ls*   *ya-laʕab*   (B after koineisation)
              (affixational passive)

(D) *wa:dʒba:t-ah*   *tinkitib*   *whu:*   *dʒa:ls*   *yi-laʕab*   (S after koineisation)
              (affixational passive)

(Homework-his)   (written-passive)   (and-he) (he-just-IND) (he-plays)

'His homework is being written for him while he is just playing'

8.5 Sociolinguistics of Diglossia in Najd

As explained in sections 2.3 and 3.3.2.1, the Najdi community is diglossic: Standard Arabic is the most formal variety, while spoken Najdi dialect is the least formal. Some S speakers, motivated by several factors to be discussed below, switched to Standard Arabic during the sociolinguistic interview. As mentioned earlier in Chapters 4, 5, and 6, the B dialect shares two variants of the merger and imperfect prefixes (/u/ and *ya-*) with Standard Arabic; and we argue that the S speakers who used /u/ and *ya-* switched to Standard Arabic, not the B dialect.

Switching between the High and Low variety in diglossic communities is characterised as mainly situational, meaning that speakers use a distinct variety in particular settings (Ferguson, 1959b; Gumperz, 1977). Weinreich (1953, p. 73) argues that speakers code-switch according to the topic first and then the audience. But do all speakers code-switch under the same circumstances? As Gumpertz puts it, "[r]ather than claiming that speakers use language in response to a fixed predetermined set of prescriptions, it seems more reasonable to assume that they build on their own and their audience's abstract understanding of situational norms,

to communicate metaphoric information about how they intend their words to be understood."
(Gumperz, 1982, p. 61).

The setting of the sociolinguistic interview was conveyed to all participants as an informal speech event. Four S speakers, however, chose to switch to Standard Arabic, most of them probably being motivated by (a) the formality of the interview and (b) the topics discussed.[36]

The effect of the observer's paradox, especially when a recording device is present, cannot be overlooked. The setting of the sociolinguistic interview has clearly been perceived by some speakers as formal, hence Standard Arabic was used.

As for code-switching due to topic, according to Ferguson (1959b) religious events are an example of the situations in which the use of the High Variety is encouraged. The use of Standard Arabic in religious discourse and code-switching to the spoken dialect is well studied (e.g., Al-Aiyed, 2018; Alshamrani, 2012; Nazzal, 2001; Saeed, 1997; Walter, 1994). The topics of discussion in the interview were probably the main reason why those speakers switched to Standard Arabic. It should be noted that the suggested topics were mostly not religious, but it was the speakers' choice to steer the conversation to religious topics or discuss the proposed ones from a religious perspective. For example, the most memorable occasion for one of these speakers is when he completed memorising the holy book. Another example is that the hobby for one of these speakers was reading religious books. It was not surprising that such topics were discussed in Standard Arabic.

The last issue we will tackle in this section is the idea of intermediate levels, between High and Low. This means that in a stretch of discourse where there is code-switching, speakers tend to use intermediate forms which are neither pure Standard nor colloquial (Bassiouney,

---

[36] One of these speakers switched to Standard Arabic for a different reason than the others. He has a Ph.D. in Arabic language, and he explicitly commented on how he has been noticed to speak standard Arabic in informal situations.

2009, p. 13). Blanc (1960) recognises five levels: Classical Arabic, modified Classical, semi-literary colloquial, koineised colloquial and plain colloquial. Meiseles (1980) proposes that there are four varieties: literary Arabic, oral literary Arabic, educated speakers' Arabic and plain vernacular. In this study, we will not try to pinpoint which intermediate level the speakers used; instead, we aim to establish that the Najdi religious register is an intermediate level between Standard and Najdi Arabic. This can be seen in the use of the three main linguistic variables. As discussed in Chapters 4, 5 and 6, the merger and imperfect prefixes are realised as /u/ and *ya-* in the religious register, but 3SM suffixed pronoun is realised as *-ah*, which is the S Najdi form. As stated in Chapter 5, in order for these speakers to use the Standard Arabic form of the pronoun (*-hu*), the entire morphology of the stem must be changed; thus, it would require extra attention to speech. In the religious register, the use of the S dialect form *-ah* along with few other features, instead of the Standard Arabic ones, are the only examples we have of an intermediate variety. Certainly, there are other examples, but it is beyond the scope of this research.[37]

Below I will transcribe an excerpt from the interview of one of these speakers. The purpose here is to: (a) establish that the speakers indeed used Standard Arabic forms, not B, and (b) highlight the fact that they used an intermediate variety. The proof of the first point is manifested in the use of lexical items and phonemes from Standard Arabic (e.g., Standard Arabic /q/ and /ʔ/ instead of Najdi /g/ and Ø, respectively). These elements will be in bold, and below each one the corresponding colloquial form will be provided between brackets. The strongest indication of an intermediate level is seen in the use of the colloquial forms in the context of Standard Arabic. The first relevant example of such forms is the use of the 3SM suffixed pronoun *-ah*, instead of *-hu*. Other examples are the use of Najdi /ðˤ/ instead of

[37] The code-switching between Standard Arabic and Najdi Arabic in the speech of religious preacher has been studied by Al-Aiyed (2018).

Standard Arabic /dˤ/, the monophthong /e:/ instead of /ay/, and the grammatical morpheme *illi* 'that' instead of *allaði:*. These elements will be underlined in the excerpt, and below each one the Standard Arabic form will be provided between brackets.

Excerpt:

- The researcher:

In Najdi dialect: *do you think that there is exaggeration in the wedding expenses?*

- Speaker 14:

ʔara:	*ʔin*	*fi:h*	*muba:laʁah*	*waʔin*	*kanat*	*ta-xtalif*	*be:n*
(aʃu:f)	*(in)*			*(win)*		*(ti-xtilif)*	

| I-see | that | there | exaggeration | and-that | it-is | it-differ | between |

alfiʔa:t	*ʔaw*	*be:n*	*alʕawa:ʔil.*	*wa-la:kin*	*baʕ<u>ðˤ</u>*	*alʕawa:ʔil*
(inna:s)	*(walla)*		*(alʕawa:yil)*	*(bas)*	*(baʕ<u>dˤ</u>)*	*(alʕawa:yil)*

| the-groups | or | between | the-families | and-but | some | the-families |

qad	*taku:n*	*ʔaqal.*
(yimkin)	*(tku:n)*	*(ʔagal)*

| might | be | less |

'I see that there is exaggeration, but it differs between groups or families. But some families might be less'

- The researcher:

In Najdi dialect: *how much do you think a wedding should cost?*

- Speaker 14:

iða	*dʒa:ʔ*	*iʃʃaxsˤ*	*li-zzawa:dʒ*	*yuqaddiru:n*	*ʕala*	*ħasab*
	(dʒa:)	*(ilwa:ħid)*		*(ygaddru:n)*		
if	come	person	for marriage	they-estimate	on	according

iʃʃaxsˤ.	*illi*	*ra:tb-ah*	*tisiʕ-ʔa:la:f*	*ʁe:r*	*illi*	*ra:tb-ah*
(ilwa:ħid)	*(allaði:)*	*(ra:tibu-hu)*		*(ʁayr)*	*(allaði:)*	*(ra:tibu-hu)*
person.	that	salary-his	nine-thousands	different	from	that salary-his

xamstˤaʕaʃar-alf	*fa-yuqaddiru:n*	*ʕala*	*ħasab*	*al-ħa:lah*	*al idʒtima:ʕiyyah.*
	(fi-ygaddru:n)				
fifteen-thousands	therefore-they-estimate	on	according	status	social.

'When a person comes for marriage, they (the family of the bride) should estimate the expenses according to the person. The one whose salary is 9000 is different from that whose salary is 15000. Therefore, the family should estimate the expenses according to the social status of the groom.'

8.6 Conclusion

The B/S dialect distinction is still maintained in Najd. However, the Bs are converging towards the S dialect, to varying degrees across the examined linguistic variables. For the merger, it is speculated that Bs will completely converge on /i/, probably after two generations. For the 3SM suffixed pronoun and the imperfect prefixes, it is difficult to predict when the Bs will completely converge on the S variants, if they ever do.

The koine of Najd is expected to form after two to three generations. Nevertheless, the dialect background of Najdis will not completely disappear. The koine will most likely have a B and an S version, distinguished by the realisations of the 3SM suffixed pronoun and the imperfect

prefixes. For the rest of the variables, the koine will exhibit either S or Standard Arabic features, but not B ones.

The religious register was used in the sociolinguistic interview by a few S speakers. This register is an intermediate variety between the High and Low variety. As we have seen above, the speakers exhibited mixed forms, and their speech is characterised as neither pure Standard Arabic nor pure colloquial Arabic.

Conclusion

Summary of the Results

After over four decades of contact between the Bs and Ss in central Najd, the B-S linguistic dichotomy is still maintained. However, it was found that the Bs are converging towards the S dialect, to varying degrees across the linguistic variables. The main reason why the direction of convergence is B to S is that the S lifestyle is associated with modernity and education, to which Bs subconsciously aspire (Al-Sowayan, 2010, p. 345). This convergence has been statistically analysed (using *generalised linear mixed models* in *R*), and it was found that it is conditioned by different social factors for each variable.

The Bs' convergence on the S variant of the merger (/i/) is conditioned by the social factors of tribe, education and level of contact with the Ss (Chapter 4). The Dawasir tribe showed more convergence than the other tribes (Qahtani and Ajmi). This is attributed to the fact that each tribe forms a social network of its own, and that the change has spread in some networks (tribes) more than the others. Education is also an essential factor, which functions as a proxy for contact with other groups (Al-Wer, 2002a). That is, the level of education (college degree or not) determines whether individuals have been exposed to other dialect groups. It was found that Bs who did not attend college were less likely to converge on /i/. Lastly, the level of contact with the Ss, which is different from the type of contact caused by education, was also significant. Using a social network index, the Bs' level of contact with the Ss was measured based on first-order ties at school/work. It was found that Bs whose first-order ties inside and outside school/work did not include Ss were less likely to converge on /i/.

The Bs' convergence on the S variant of the 3SM suffixed pronoun (*-ah*) is quite low compared to the merger. The significant social factor conditioning this variation was age

(Chapter 5). It was found that young Bs are leading the change from *-ih* to *-ah*. We still cannot pinpoint the reasons why age is the explanatory factor for this variable, rather than tribe or contact, as per the merger. However, one possible explanation is that because *-ah* already exists in the B dialect (perfective paradigm), the change towards *-ah* is a natural drift due to redundancy. In other words, the perfective/ imperfective traits of the verb are reflected in the stem, and thus the use of *-ih* with imperfective verbs and *-ah* with perfective is redundant. Therefore, the leaders of this internally motivated change could be young speakers.

In the variable of imperfect prefixes (*ya-* / *yi-*) the Bs did not show any convergence on the S variant *yi-* (Chapter 6). As we have established, this variation (taltala) has been preserved for hundreds of years and is less likely to disappear soon. According to Rabin (1951), the dialects of the ancient tribes of Yemen, from which the current B tribes descend, were reported to feature *ya-*. On the other hand, the dialect of the ancient tribe of Tamim and other tribes in the centre of the Arabian Peninsula, from which the current S tribes descend, were reported to feature *yi-*.

As for sociolinguistic salience, the data from the SCAT, the online identification task and multiple interviews reveals that the merger is salient (i.e., it is a marker), while the other variables are not (indicators). These results agree with the interview data: the Bs showed considerable convergence on /i/, a fact which is consistent with its salience, while for the other variables there is slight or no convergence. The salience results are as follows:

- In the data from the SCAT, all Bs were able to associate the variants of the merger correctly with an average response time of 0.8 second, while for the other variables some participants mislabelled the variants, and the average response time was above 1 second for all variants.

- In the online identification task, the majority of participants were able to associate variants of the merger with the corresponding dialect correctly (30/34 for /i/ and 29/32 for /u/); however only 50% of the participants were able to associate variants of the 3SM suffixed pronoun correctly (17/34 for *-ih* and 16/34 for *-ah*). Unexpectedly, although the interview data for the variable of imperfect prefixes did not show any convergence, and the data from the SCAT indicates that it is not salient, a large number of participants were able to associate the variants correctly (26/34 for *ya-* and 29/34 for *yi-*).

- In the multiple interviews, the data from the second interviews (conducted by a S interviewer) was compared to the default production pattern, which is of the interviews conducted by a B. All 3 Bs showed a considerable rate of convergence on the S variant of the merger (/i/) in the default production pattern, and this rate increased significantly further when they were interviewed by a S speaker. As for the 3SM suffixed pronoun, one speaker showed convergence on *-ah* in the default pattern, and this rate increased when he was interviewed by a S speaker. The other two speakers did not show any use of *-ah* in either interview. Lastly, all B speakers used their variant of the imperfect prefix *ya-* consistently in both interviews.

Now we turn to the Ss. The data reveals consistent use of their variants, except for 4 speakers who code-switched to Standard Arabic, with which the B dialect shares two variants (/u/ and *ya-*). The reasons behind this code-switching are the formality of the interview and the topics discussed. These speakers perceived the interview as a formal (or semi-formal) event and preferred to chat about religious topics, and hence Standard Arabic was used. As established in Chapter 8, code-switching speakers usually use an intermediate variety, which is neither pure Standard Arabic, nor pure colloquial Arabic.

Based on the quantitative results and qualitative observations of the other linguistic variables, it is apparent that a Najdi koine has not yet been formed, but it is likely to be complete after two or three generations (Chapter 8). Nevertheless, there will be still B and S versions of the koine. When the koine is stable, it will most likely resemble S or Standard Arabic features, except for the imperfect prefixes which have been and will likely be preserved for a long time into the future.

Original Contributions of the Present Study

To the best of my knowledge, no empirical studies have been conducted on the sociolinguistic variation in central Najd, especially the current linguistic variables. This thesis is the first sociolinguistic investigation of the dialects of central Najd. The original contributions of the present study are the following:

- It has been empirically shown that the Bs are converging on the S dialect, and not vice versa.

- The study provides, based on evidence from the current data, a predictive linguistic description of Najdi koine.

- Linguistic variation among the B tribes of Najd has been discovered. The B tribes were then grouped into Bs who are originally from the southwestern parts of the Arabian Peninsula and WBs who are originally from the mid-western regions. This categorisation was based on whether theses tribes have retained or lost the opposition between /u/ and /i/. The B tribes (from the southwest) have retained the opposition between the vowels, while the WB tribes have merged them to /i/, as is the case in the S dialect.

- Tribal affiliation is stated to be a more essential social factor in Saudi Arabia than in most Arab countries (Bassiouney, 2009, p. 116). As far as I know, no study has

incorporated tribal affiliation as a factor in the quantitative analysis, either in Saudi Arabia or in countries where it is relevant. Tribal affiliation has been included as a factor in the present study, and the sociolinguistic variation between tribes in the merger variable was statistically significant. Tribal affiliation as a factor might correspond to social network in some cases, but there are several differences, which were discussed in 3.5.3.

- The merger of /u/ and /i/ has been acknowledged by many linguists, but it seemed to have received little attention sociolinguistically. In this study, we have explored the dialect contact between speakers of one dialect which has retained the opposition between /u/ and /i/ (B) and those of another in which the opposition is lost in almost all phonological environments (S).

- The realisations of the 3SM suffixed pronoun (-*ih* / -*ah*) are acknowledged in the literature to exist in Najd; however, this is the first study to associate them with the B/S linguistic dichotomy in Najd.

- The B assimilation of /a/ with /i/ as in *ħili:b* 'milk'(as opposed to S *ħali:b*) is a newly discovered feature of the B dialect of Najd. This feature is known to have existed in the ancient B dialects of the Arabian Peninsula, and we have found that it has been preserved in the Najdi B dialect (see section 8.3.5).

- Tanwīn is probably the most conservative feature of the B dialect of Najd. The data from this study, however, reveals that it is also preserved in the S dialect (8.3.4).

- This study also discusses the sociolinguistics of diglossia in Najd, which in this study is associated with the religious register.

Recommendations for Further Research

The WB tribes have been marginal to this study. They were found to share one feature with the S dialect and several more with the B dialect. The history and dialect of the WB tribes

239

could be further explored. As the original WB tribes still inhabit the mid-western parts of the Arabian Peninsula and small groups have migrated to central Najd, it is worthwhile studying to what extent they (WBs in central Najd) have accommodated to the B or S dialects.

It was pointed out to me by several S speakers that there are linguistic differences between regional S dialects within central Najd. For example, although the dialects of the city of Dir'iyah and Ad-Dlam are both S, they were said to be different in some aspects (e.g., the quality and length of some vowels). As a B Najdi speaker, it was difficult for me to spot these differences.

There are groups who migrated to central Najd from other parts of Saudi Arabia. For example, a large number of migrants from Qassim currently live in Riyadh Province. Although S in nature, the Qassimi dialect has been classified as a different regional dialect from central S (Ingham, 1994). Whether they have accommodated to the S dialect would also be a topic for further research.

Finally, female speakers have been excluded from the current study mainly due to cultural reasons. Because there is gender segregation in almost all social settings, female Najdi speakers are expected to exhibit different sociolinguistic behaviours from those of male speakers. Hence, the sociolinguistic characteristics of female speakers in Najd need to be further studied. Due to the conservativeness of the Najdi community, such studies are best conducted by a female researcher.

Appendix A

Sample of Data Transcription for One Speaker:

The words are transcribed using characters that are recognisable by R.

3 = ʕ, 9 = sˤ, 2 = ʁ, 0 = ðˤ, 6 = tˤ, 7 = ħ, x = χ, ? = ʔ

Speaker #	Ling Variable	Variant	Type of Variant	Word
8	merger (/u/ - /i/)	/i/	S	x-i-dh
8	Imperfect Prefix	*ya-*	B	ya-gli6uun
8	merger (/u/ - /i/)	/i/	S	sh-i-2il
8	Imperfect Prefix	*na-*	B	na-6la3
8	Imperfect Prefix	*ya-*	B	ya-r7im-ih
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ih*	B	ya-r7im-ih
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ih*	B	nzuur-ih
8	merger (/u/ - /i/)	/i/	S	k-i-ll
8	merger (/u/ - /i/)	/i/	S	k-i-nt
8	Imperfect Prefix	*na-*	B	na-sma3
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ah*	S	?a7fi0-ah
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ah*	S	?agu:l-ah
8	merger (/u/ - /i/)	/i/	S	g-i-99ah
8	merger (/u/ - /i/)	/u/	B	g-u-99athum
8	Imperfect Prefix	*ya-*	B	ya-b2a
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ih*	B	3amm-ih
8	Imperfect Prefix	*ya-*	B	ya-nfa3
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ih*	B	rab3-ih
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ih*	B	to:rid-ih
8	merger (/u/ - /i/)	/u/	B	k-u-llaha
8	3MS suffixed pronoun (*-ih* / *-ah*)	*-ih*	B	3imdaan-ih
8	Imperfect Prefix	*ya-*	B	ya-r7al

Social Network Index:

Participant: # [ ]

Group	Tribe	Age	Education	Job

- Mother from the same tribe? | Yes | No |

- Typically, how many hours do you spend at home?

Less than 3 hrs	4-5 hrs	5-6 hrs	More than 6 hrs

- Who lives in your neighborhood? do you interact with them?   Yes   No

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- Who do you spend time with at work/school?

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- Who do you spend time with outside work?

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- How many times do you visit members of your big family?

Every year	Every 6 mths	Every 3 mths	Every mth or more
3	2	1	0

For Bedouins

Do you have camels?

yes	no

How many times do you go to the camels?

Every month or more	Every 2 weeks	Weekends	Every 3 days or more

Sample of Social Network Index:

SOCIAL NETWORK INDEX

Participant: # 4

Group	Tribe	Age	Education	Job
S	Arrunis	18	Secondary	Student

- Mother from the same tribe?   Yes   No

- Typically, how many hours do you spend at home?

Less than 3 hrs	4-5 hrs	5-6 hrs	More than 6 hrs
			✓

- Who lives in your neighborhood? do you interact with them?   Yes   *No*

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- Who do you spend time with at work/school?

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	*1*	0

- Who do you spend time with outside work?

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	*1*	0

- How many times do you visit members of your big family?

Every year	Every 6 mths	Every 3 mths	Every mth or more
3	2	1	*0*

For Bedouins

Do you have camels?

yes	no

How many times do you go to the camels?

Every month or more	Every 2 weeks	Weekends	Every 3 days or more

SOCIAL NETWORK INDEX

Participant: # **8**

Group	Tribe	Age	Education	Job
B	9	17	Secondary	Stu

- Mother from the same tribe? | Yes | No |

- Typically, how many hours do you spend at home?

Less than 3 hrs	4-5 hrs	5-6 hrs	More than 6 hrs
	✓		

- Who lives in your neighborhood? do you interact with them?   (Yes) No

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- Who do you spend time with at work/school?

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- Who do you spend time with outside work?

Not relatives or same tribe	Few of same tribe	Mostly same tribe	Mostly relatives
3	2	1	0

- How many times do you visit members of your big family?

Every year	Every 6 mths	Every 3 mths	Every mth or more
3	2	1	0

For Bedouins

Do you have camels?

yes	no

How many times do you go to the camels?

Every month or more	Every 2 weeks	Weekends	Every 3 days or more
			✓

Pictures for the Elicitation Task of Guttural Effect:
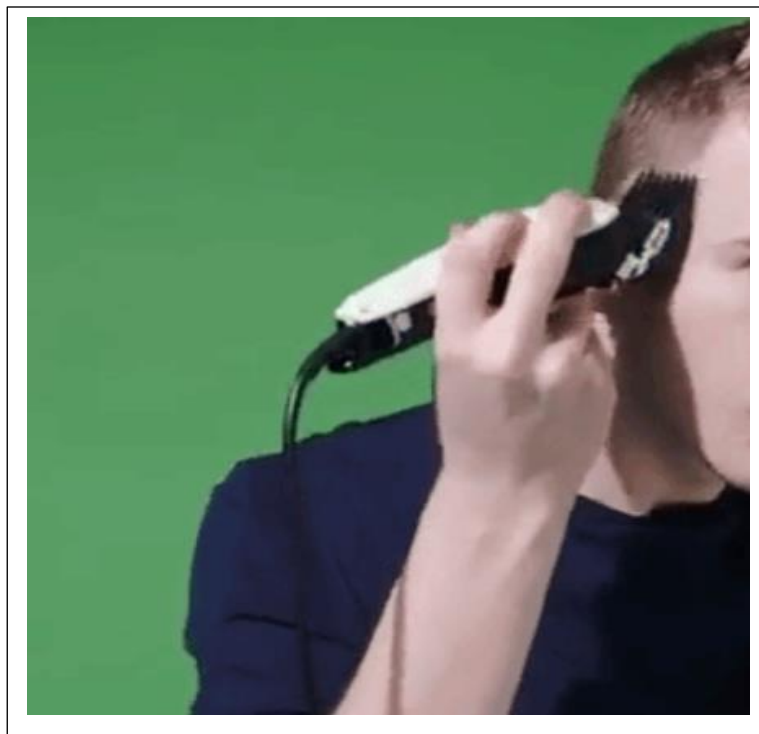
Target word: *yħalig* 'to shave'

Target word: *yħarθuːn* 'to plow'

Target word: *yʕagid* 'to make a knot'

Target word: *yhadim* 'to demolish'

Target word: *tʁazil* 'wool spinning'
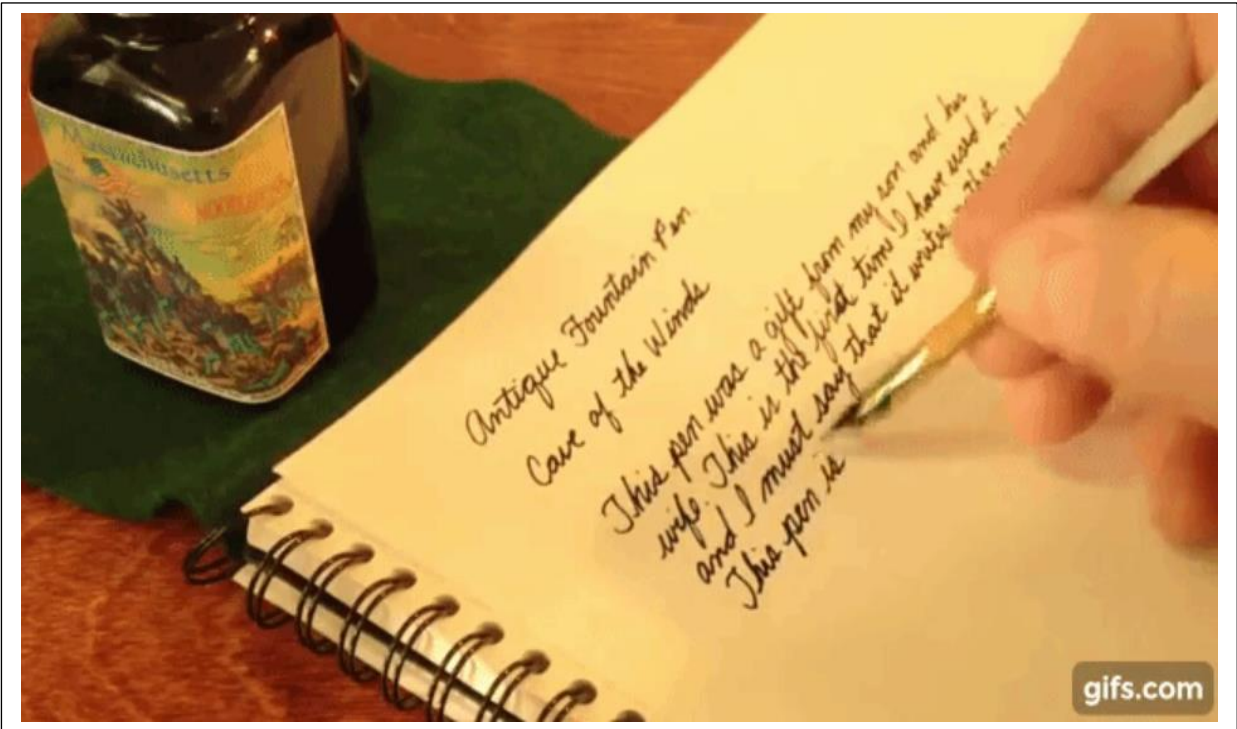
Target word: *yaktib* 'to write' **(Distractor)**

Target word: *yxalitˤ* 'to mix'

Target word: *yxatim* 'to seal'

Target word: *yʕaṣʕir* 'to squeeze'

Target word: *yxabiz* 'to bake'

Target word: *yxarif* 'to harvest dates'

Target word: *yaksir* 'to break' **(Distractor)**

Target word: *yʕadʒin* 'to knead'

Target word: *yħasib* 'to count'

Target word: *yħalib* 'to milk'

Target word: *yħafir* 'to dig'
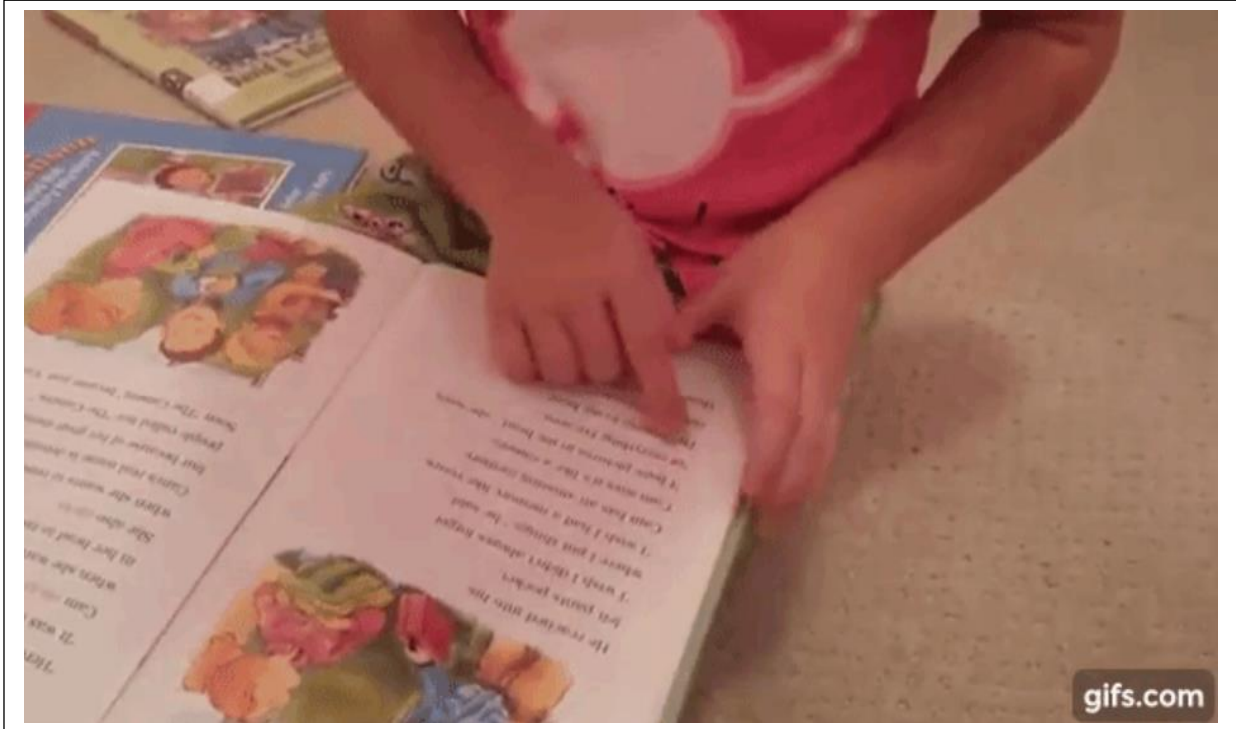
Target word: *tagra:* 'to read' (**Distractor**)

Target word: *yʁasil* 'to wash'

The Picture Used for the Elicitation of Lowered Versions of /eː/ and /oː/:

Target words: *fɛːsˤal* 'King Faisal', *xɛːmah* 'tent', *nibɛːtˤah* 'slingshot', *ħɛːtˤah* 'free kick wall', *ʁɛːmah* 'cloud', *sˤiffɛːrah* 'whistle', *bɛːðˤah* 'an egg'.

Checklist to Help the Researcher Conduct the Interview and the SCAT:

Interview

- Talk about research & information sheet.
- Start recording & oral consent.
- Social network index.
- Interview.
- Elicitation task.

Experiment

- Explain experiment.
- Enter participant number and code.
- Do trial.
- Do the experiment.

Appendix B

Copy of the Online Questionnaire:

Q.1

السلام عليكم و رحمة الله وبركاته

يهدف هذا الاستبيان الممتع لمعرفة مدى إمكانية سكان منطقة نجد في تحديد الفروق بين لهجة البدو و الحضر في نجد.

قم أولا بإدخال معلوماتك في الأسفل, بعد ذلك ستستمع إلى مقطع صوتي لشخص ما ثم يطلب منك إختيار هل هو بدوي أو حضري أو غير ذلك . بعد كل مرة سيطلب منك اختيار مستوى تأكدك من الإجابة. شكرا لمشاركتك

Q.2   الرجاء ادخال اسم القبيلة

في حال عدم الرغبة, الرجاء تحديد هل أنت من الحاضرة أو البادية أو غير ذلك

بإمكانك إدخال أكثر من خيار

☐ القبيلة [_____]

☐ أنا من حاضرة نجد

☐ أنا من بادية نجد

☐ غير ذلك [_____]

Q.3   العمر؟

○ 15-25

○ 26-35

○ 36-45

○ 46-55

○ أكبر من 55

Q.4   التعليم؟

○ جامعي أو أعلى

○ ثانوي

○ أقل

▶ 0:00 / 0:00 ● 🔊 ⋮

Q.5

هل الشخص هذا؟

○ بدوي

○ حضري

○ غير ذلك (الرجاء التحديد) [_____]

○ لا أعلم

256

Q.6   مامدى تأكدك من الإجابة السابقة

غير متأكد   متأكد جدا
0   1   2   3   4   5   6   7   8   9   10
○   ○   ○   ○   ○   ○   ○   ○   ○   ○   ○

Q.7

▶ 0:00 / 0:00 ●━━━━━ 🔊 ⋮

هل الشخص هذا؟

○ بدوي

○ حضري

○ غير ذلك (الرجاء التحديد) [_____]

○ لا أعلم

Q.8   مامدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا
0   1   2   3   4   5   6   7   8   9   10
○   ○   ○   ○   ○   ○   ○   ○   ○   ○   ○

▶ 0:00 / 0:01 ●━━━━━ 🔊 ⋮

Q.9
هل الشخص هذا؟

○ بدوي

○ حضري

○ غير ذلك (الرجاء التحديد) [_____]

○ لا أعلم

Q.10   مامدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا
0   1   2   3   4   5   6   7   8   9   10
○   ○   ○   ○   ○   ○   ○   ○   ○   ○   ○

257

▶ 0:00 / 0:00 ● 🔊 ⋮

Q.11

هل الشخص هذا؟

○ بدوي

○ حضري

○ غير ذلك (الرجاء التحديد) [_____]

○ لا أعلم

Q.12   مامدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا

0   1   2   3   4   5   6   7   8   9   10

○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

▶ 0:00 / 0:00 ● 🔊 ⋮

Q.13

هل الشخص هذا؟

○ بدوي

○ حضري

○ غير ذلك (الرجاء التحديد) [_____]

○ لا أعلم

Q.14   مامدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا

0   1   2   3   4   5   6   7   8   9   10

○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

▶ 0:00 / 0:00 ● 🔊 ⋮

Q.15

هل الشخص هذا؟

○ بدوي

○ حضري

○ غير ذلك (الرجاء التحديد) [_____]

○ لا أعلم

Q.16   ما مدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا
0   1   2   3   4   5   6   7   8   9   10
◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯

▶   0:00 / 0:00   ━━●━━━   🔊   ⋮

Q.17

هل الشخص هذا؟

◯ بدوي

◯ حضري

◯ غير ذلك (الرجاء التحديد) [_____]

◯ لا أعلم

Q.18   ما مدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا
0   1   2   3   4   5   6   7   8   9   10
◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯

▶   0:00 / 0:00   ━━●━━━   🔊   ⋮

Q.19

هل الشخص هذا؟

◯ بدوي

◯ حضري

◯ غير ذلك (الرجاء التحديد) [_____]

◯ لا أعلم

Q.20   ما مدى تأكدك من الإجابة السابقة؟

غير متأكد   متأكد جدا
0   1   2   3   4   5   6   7   8   9   10
◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯   ◯

Appendix C

- Language & Linguistic Science ethics committee research approval.

- Consent form in English.

- Information sheet in English.

- Consent form in Arabic.

- Information sheet in Arabic.

**Language
Science ethics**

UNIVERSITY *of York*

**&Linguistic
committee**

research approval request form

If you are collecting any data from human subjects, you must fill out this form. You may ***not*** begin data collection until you receive notification that your research has been approved.

It is your responsibility to familiarise yourself with the University's Code of practice and principles for good ethical governance. There is a set of Guidance Notes accompanying this form; it is recommended that you read through the notes first to help you plan your responses to the questions, and to reflect on any ethical questions raised by your planned research. The particulars of your answers to some of the questions may differ depending on the research design of your individual project, but the Guidance Notes offer advice and in some cases sample wording for answers. If you have any additional queries, do not hesitate to contact your supervisor or the Chair of the Ethics Committee at <linguistics-ethics@york.ac.uk>

NB: If you are collecting data from NHS patients or staff, or Social Service users or staff, you will need to apply for approval through the Integrated Research Application System (IRAS) at https://www.myresearchproject.org.uk/Signin.aspx
- Staff: please fill in the IRAS form first and send your completed form to L&LS
- Students: applications for approval through IRAS should normally be **pre-reviewed** by the L&LS ethics committee

Completed forms should be submitted to the L&LS Ethics Committee as follows:
1. **One electronic copy of the compliance declaration and one electronic copy of the Ethics form** including the Information sheet and consent form entitled "SURNAME ethics submission MONTH YEAR" via email to <linguistics-ethics@york.ac.uk> Any other questionnaire submitted should be pasted at the end of the ethics form to create a single submitted document.
2. **One signed Compliance form**, to be submitted to the department office, **once approval has been granted**.

Initial decisions will normally be made and communicated within two weeks of the receipt of your complete application.

SUBMISSION FORM

1a. Please provide the following details about the principal investigator at YORK

Name of Applicant:	**Nasser Mohammed Alajmi**
Email address:	**nma514@york.ac.uk**
Telephone:	
Staff/Student Status:	**PhD Student**
Dept/Centre or Unit:	Language & Linguistic Science
Head of Department:	Dunstan Brown, dunstan.brown@york.ac.uk
Chair of Dept. Research Committee:	Marilyn Vihman, marilyn.vihman@york.ac.uk

1b. Any other applicants (for collaborative research projects)

Name of Applicant:	
Email address:	
Telephone:	
Staff/Student Status:	
Dept/Centre or Unit:	
Head of Department:	
Head of Research: (if applicable)	

2. (STUDENTS ONLY) Please provide the following supervisory details for your project:

PLEASE NOTE: By submitting this form, you confirm that your supervisor has approved the summary of your research proposal and the participant information sheet.

1st Supervisor, email address	Paul Kerswill, **paul.kerswill@york.ac.uk**
2nd Supervisor, email address	

3. Please provide the following details about your project:

Title of Project:	The Bedouin/Sedentary Dichotomy in Najd: A Sociolinguistic Study
Date of Submission to L&LS Ethics Committee:	
Project Start Date:	
Duration:	
Funded Yes/No:	**No**
If funded, give source:	
External Ethics Board Jurisdictions:	

4.   *Summary of research proposal*

4.1 Aims and objectives of the research
Please outline the questions or hypotheses that will be examined in the research.

(PLEASE DO THIS **IN A WAY THAT WILL BE COMPREHENSIBLE TO NON-LINGUISTS.**)
[500 words max.]

In Najd, the central region of Saudi Arabia, the dialects have been classified into Bedouin and Sedentary dialects four decades ago before the oil boom. After the oil boom in the 1970s, the rapid rate of urbanization has forced both groups to give up their old lifestyles and live side by side in the city. In the field of language variation and change, such cases give rise to several research questions some of which are:

- Is there still a systematic linguistic variation between Bedouin and Sedentary dialects?
- What are the linguistic features of each dialect?
- What are the salient linguistic features of each dialect (i.e., people are aware of)?
- Which group accommodates (in speech) to the other?

Only male speakers will be examined and recorded for the following considerations:
- The two genders in the entire targeted community are completely segregated, which is also the case in most regions in Saudi Arabia. Therefore, women as a semi-isolated group have probably developed their own linguistic norms.
- It is also culturally and religiously difficult for the researcher to approach women and record them.

4.2 Methods of data collection
Outline how the data will be collected from or about human subjects.

The data will be collected by the means of (1) <u>sociolinguistic interview</u> (usually one-on-one interview, conducted in person, in which the interviewee will be asked questions) and given picture naming tasks (a task in which participants will be shown pictures and asked what is the person in the picture doing?) and (2) <u>sociolinguistic salience test</u> (an experimental design in which subjects listen to different audio stimuli (i.e., words) and are asked to associate each word with a dialect. The experiment will take 3-5 minutes. The researcher will recruit friends to produce 20 stimuli (neutral words like 'all' and 'remember'), which will be evenly divided to 10 Bedouin and 10 Sedentary. The experiment will be run using Psychopy. The procedure of the experiment is as follows: listeners will listen to the first stimulus and then be asked to click on the right mouse button if he thinks the speaker is Bedouin or left button if he thinks the speaker is Sedentary - and they will continue to do this for the rest of stimuli. The participants will be interviewed first and will then participate in the salience experiment.

4.3 Recruitment of participants
How many participants will take part in the research? How will they be identified and invited to take part in the study? How will informed consent be obtained?

There will be 50 male participants, who will be chosen based on the following criteria:
- Born and Live in Al-Kharj (a small city in Najd and the hometown of the researcher)

- Fall into one of two age ranges (18-30) and (50-70).
- Belong to specific tribes.

The researcher will get help from friends and relatives in Al-Kharj who will be asked to look for participants who match the criteria. Identifying whether a participant is Bedouin or Sedentary will be based on tribal names (last names) which is a strong indication of their group affiliation. Their contacts will be given to the researcher and they will be contacted to set a time for the interview. Hence, when approaching the participants, the researcher will then have the advantage of being a member of the community and coming through a friend or relative.

Due to cultural and political considerations, it will weaken the position of the researcher and may cause participants to withdraw if the consent form will be handed to them in paper form. In other words, people in Saudi Arabia are very suspicious when it comes to signing papers. Therefore, it would be best if their consent will be taken in an audio recording. Before or at the beginning of the interview, the researcher will read a list of statements to the participants and they should respond to them by *yes* or *no*. The consent form and the translation is included below.

4.4 Participant information sheets and consent forms
Please attach (to the end of this form):
(1) the project information sheet to be given to all participants and
(2) the informed consent form **(n.b. failure to submit these documents will result in your application being returned without decision).**

i. Please confirm you have included the participant information sheet and the consent forms to be given to all participants with your submission to L&LS. If these have not been attached, please explain why this is the case

The information sheet (and the translation) has been attached and will be handed to them before the beginning of the interview. The consent form (and the translation) has been attached but it will be read to the participants instead of signing papers, as specified above.

ii. Are the results to be given as feedback or disseminated to your participants (if yes please specify when, in what form, and by what means)

Generally no, but if someone if interested I can tell them the results which will be in the form of statements and generalizations about the research.

4.5 Anonymity
In most instances the Committee expects that anonymity will be offered to research subjects. Please set out how you intend to ensure anonymity. If anonymity is not being offered please explain why this is the case.

The subjects will be offered anonymity. I will record the subject's name and contact details, but store these in an encrypted file separate from the main data file. These details will be linked to the data file via a code letter/number.

4.6 Data Collection

All personal and sensitive data must be collected and stored in accordance with the Data Protection Act 1998. Please set out all the types of data you will be collecting (e.g. interviews, questionnaires, recordings)

i. Please detail type(s) of data.

The data will be on the form of audio recordings for the interview and Excel sheets for the experiment. Additionally, the researcher will collect information like age, job and tribal affiliation in paper prior to the interview, but it will be transcribed in codes (with no names) in an Excel sheets later .

ii. Where is the data to be collected and where will it be stored electronically?
Please describe what protection there will be in relation to electronic storage.

The data will be collected in Saudi Arabia. It will be stored by the researcher electronically. The recordings, the papers (which will be scanned) and the excel sheets will be stored on the researcher's personal laptop, which is password protected and additionally encrypted using Bitlocker. The data will be also backed up in the Google Drive account of the researcher and which is password protected as well. The data will be transferred to the university server as soon as possible. There will be no names except in the separately encrypted file (see 4.5) and the data will be coded as needed. The separate encrypted file will not be stored on the researcher's own computer but rather on Google Drive only.

iii. Where is the data to be stored in paper form? Please describe how this will be protected.

Once the papers for the experiments are scanned and backed-up, they will be destroyed.

iv. At what point do you propose to destroy the data?

I am not destroying my data. In sociolinguistics, it is necessary for the purpose of further research not to destroy data.

v. If you are sharing data with others outside your department, what steps are you taking to ensure that it is protected?

I am not sharing my data.

vi. If the data is to be exported outside the European Union, what steps are you taking to ensure that it is protected? **(Note: you must identify how you will comply with Data Protection Act 1998 requirements.)**

My data will be collected outside the European Union and will be protected as indicated

above.

4.7 Perceived risks or ethical problems
Please outline any anticipated risks or ethical problems that may adversely affect any of the participants, the researchers and or the university, and the steps that will be taken to address them. **(Note: all research involving human participants can have adverse effects.)**
i. Risks to participants (e.g. emotional distress, financial disclosure, physical harm, transfer of personal data, sensitive organisational information…)

If a participant talk about a sensitive issue, the researcher will delete this part form the recording immediately after the interview.

ii. Risks to researchers (e.g. personal safety, physical harm, emotional distress, risk of accusation of harm/impropriety, conflict of interest…)

Most participants are direct friends of a friend (or relative). Most of the time, the mediator will be present. If not, the researcher will inform a family member of the who, where and when of the interview.

iii. University/institutional risks (e.g. adverse publicity, financial loss, data protection…)

I anticipate no risks.

iv. Financial conflicts of interest (e.g. perceived or actual with respect to direct payments, research funding, indirect sponsorship, board or organisational memberships, past associations, future potential benefits, other…)

None.

v. Please draw the committee's attention to any other specific ethical issues this study raises.

None.

5. Ethics checklist

Please confirm that all of the steps indicated below have been taken, or will be taken, with regards to the above named project submitted for ethical approval. If there are any items that you cannot confirm, or are not relevant to your project, please use the space provided below to explain.

Please tick if true, otherwise leave blank:

✔ Informed consent will be sought from all research participants where appropriate

✔ All data will be treated anonymously and stored in a secure place

✔ All Relevant issues relating to Data Protection legislation have been considered (see http://www.york.ac.uk/recordsmanagement/dpa/)

✔ All quotes and other material obtained from participants will be anonymised in all reports/publications arising from the study where appropriate

✔ All reasonable steps have been taken to minimise risk of physical/psychological harm to project participants.

✔ All reasonable steps have been taken to minimise risk of physical/mental harm to researchers

✔ Participants have been made aware of and consent to all potential futures uses of the research and data

☐ With respect to indemnity Sue Final (University IP Manager **email:**sue.final@york.ac.uk)has been made aware of the research, **where appropriate**

✔ There are no known conflicts of interest with respect to finance/funding

> If applicable, please explain in the space below why any of the above items have not yet been confirmed:

6. Other comments

Are there any issues that you wish to draw to the Committee's attention (it is your responsibility to draw the committee's attention to any ethical issues that may be of perceived or actual interest)?
None.

7. Submission Checklist for Applicants

Finally, please **sign** the form and ensure that **all of the indicated documents** below are sent **electronically** to via email to <linguistics-ethics@york.ac.uk>.

✓ L&LS Application form

✓ Consent form for participants (will be presented orally)

✓ Information Sheet for participants

8. Signed undertaking

In submitting this application I hereby confirm that there are **no actual or perceived conflicts of interest** with respect to this application (and associated research) other than those already declared.

Furthermore, I hereby undertake to ensure that the above named research project will meet the commitments in the checklist above. In conducting the project, the research team will be guided by the Social Research Association's/AHRC's/ESRC's ethical guidelines for research.

Nasser Mohammed Alajmi   (Name of Lead Researcher/Principal Investigator)

30/5/2017   (Date)

Paul Kerswill   (Name of Supervisor (if relevant))

30/5/2017   (Date)

The Sociology of Najd

Lead researcher: Nasser Alajmi

Consent form

The researcher will read the following questions to the participants and they will be required to answer with *yes* or *no*. The researcher will tell the participants that if there is anything they do not understand, or if they want more information, they should ask the researcher.

Have you read and understood the information leaflet about the study?	Yes ✿ No ✿
Have you had an opportunity to ask questions about the study and have these been answered satisfactorily?	Yes ✿ No ✿
Do you understand that the information you provide will be held in confidence by the research team, and your name or identifying information about you will not be mentioned in any publication?	Yes ✿ No ✿
Do you understand that you may withdraw from the study at any time before the end of the data collection session without giving any reason, and that in such a case all your data will be destroyed?	Yes ✿ No ✿
Do you understand that the information you provide may be kept after the duration of the current project, to be used in future research on language?	Yes ✿ No ✿
Do you agree to take part in the study?	Yes ✿ No ✿
If yes, do you agree to your interview being recorded on audio?	Yes ✿ No ✿
Do you agree to excerpts from your audio recordings to be used in presentations or in teaching by the researcher, without disclosing your real name? *(You may take part in the study without agreeing to this).*	Yes ✿ No ✿
Do you agree to the researcher's keeping your contact details after the end of the current project, in order that s/he may contact you in the future about possible participation in other studies? *(You may take part in the study without agreeing to this).*	Yes ✿ No ✿

UNIVERSITY *of York*

DEPARTMENT OF
LANGUAGE AND
LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
nma514@york.ac.uk

INFORMATION SHEET

PLEASE KEEP THIS INFORMATION SHEET AND A SIGNED COPY OF THE CONSENT FORM
FOR YOUR RECORDS

You are invited to take part in a research study. Before you decide whether to participate it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully. If there is anything you do not understand, or if you want more information, please ask the researcher.

Title of study:

The Sociology of Najd

Researcher:

Nasser Mohammed Alajmi

Ph.D. Student at the University of York

What is the research about?

The research is about the sociology of Najd including old and current lifestyles, classification of social groups and language.

Who is carrying out the research?

The researcher is a student at the linguistics department in the University of York who needs this data for his research.

Why have you been chosen to participate?

You are chosen because you are born and raised in Najd.

What does the study involve?

The study will involve a one-on-one interview which will take 45-60 minutes and a picture naming task (where you will simply name what is happening in a picture) that will take 5-7 minutes. After that, you will participate in a 5 minutes experiment in which you listen to some recorded voices and be asked to choose whether the speaker is Bedouin or Sedentary.

Do I have to take part?
You do not have to take part in the study. If you do decide to take part you will be given this information sheet to keep and will be asked to state in audio recording that you agree to take part in the interview and experiment. If you decide to take part you will still be free to withdraw without giving a reason, even during the session itself. If you withdraw from the study, we will destroy your data and will not use it in any way.

What are the possible risks of taking part?

There are no possible risks. You will be interviewed one-on-one and do the audio experiment. Whenever you feel uncomfortable you can withdraw.

Are there any benefits to participating?
you will help the researcher understand sociology of Najd which is unknown to sociolinguists and to the world.

What will happen to the data I provide?
After analysed and compared statistically, the data you provide will be used alongside the data of other participants to help us understand the sociology of Najd. Your data will be stored securely in the University of York, Department of Language and Linguistic Science. The data will be kept for further research, but your identity will still be anonymous in the data.

What about confidentiality?

Your identity will be kept strictly confidential. No real names will be used in any presentations, publications or in my dissertation.

Will I know the results?

The results of my analysis of the data collected from all participants can be given to you after I finish my dissertation. However, I cannot give you the data, results, or analysis of other particular participants.

*This study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science at the University of York. If you have any questions regarding this, you can contact the chair of the L&LS Ethics Committee, Márton Sóskuthy, (**email:** marton.soskuthy@york.ac.uk; **Tel:** (01904) 324171).*

If you have further questions regarding this study, please feel free to contact:

<div align="center">

Nasser Mohammed Alajmi
Department of Language and Linguistic Science
University of York, Heslington, York, YO10 5DD

email: nma514@york.ac.uk

Supervisor:
Professor Paul Kerswill
paul.kerswill@york.ac.uk

</div>

دراسة المجتمع النجدي

الباحث الرئيسي: ناصر بن محمد العجمي

استمارة الموافقة على اجراء دراسة

سوف يقوم الباحث بقراءة الأسئلة التالية على المشاركين و يكون عليهم الإجابة بنعم أو لا. سيخبر الباحث المشاركين أنه في حالة استصعب عليهم فهم شيء أو أرادوا الإ ستزادة بمعلومات عن البحث فإنه بإمكانهم سؤال الباحث.

هل قرأت وفهمت نشرة المعلومات الموضحة ☐ نعم ☐ لا للدراسة؟

هل أتيحت لك الفرصة لطرح أسئلة عن الدراسة، وهل تمت الاجابة عليها بشكلٍ مرضٍ؟   نعم ☐   لا ☐

هل أنت على دراية بأن المعلومات التي ستقدمها لنا ستحفظ بشكل متقن من قبل فريق البحث، ولن يتم ذكر اسمك او المعلومات الخاصة بك في وأي وعاء نشر؟   نعم ☐   لا ☐

هل أنت تعلم بإمكانك الانسحاب من الدراسة في أي وقت قبل نهاية جمع البيانات دون إبداء أي سبب، وأنه في مثل هذه الحالة سيتم التخلص من جميع البيانات الخاصة بك؟   نعم ☐   لا ☐

هل أنت تعلم ان المعلومات التي ستقدم لنا قد تبقى مدة المشروع او بعده لاستخدامها في البحث المستقبلية في اللغة؟
نعم ☐ لا ☐

هل انت موافق على المشاركة في الدراسة؟   نعم ☐   لا ☐

إذا كانت الاجابة بنعم، هل توافق على تسجيل مقابلتك صوتياً؟   نعم ☐   لا ☐

هل توافق على أخذ مقتطفات من تسجيلك الصوتي وذلك لاستخدامها من قبل الباحث في العروض التقديمية او التدريس، وبدون الكشف عن اسمك الحقيقي؟   نعم ☐   لا ☐
(بإمكانك المشاركة في الدراسة بدون هذه الموافقة).

هل توافق على الاحتفاظ ببيانات الاتصال بك بعد نهاية هذا المشروع؛ وذلك حتى يتمكن التواصل معك في المستقبل بغرض المشاركة المتحملة منك في دراسات أخرى؟   نعم ☐   لا ☐

UNIVERSITY *of York*

**DEPARTMENT OF
LANGUAGE AND
LINGUISTIC SCIENCE**

Heslington, York, YO10 5DD, UK
nma514@york.ac.uk

قسم اللغات وعلم

اللغة بجامعة يورك

ببريطانيا

معلومات عن البحث

أنت مدعو للمشاركة في هذا الدراسة العلمية. قبل أن تقرر في المشاركة في الدراسة من المهم أن تعرف على ماذا تشمل هذه الدراسة و ما هي أهدافها. الرجاء قراءة المعلومات التالية بتأني وفي حال وجود أسئلة أو استفسارات لا تتردد في سؤال الباحث.

عنوان البحث :

دراسة المجتمع النجدي

الباحث:

ناصر محمد العجمي

طالب دكتوراة بجامعة يورك قسم اللغات

عن ماذا يتكلم البحث؟
البحث يهدف إلى دراسة و فهم تركيبة المجتمع النجدي.

من الذي يقود البحث؟
البحث هو عبارة عن جزء من رسالة دكتوراة يقوم بها الباحث في مجال المجتمع و اللغة.

لماذا تم اختيارك للمشاركة في الدراسة؟
سبب اختيارك هو انك مولود و تعيش في نجد

على ماذا تحتوي الدراسة؟
تتكون الدراسة من جزئين من جمع البيانات. الجزء الأول هو مقابلة شخصية مع الباحث يتم فيها سؤالك أسئلة عامة عن حياتك ثم سيطلب منك التعرف على مجموعة من الصور. المقابلة سوف تستمر لمدة تتراوح بين 45 إلى 60 دقيقة. الجزء الثاني هو تجربة علمية بسيطة حيث يطلب منك الإستماع إلى جمل و عبارات و يطلب منك الإجابة على أسئلة متعلقة بها.

هل يتوجب على المشاركة؟
لا. لا يتوجب عليك المشاركة. إذا وافقت على المشاركة, ستعطى لك أوراق المعلومات هذه و سيطلب منك الإقرار على أنك وافقت على المشاركة عن طريق تسجيل صوتك و أنت ترد بالموافقة على المشاركة. بعد موافقتك يحق لك العدول عن رأيك و الانسحاب من المشاركة سواءا قبل بداية الدراسة أو في أثناءها أو حتى بعد الانتهاء منها. في حال إنسحابك سيتم التخلص من التسجيلات و جميع بياناتك و لن تستخدم أبدا في البحث.

ماهي المخاطر المترتبة على المشاركة في الدراسة؟
لا يوجد أي مخاطر من المشاركة في الدراسة هذه. هي بكل بساطة مقابلة شخصية و تجربة صوتية و يحق لك الإنسحاب في أي وقت.

هل هنالك أي فوائد من المشاركة في الدراسة؟
في الغالب ليس هنالك فوائد للمشارك ولكنك ستساعد الباحث على فهم و دراسة المجتمع النجدي اللتي تعتبر مجهولة نوعا ما لعلماء المجتمع و اللغة.

ماذا سيحصل للبيانات اللتي ستؤخذ مني؟

274

البيانات اللتي ستؤخذ منك سيستخدمها الباحث في دراسته و سيقارنها مع بيانات المشاركين الآخريين ليتوصل إلى استنتاجات حول مجتمع نجد. البيانات ستخزن مع الباحث أثناء تحليلها و دراستها. بعد ذلك سيتم تخزين البيانات هذه في موقع الجامعة و ذلك للحفاظ على سريتها و استخدامها لدراسات قادمة إن دعت الحاجة.

<u>ماذا عن سرية المعلومات و الحفاظ على هويتي؟</u>
هويتك ستكون مجهولة تماما و لن يتم ذكر اسمك في أي عروض أو مقارنة بحوث أو رسالة علمية. سيتم ربط بياناتك برقم مثل (مشارك رقم 3) ولن يتم ذكر اسمك على الإطلاق.

<u>هل سأعرف النتيجة؟</u>
نتيجة تحليل بياناتك و بيانات المشاركين الآخرين ستعطى لك إذا طلبتها بعد الإنتهاء من رسالة الدكتوراة. يجب التنبيه على أنه ليس من حقي ولا من حقك أن أعطيك بيانات الأشخاص الاخرين أو شخص معين.

<u>هذه الدراسة تمت مراجعتها من قبل لجنة أخلاقيات البحث العملي بقسم اللغات و علم اللغة بجامعة يورك ببريطانيا. إذا كان لديك أي أسئلة</u> تجاه هذا الإعتماد يمكنك التواصل مع رئيس اللجنة:
*Márton Sóskuthy, (**email:** marton.soskuthy@york.ac.uk; **Tel:** (01904) 324171).*

إذا كان لديك أسئلة عن الدراسة لاتتردد في سؤال الباحث:

ناصر محمد العجمي
قسم اللغات و علم اللغة بجامعة يورك ببريطانيا

ايميل: *nma514@york.ac.uk*

المشرف الأكاديمي:
Professor Paul Kerswill
paul.kerswill@york.ac.uk

Appendix D

R Codes for Statistical Analysis (The Merger):

```
View(Bs.u_i)

attach(Bs.u_i)

Bs.u_i$Participant <- as.factor(Bs.u_i$Participant)

Bs.u_i$Tribe <- as.factor(Bs.u_i$Tribe)

Bs.u_i$Age <- as.factor(Bs.u_i$Age)

Bs.u_i$Edu <- as.factor(Bs.u_i$Edu)

Bs.u_i$DV <- as.factor(Bs.u_i$DV)

Bs.u_i$socio.integ <- as.factor(Bs.u_i$socio.integ)

Bs.u_i$Word <- as.factor(Bs.u_i$Word)

str(Bs.u_i)

library(lme4)

mod.(Bs'.merger) <- glmer(Dependent.Variable ~ Tribe + Age +
Edu + Contact + (1 | Speaker) + (1 | Item), data = Bs.u_i,
family = binomial,

control=glmerControl(optimizer="bobyqa"))

summary(Bs'.merger)

library(Hmisc)

somers2(probs, as.numeric(Bs.u_i$DV)-1)

library(effects)

plot(allEffects(Bs'.merger), type="response")
```

R Codes for Statistical Analysis (3SM Suffixed Pronoun):

```
View(Bs.ih_ah)

attach(Bs.ih_ah)

Bs.ih_ah$Participant <- as.factor(Bs.ih_ah$Participant)

Bs.ih_ah$Tribe <- as.factor(Bs.ih_ah$Tribe)

Bs.ih_ah$Age <- as.factor(Bs.ih_ah$Age)

Bs.ih_ah$Edu <- as.factor(Bs.ih_ah$Edu)

Bs.ih_ah$DV <- as.factor(Bs.ih_ah$DV)

Bs.ih_ah$socio.integ <- as.factor(Bs.ih_ah$socio.integ)

Bs.ih_ah$Word <- as.factor(Bs.ih_ah$Word)

library(lme4)

mod.(Bs'.3SM) <- glmer(Dependent.Variable ~ Age + Tribe +
Contact + (1 | Speaker) + (1 | Item), data = Bs.ih_ah, family
= "binomial", control=glmerControl(optimizer="bobyqa"))

summary(Bs'.3SM)

library(Hmisc)

somers2(probs, as.numeric(Bs.ih_ah$DV)-1)

plot(Age, fitted.values(Bs'.3SM))

plot(Bs'.3SM)

plot(allEffects(Bs'.3SM), type="response")
```

Chi Square Test for the Interviewer Effect, The Merger (Excel):

- Ajmi Speaker:

	A	B	C	D	E
		B inter	S inter	Total	
1		B inter	S inter	Total	
2	B tokens	25	26	51	0.704225
3	S toekn	3	17	20	0.28169
4	total	28	43	71	
5					
6					
7		19.71831	30.28169		
8		7.887324	12.11268		
9					
10					
11			0.008058		
12					

C11 = CHITEST(B2:C3,B7:C8)

- Dosari Speaker:

	A	B	C	D	E
1		B inter	S inter	Total	
2	B tokens	29	7	36	0.455696
3	S toekn	15	28	43	0.544304
4	total	44	35	79	
5					
6					
7		20.05063	15.94936709		
8		23.94937	19.05063291		
9					
10					
11					
12			4.70269E-05		
13					

C12 = CHITEST(B2:C3,B7:C8)

- Qahtani Speaker:

	A	B	C	D	E
1		B inter	S inter	Total	
2	B tokens	43	23	66	0.804878
3	S toekn	4	12	16	0.195122
4	total	47	35	82	
5					
6		37.82927	28.17073		
7		9.170732	6.829268		
8					
9					
10			0.003578		
11					

C10 = CHITEST(B2:C3,B6:C7)

References List

Abboud, P.F. (1964). *The syntax of Najdi Arabic*. (Ph.D. Thesis), Austin, Texas.

Abboud, P.F. (1978). The vowel of the imperfective prefix in Najdi Arabic. In *Linguistic and Literary Studies in Honour of Archibald A. Hill, ii Descriptive Linguistics*, eds. M. Jazayeri, E. Polome and W. Winter. The Hague, Paris, New York: Mouton, (pp.129-38).

Abboud, P.F. (1979). The verb in Northern Najdi Arabic. In *BSOAS* 42, ii, (pp. 467-499).

Abdei-Jawad, H. (1986). The emergence of an urban dialect in the Jordanian urban centers. *International Journal of the Sociology of Language*, 61, (pp. 53-63).

Abu-Haidar, F. (1991). *Christian Arabic of Baghdad*, Wiesbaden: Otto Harrassowitz.

Abu-Haidar, F. (2006a). Baghdad Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 222-231). Leiden: Brill.

Abu-Haidar, F. (2006b). Bedouinization. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 269-274). Leiden: Brill.

Abu-Hasan, H. (2001). *Lahjat ahil Al-Qasim: Dirasah lughawiyyah*. Riyadh: Imam Muhamad Bin Saud University.

Abu-Manga, A. (2009). Sudan. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 375-381). Leiden: Brill.

Abu-Mansour, M. (2008). Meccan Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp.179-187). Leiden:Brill.

Ad-Darsoni, S. (2009). *Mu'jam lisaan qahtaan*. Jazan University. Jazan, Saudi Arabia.

Ad-Darsoni, S. (2013). *Mu'jam allahajaat almahkiyyah fi almamlakah al'arabiyah assi'udiyah*. Jazan University. Jazan, Saudi Arabia.

Agameya, A. (2009). Subject. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 354-359). Leiden: Brill.

Al-Ahdal, H. (1989). *A sociolinguistic description of speech in Makkah*. (Ph.D. thesis), University of Reading.

Al-Aiyed, M. (2018). *Diglossic Code-Switching between Standard Arabic and Najdi Arabic in Religious Discourse*. (Unpublished Ph.D. Thesis), Durham University.

Al-Azraqi, M. (1998). *Aspects of the syntax of the dialect of Abha (south west Saudi Arabia)*. (unpublished Ph.D. thesis), Durham University.

Al-Azraqi, M. (2007). The use of Kaškašah/Kaskasah and alternative means among educated urban Saudi speakers. In C. Miller, E. Al-Wer, D. Caubet, & J. C. E.Watson (eds.), *Arabic in the city: Issues in dialect contact and language variation*. London: Routledge. (pp. 230-245).

Al-Essa, A. (2008). *Najdi speakers in Hijaz: A sociolinguistic investigation of dialect contact in Jeddah*. (Unpublished Ph.D. thesis): University of Essex, UK.

Al-Ghamdi, N. (2014). *A sociolinguistic study of dialect contact in Arabia: Ghamdi immigrants in Mecca*. (Unpublished Ph.D. Thesis), University of Essex.

Al-Hazmy, A. M. (1975). *A critical and comparative study of the spoken dialect of the Harb tribe in Saudi Arabia*. (unpublished Ph.D. thesis), University of Leeds.

Al-Hugail, A. (1967) *kanz alansab wa majma' aladab*. Riyadh: Al-Hugail.

Al-Jehani, N. (1985). *Sociostylistic stratification of Arabic in Makkah*. (Ph.D. thesis), University of Michigan, Ann Arbor.

Al-Qahtani, K. (2015) *A sociolinguistic study of the Tihami Qahtani dialect in Asir, Southern Arabia*. (unpublished Ph.D. thesis), University of Essex.

Al-Rojaie, Y. (2013). Regional dialect leveling in Najdi Arabic: The case of the deaffrication of [k] in the Qasimi dialect. *Language Variation and Change, 25*(1), (pp. 43-63).

Al-Shamrani, H. (2012). Diglossia in Arabic TV stations. In *Journal of King Saud University–Languages and Translation*. 24, (pp. 57-69).

Al-Sharkawi, M. (2017) *History and Development of the Arabic Language: From pre-Islamic times to the age of conquests*. London: Routledge.

Al-Shehri, A. (1993). *Urbanisation and linguistic variation and change: A sociolinguistic study of the impact of urbanisation on the linguistic behaviour or urbanised rural immigrants in Hijaz, Saudi Arabia*. (Ph.D. thesis), University of Essex, Colchester.

Al-Sowayan, S. (2010). *As-sahraa' al-Arabiah: thaqāfatha wa shi'riha Abr Al-isur, Qira'ah anthrobologiah*. Beirut: Arabian Network Publishing.

Al-Sweel, A. (1987). Verbal and nominal forms of Najdi Arabic. *Anthropological Linguistics,* 29, (pp. 71-90).

Al-Sweel, A. (1990). Some aspects of Najdi Arabic phonology. *ZAL*, 21, (pp. 71-82).

Al-Uthaymin, A. (1999). *The history of kingdom of Saudi Arabia*. Riyadh: Al-Obaikan.

Al-Wer, E. (1999). Why do different variables behave differently? In Y. Suleiman (Ed.), *Language and society in the Middle East and North Africa* (pp. 38-57). London: Routledge.

Al-Wer, E. (2002a). Education as a speaker variable. In A. Rouchdy (Ed.), *Language contact and language conflict in Arabic: variations on a sociolinguistic theme* (pp. 41-53). London: Routledge Curzon.

Al-Wer, E. (2002b). Jordanian and Palestinian dialects in contact: Vowel raising in Amman. In M. Jones, & E. Esch (Eds.), *Language change: The interplay of internal, external and extra-linguistic factors* (pp. 63-79). Berlin: Mouton de Gruyter.

Al-Wer, E. (2007a). Jordanian Arabic (Amman). In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 505-517). Leiden: Brill.

Al-Wer, E. (2007b). The formation of the dialect of Amman. In C. Miller, E. Al-Wer, D. Caubet, & J. Watson (Eds.), *Arabic in the city: Issues in dialect contact and language variation* (pp. 55-76). London: Routledge.

Al-Wer, E. (2013). Sociolinguistics. In J. Owens (Ed.), *The Oxford handbook of Arabic linguistics* (pp. 241-263). Oxford: Oxford University Press.

Appleyard, D. (2008). Personal pronoun (Standard Arabic). In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 588-593). Leiden: Brill.

Ashwan, M. (1990). *The Population growth of Riyadh City in Saudi Arabia*. (Unpublished Ph.D. thesis): University of Durham, UK.

Auer, P., and Hinskens, F. (2005). The role of interpersonal accommodation in a theory of language change. In P. Auer, F. Hinskens, & P. Kerswill (Eds.), *Dialect change: Convergence and divergence in European languages* (pp. 335-357). Cambridge: Cambridge University Press.

Auer, P., Barden, B., and Grosskopf, B. (1998). Subjective and objective parameters determining "salience" in long-term dialect accommodation. *Journal of Sociolinguistics*, 2 (2), (pp. 163-187).

Baayen, H. (2008) *Analyzing Linguistic Data: A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.

Badawi, M. (1973). *Mustawayaat al-arabiyya al-mu'asirah fi Misir*. Cairo: Dar al-Ma'arif.

Bardovi-Harlig, K. (1987). Markedness and salience in second-language acquisition. *Language Learning* 37, (pp. 385-407).

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language,68* (3), (pp. 255-278).

Bassiouney, R. (2009). *Arabic sociolinguistics*. Edinburgh: Edinburgh University Press.

Bates D, Maechler M, Bolker B, Walker S (2015). lme4: Linear Mixed-Effects Models Using Eigen and S4. R package version 1.1-10, URL http://CRAN.R.project.org/package=lme4.

Baugh, J. (2001). A dissection of style-shifting. In P. Eckert & J. R. Rickford (Eds.), *Style and sociolinguistic variation*. Cambridge: Cambridge University Press. (pp. 109-118)

Becker, K. (2013). The sociolinguistic interview. In C. Mallinson, B. Childs, & G. V. Herk (Eds.), *Data collection in Sociolinguistics: Methods and applications*. (pp. 91-100)

Behnstedt, P. (2009). Syria. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 402-409). Leiden: Brill.

Bell, Allan (1977). *The language of radio news in Auckland: a sociolinguistic study of style, audience and subediting variation*. (Unpublished PhD thesis). University of Auckland.

Bell, Allan. (1984) Language style as audience design. *Language in Society*, 13. (pp. 145-204).

Blanc, H. (1960). Style variations in Arabic: A sample of interdialectal conversation. In C. A. Ferguson (ed.), *Contributions to Arabic linguistics*, Cambridge, MA: Harvard University Press, (pp. 81-156).

Blanc, H. (1964). *Communal dialects in Baghdad.* Cambridge: Centre for Middle Eastern Studies.

Blanc, H. (1969). The Fronting of Semitic g and the qāl–gāl Dialect Split in Arabic. In *Proceedings of the International Conference on Semitic Studies held in Jerusalem, 19-23 July 1965*, Jerusalem: Israel Academy of Sciences and Humanities, (pp. 7–37).

Blanc, H. (1970). The Arabic dialect of the Negev Bedouins. *Proceedings of the Israel Academy of Sciences and Humanities* 4:7.112‑150.

Blom, J. P. and Gumperz, J. J. (1972) Social meaning in linguistic structures: Code-switching in Norway. In J. Gumperz and D. Hymes (eds.), *Directions in Sociolinguistics*, New York: Holt, Rinehart & Winston, (pp. 407-34).

Bott, E. (1971). *Family and social network* (2nd ed.). London: Tavistock.

Boucherit, A. (2006). Algiers Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 58-66). Leiden: Brill.

Boucherit, A. and Lentin, J. (1989). Les dialectes féminins dans le monde arabe: des dialectes minoritaires et leur évolution' in Koskas, E. and Leeman, D. (eds.) *Genre et Langage. Actes du Colloque tenu à Paris X-Nanterre*, Linx 21, (pp. 17-37).

Cadora, F. J. (1992). *Bedouin, village and urban Arabic: An ecolinguistic study*. Leiden: E.J. Brill.

Cadora, F.J. (1989). Linguistic Change and the Bedouin-Sedentary Dichotomy in Old Arabic Dialects, *Anthropological Linguistics*, (pp. 264-284).

Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. In *Language Variation and Change*. 21(1), (pp. 135-156).

Caubet, D (1998). Etude sociolinguistique des traits préhilaliens dans un dialecte en voie d'urbanisation' in Aguadé & al. (eds.), (pp. 165-175).

Caubet, D. (2008). Moroccan Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 273-287). Leiden: Brill.

Cedergren, H.J. and Sankoff, D. (1974) Variable rules: Performance as a statistical reflection of competence. *Language* 50(2), (pp. 333-355).

Chakrani, B. (2011). Covert Language Attitudes: A New Outlook on the Sociolinguistic Space of Morocco. *Selected Proceedings of the 40th Annual Conference on African Linguistics*. Eyamba G. Bokamba et al. (Eds.), (pp. 168-177). Somerville, MA: Cascadilla Proceedings Project.

Chambers, J. K. (2009) *Sociolinguistic Theory: Language Variation and its Social Significance*. 3rd ed. Oxford: Wiley-Blackwell.

Cleveland, R. L. (1963). A classification for the Arabic dialects of Jordan. *Bulletin of the American Schools of Oriental Research*, 167, (pp. 56-63).

Cohen, D. (1970). Koinè, langues communes et dialectes arabes. In David Cohen, *Etudes de linguistique semitique et arabe,* (pp. 105-125). The Hague & Paris: Mouton.

Coupland, N. (1984). Accommodation at work: Some phonological data and their implications. *International Journal of the Sociology of Language*, 46, (pp. 49-70).

Daher, J. (1999). Gender in linguistic variation: The variable (q) in Damascus Arabic. *Perspectives on Arabic linguistics, XI*, (pp. 183-206).

De Jong, R. (2009). Sinai Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 237-251). Leiden: Brill.

Dickins, J. (2007). Khartoum Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 559-571). Leiden: Brill.

Dillard, J. (1972). *Black English: Its history and usage in the United States*. New York: Random House.

Doughty, J. (1924). *Travels in Arabia Deserta*. London: Jonathan Cape and the Medici Society.

Douglas-Cowie. Ellen (1978). Linguistic code-switching in a Northern Irish village: Social interaction and social ambition. In Peter Trudgill (ed.), *Sociolinguistic Patterns in British English*. London:Edward Arnold . (pp. 37-51).

Eckert, P. (2000). *Linguistic variation as social practice*. Oxford, Blackwell.

Edzard, L. (2008). Omani Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp.478-491). Leiden:Brill.

El-Hassan, S. A. (1977). Educated Spoken Arabic in Egypt and the Levant: A critical review of diglossia and related concepts". *Archivum Linguisticum* (n.s.) 8:2. (pp. 112-132).

Farghal, M. (2007). Kinship terms. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 583-588). Leiden: Brill.

Ferguson, C. (1959a). The Arabic koine. *Language.* 35, (pp. 616-630).

Ferguson, C. (1959b). Diglossia. *Word, 15*, (pp. 325-40).

Fischer, W & Jastrow, O, (1980) *Handbuch der arabischen dialekte*. Wiesbaden: Harrassowitz.

Fischer, W. (2006). Classical Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 397-405). Leiden: Brill.

Germanos, M. and Miller, C. (2015). Is religious affiliation a key factor of language variation in Arabic speaking countries? *Language & Communication 42*. (http://dx.doi.org/10.1016/j.langcom.2014.12.001)

Gibson, M. (2009). Tunis Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 563-571). Leiden: Brill.

Giles, H and Andrew, B. (2004). Assessing language attitudes. In Alan Davies and Catherine Elder (eds.), *The Handbook of Applied Linguistics*. (pp. 187-209), Oxford: Blackwell.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics, 15*, (pp. 87-105).

Giles, H. (Ed.). (1984). The dynamics of speech accommodation [Special issue]. *International Journal of the Sociology of Language*, 46, (pp. 1-155).

Graff, D., Labov, W., and Harris, W.A. (1986). Testing listeners' reactions to phonological markers of ethnic identity: a new method for sociolinguistic research. In Sankoff, D. (Ed.), *Current Issues Linguist Theory*. John Benjamin, Amsterdam/Philadelphia, (pp. 45-58).

Grand' Henry, J. (2009). Taltala. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 430-431). Leiden: Brill.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. K. L. (1998). Measuring Individual
Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality
and Social Psychology*, 74, (pp. 1464-1480).

Gumperz, J. (1977). The sociolinguistic significance of conversational code-switching. *RELC
Journal*. 8, 2: (pp. 1-34).

Gumperz, J. (1982). Social network and language shift. In J. J. Gumperz (ed.), *Discourse
Strategies*. Cambridge: Cambridge University Press, (pp. 38-58).

Haeri, N. (1991). *Sociolinguistic variation in Cairene Arabic: Palatalization and the qaf in
the speech of men and women*. (Unpublished Ph.D. thesis). The University of
Pennsylvania, Philadelphia.

Haeri, N. (1997). *The sociolinguistic market of Cairo: Gender, class and education*. London:
Kegan Paul International.

Hamza, F. (1933) *The heart of Arabia*, London.

Hary, B. (1996). The importance of the language continuum in Arabic multiglossia.
*Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said
Badawi*, ed. Alaa Elgibali, (pp. 69-90). Cairo: American University of Cairo Press.

Hinskens, F. (1992). *Dialect levelling in Limburg: structural and sociolinguistic aspects*.
(Ph.D. thesis). University of Nijmegen.

Holes, C. (1983), Bahraini dialects: Sectarian differences and the sedentary/nomadic split.
Zeitschrift for arabische Linguistik 10, (pp. 7-38).

Holes, C. (1984). Bahraini dialects: Sectarian differences exemplified through texts.
Zeitschrift für Arabische Linguistik, 13, (pp. 27-67).

Holes, C. (1986), The social motivation for phonological convergence in three Arabic
dialects. *International Journal of the Sociology of Language*. 61, (pp. 33-51).

Holes, C. (2005). Dialect and national identity: The cultural politics of self-representation in Bahraini musalsalāt. In P. Piscatori (Ed.), *Monarchies and nations: Globalization and identity in the Arab States of the Gulf* (pp. 52-72). London: I.B. Tauris.

Holes, C. (2006). Bahraini Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 241-255). Leiden: Brill.

Holes, C. (2007). Kuwaiti Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 608-620). Leiden:Brill.

Holes, C. (2016). *Dialect, Culture and Society in Eastern Arabia. Volume III: Phonology, Morphology, Syntax, Style.* Leiden: Brill.

Hosmer, D. & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: John Wiley & Sons, Inc.

Husàm ad-Dìn, K. (1990). *Al-Qaràba* [Kinship]. Cairo: Anglo-Egyptian Library.

Ibrahim, Z. (2008). Lexical variation: modern standard Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 13-21). Leiden: Brill.

Ingham, B. (1971). Some characteristics of Meccan speech. *Bulletin of the School of Oriental and African Studies*, xxxiv (2): (pp. 273-297).

Ingham, B. (1979). Notes on the dialect of the Muṭair in Eastern Arabia. In *ZAL*, 2, (pp. 23-35).

Ingham, B. (1982a). *North East Arabian dialect*. London and Boston: Kegan Paul International.

Ingham, B. (1982b). Notes on the dialect of the Dhafīr of north-eastern Arabia. In *BSOAS*, 45, ii, (pp. 249-97).

Ingham, B. (1986), Notes on the dialect of the Ā1 Murrah of eastern and southern Arabia. In *BSOAS*, 49, ii, (pp. 271-89).

Ingham, B. (1994). *Najdi Arabic: Central Arabian*. Amsterdam: John Benjamins Publishing Company.

Ingham, B. (2009). Saudi Arabia. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 123-130). Leiden: Brill.

Isaksson, Bo. (1991). The personal markers in the modern Arabic dialects of the Arabian Peninsula. *Orientalia Suecana*, 40, (pp. 117-145).

Iványi, T. (2006). Diphthongs. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 640-643). Leiden: Brill.

Jabeur, M. (1987). *A Sociolinguistic study in Tunisia: Rades.* (Unpublished Ph.D. thesis). The University of Reading, Reading.

Jabri, A. (2007). *Lahjat al-arab fi al-qur'an al-kareen: dirasah istiqra'iyah tahleeliah*. Beirut: Daar Alkutub Al-ilmiah.

Jagger, H. J. (1940). *English in the Future*. London: Thomas Nelson.

Jastrow, O. (2007). Iraq. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 414-424). Leiden: Brill.

Johnson, D. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass, 3*(1), (pp. 359-383).

Johnstone, B. and Scott K. (2008). Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics* 12(1), (pp. 5-33).

Johnstone, T. M. (1964). Further studies on the Dōsiri dialect of Arabic as spoken in Kuwait. *Bulletin of the School of Oriental and African Studies*, 27 (1), (pp. 77-113).

Johnstone, T. M. (1967a). *Eastern Arabian dialect studies*. London: Oxford University Press.

Johnstone, T. M. (1967b). Aspects of syllabication in the spoken Arabic of 'Anaiza. In *BSOAS*, 40, (pp. 1-16).

Kerswill, P. (1985). A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge Papers in Phonetics and Experimental Linguistics* 4, Dept. of Linguistics, University of Cambridge.

Kerswill, P. (2013). Koineization. in JK Chambers & N Schilling (eds), *Handbook of Language Variation and Change*. 2nd ed., Wiley-Blackwell, Oxford, (pp. 519-536).

Kerswill, P. and Williams, A. (2002). 'Salience' as an explanatory factor in language change: Evidence from dialect levelling in urban England. In M. C. Jones & E. Esch (Eds.), *Language change: The interplay of internal, external and extra-linguistic factors* (pp. 81-110). Berlin: Walter de Gruyter.

Kerswill, P. (2010). Contact and new varieties. In R. Hickey (Ed.), *The handbook of language contact* (pp. 230-251). Oxford: Wiley-Blackwell.

Kressel, G. (1996). *Ascendancy Through Aggression: The Anatomy of a Blood Feud Among Urbanized Bedouins*. Wiesbaden: Harrassowitz.

Labov, W. (1963) The social motivation of a sound change. *Word*, 19, (pp. 273-309).

Labov, W. (1966). *The social stratification of English in New York City.* Washington, D.C.: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.

Labov, W. (1984). Field methods of the Project on Linguistic Change and Variation. In J. Baugh, & J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28-53). Englewood Cliffs: Prentice Hall.

Labov, W. (1986) Language structure and social structure. In S. Lindenberg, J. Coleman, and S. Nowak (eds.), *Approaches to Social Theory*. New York: Russell Sage.

Labov, W. (1994). *Principles of linguistic change: Internal factors.* Oxford: Blackwell.

Labov, W. (2001). The anatomy of style-shifting. In Penelope Eckert and John R. Rickford (eds.) *Style and Sociolinguistic Variation*. Cambridge: Cambridge University Press, (pp. 85-108).

Labov, William, P. Cohen, C. Robins and J. Lewis. (1968). A study of the non-standard English of Negro and Puerto Rican Speakers in New York City. Cooperative Research Report 3288. Vols I and II. Philadelphia: U.S. Regional Survey (Linguistics Laboratory, U. of Pa.).

Larcher, P. (2009). Verb. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 638-645). Leiden: Brill.

Lentin, J. (2006). Damascus Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 546-555). Leiden: Brill.

Lipinski, E. (1997). *Semitic languages: outline of comparative grammar*. Leuven: Peters Publishers & Department of Oriental Studies.

Llamas, C., Watt, D., and Johnson, D.E. (2009). Linguistic accommodation and the Salience of national identity markers in a border town. *J. Lang. Soc. Psychol*. 28, (pp. 381-407). doi:10.1177/0261927X09341962.

Llamas, C., Watt, D., and MacFarlane, A.E. (2016) Estimating the Relative Sociolinguistic Salience of Segmental Variables in a Dialect Boundary Zone. *Front. Psychol*. 7:1163. doi: 10.3389/fpsyg.2016.01163.

MacLeod, B. (2015). A critical evaluation of two approaches to defining perceptual salience. *Ampersand*, 2, (pp. 83-92).

Mahmoud, Y. (1986). Arabic after diglossia. *The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday. I. From phonology to society*, ed. Joshua Fishman a.o. (pp. 239-251). Berlin: Mouton de Gruyter.

Mair, C. (2013). Speculating on the future of English as a contact language. In Schreier, D. & Hundt, M. (eds.), *English as a Contact Language*. Cambridge: Cambridge University Press, (pp. 314–28).

Martinet, A. (1952). Function, structure, and sound change. *Word*, 8, (pp. 1-32).

Meiseles, G. (1980). Educated spoken Arabic and the Arabic language continuum. *Archivum Linguisticum*, 11(2): (pp. 118-48).

Messaoudi, L. (1998). Traits linguistiques du parler ancien de Rabat' in Aguadé & al. (eds.), (pp. 157-163).

Messaoudi, L. (2001). Urbanisation linguistique et dynamique langagières dans la ville de Rabat. in Bulot, T., Bauvois, C. and Blanchet, P. (eds.) *Sociolinguistique urbaine. Variations linguistiques, images urbaines et sociales*. Rennes: Presses de l'Université de Rennes.

Miller, C. (2004). Variation and changes in Arabic urban vernaculars. In M. Haak, K. Versteegh, & R. De Jong (Eds.), *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday* (pp. 177-206). Leiden: Brill.

Miller, C. (2005). Between accommodation and resistance: Upper Egyptians migrants in Cairo. In *Linguistics* 43: 5, (pp. 903–956).

Miller, C. (2006). Dialect koine. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 593-597). Leiden: Brill.

Miller, C. (2007). Arabic urban vernaculars: Development and change. In C. Miller, E. AI-Wer, D. Caubet & J. Watson (Eds.), *Arabic in the city: Issues in dialect contact and language variation*, (pp. 1-31). London, UK and New York, NY: Routledge.

Milroy, L (1980). *Language and Social Networks*. Oxford: Basil Blackwell.

Milroy, L. and Gordon, M. (2003) *Sociolinguistics: Method and Interpretation*. London: Blackwell.

Milroy, L. and Milroy, J. (1992) Social network and social class: Towards an integrated sociolinguistic model. *Language in Society*, 21, (pp. 1-26).

Moore, J. (1990). Doubled verbs in modern standard Arabic. in Eid and McCarthy (Eds.), *Perspectives on Arabic Linguistics II* (pp. 55-93). Amsterdam: John Benjamins.

Mufwene, S. (1991). Pidgins, creoles, typology, and markedness. In Francis Byrne & Thom Huebner (eds.) *Development and structures of creole languages. Essays in honor of Derek Bickerton*. Amsterdam: John Benjamin, (pp.123-143).

Murad, M. (2007). *Language attitudes of Iraqi native speakers of Arabic: A sociolinguistic investigation.* (Unpublished master's thesis), University of Kansas, Lawrence.

Naïm, S. (2006). Beirut Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 274-286). Leiden: Brill.

Nassar, H. (1956). Al-Mujam al-Arabi Nash'atuhu wa-tatawwuruhu. Cairo: Dar al-Kitab al-Arabi.

Nazzal, A. (2001). *The pragmatic functions of Qur'anic verses: The case of insha'allah in Arabic discourse as a species of indirectness.* (Unpublished Ph.D. Thesis), State University of New York, Albany.

Niedzielski, N. and Giles, H. (1996). Linguistic accommodation. In H. Goebl, P. H. Nelde, Z. Stary, & W. Wölck (Eds.), *Kontaktlinguistik: Ein internationales Handbuch zeitgenössischer Forschung [An International Handbook of Contemporary Research]* (Vol. 1, pp. 332-342). Berlin: Mouton de Gruyter.

Nolan, F., Kerswill, P. and Wright, S. (1991). The interaction of sociophonetic features and connected speech processes. Final report submitted to the Economic and Social Research Council, ref. R000231056.

Oppenheim, M., Bräunlich, E. and Caskel, W. (2004). *Die Beduinen*. London: Al-Warrak Pub. Oxford: Blackwell.

Owens, J. (1998). *Neighborhood and Ancestry: Variation in the Spoken Arabic of Maiduguri, Nigeria*. Philadelphia, PA: Benjamins.

Owens, J. (2001). Arabic sociolinguistics. *Arabica, 48*(4), (pp. 419-469).

Owens, J. (2006). *A Linguistic History of Arabic*. Oxford: Oxford University Press.

Palva, H. (1994) Bedouin and Sedentary elements in the dialect of Es-Salt: diachronic notes on the sociolinguistic development, in D. Caubet and M. Vanhove (eds) *Actes des Premières journées internationales de dialectologie arabe de Paris*, Paris: INALCO, (pp. 459-469).

Palva, H. (2006). Dialects: Classification. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 604-613). Leiden: Brill.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. (2019). *PsychoPy2: experiments in behavior made easy. Behavior Research Methods*. 10.3758/s13428-018-01193-y.

Pereira, C. (2009). Tripoli Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 548-556). Leiden: Brill.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus.* New York: Springer.

Prochazka, T. (1988). *Saudi Arabian dialects*. London: Kegan Paul.

Quirk, R. (1972). English in twenty years. In Quirk, R. (ed.), *The English Language and Images of Matter.* London: Oxford University Press, (pp. 68-76).

R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org.

Rabin, C. (1951). *Ancient West Arabian*. London: Taylor's Foreign Press.

Retsö, J. (2006). ʿArab. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 126-132). Leiden: Brill.

Rickford, J. R. and McNair-Knox, F. (1994) Addressee- and topic-influenced style-shift: A quantitative sociolinguistic study. In D. Biber and E. Finegan (eds.), *Sociolinguistic Perspectives on Register*. Oxford: Oxford University Press, (pp. 235-76).

Romaine, S. (1978). Post-vocalic /r/ in Scottish English: Sound change in progress. In P. Trudgill (ed.), *Sociolinguistic Patterns in British English*. London: Arnold, (pp. 144-57).

Rosenhouse, J. (2006). Bedouin Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 259-269). Leiden: Brill.

Sadiq, S. (2016). *Dialect convergence in Egypt: The impact of Cairo Arabic on Minya Arabic.* (Unpublished Ph.D. thesis). The University of York, UK.

Saeed, A. (1997). *The pragmatics of codeswitching from fusha Arabic to ammiyyah Arabic in religious oriented discourse*. (Unpublished Ph.D. thesis), Ball State University, Indiana.

Samarin, W. J. (1971). Salient and substantive pidginization. In D. Hymes (ed.), *Pidginization and creolization of languages*. Cambridge University Press.

Sankoff, D., Tagliamonte, S, and Smith, E. (2005). *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto.

Schilling-Estes, N. (1998) Self-conscious speech in Ocracoke English. *Language in Society*, 27 (1), (pp. 53-83).

Schirmunski, V. (1930). Sprachgeschichte und Siedelungsmundarten. *Germanisch Romanische Monatschrift* 18: (pp.113-22) (Part I), (pp.171-188) (Part II).

Schmidt, R. (1974). *Sociolinguistic variation in spoken Arabic in Egypt: Re-examination of the concept of diglossia*. (Ph.D. thesis), Brown University, Providence.

Shahin, K. (2008). Palestinian Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, pp. 526-538). Leiden: Brill.

Siegel, J. (1985) Koines and koineization. *Language in Society*. 14: (pp. 357-378).

Siegel, J. (2001) Koine formation and creole genesis. In N. Smith and T. Veenstra (eds.), *Contact and Creolization*. (pp. 175-197). Amsterdam: John Benjamins.

Siegel, J. (2010). *Second Dialect Acquisition*. Cambridge University Press, Cambridge, New York.

Smith, W. Robertson. (1990). *Kinship and marriage in early Arabia*. Stanley A. Cook. London: Darf Publishers.

Tagliamonte, S. A. (2011). *Variationist sociolinguistics: Change, observation, interpretation.* Malden: John Wiley & Sons.

Taine-Cheikh, C. (2007). Hassàniyya Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 2, pp. 240-250). Leiden: Brill.

The General Authority of Statistics, Saudi Arabia. (2010). *Detailed results of Riyadh Province (general population and housing census).*

Torbert, B.C. (2004). *Southern Vowels and the Social Construction of Salience*. (unpublished Ph.D. thesis). Duke University.

Trager, G. L., & Rice, F. A. (1954). The personal-pronoun system of Classical Arabic. *Language*, 30 (2), (pp. 224-229).

Trudgill, P & Foxcroft, N (1978). On the sociolinguistics of vocalic mergers: Transfer and approximation in East Anglia. In P. Trudgill (ed.), *Sociolinguistic Patterns in British English, London: Edwin Arnold*, (pp. 69-79).

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, P. (1981). Linguistic accommodation: Sociolinguistic observations on a sociopsychological theory. In Masek, C. S., Hendrick,R. A.. & Miller, M. F. (eds.). *Papers from the parasession on language and behavior*. Chicago, III.: Chicago Linguistic Society. (pp. 218-37).

Trudgill, P. (1986) *Dialects in Contact*. Oxford: Blackwell.

Trudgill, P. (1998). The chaos before the order: New Zealand English and the second stage of new-dialect formation. In E.H. Jahr (ed.), *Advances in Historical Sociolinguistics*. (pp. 197-207) Berlin: Mouton de Gruyter.

Trudgill, P. (2002). Linguistic and social typology. In J. K. Chambers, P. Trudgill and N. Schilling -Estes (eds.). *The Handbook of language variation and change*. Malden, Mass: Blackwell. (pp. 708-728).

Trudgill, P. (2004). *New Dialect Formation: The Inevitability of Colonial Englishes.* Edinburgh: Edinburgh University Press.

United Nations Population Division, (2018). *Urban population, development and the environment.* Department of Economic and Social Affairs, United Nations Secretariat, New York. USA.

Versteegh, K. (1984). *Pidginization and Creolization: the case of Arabic*. Amsterdam-Philadelphia: John Benjamins.

Versteegh, K. (2001). *The Arabic language*. Edinburgh, UK: Edinburgh University Press.

Versteegh, K. (2017). Religion as a Linguistic Variable in Christian Greek, Latin, and Arabic. in *Philologists in the World. A Festschrift in Honour of Gunvor Mejdell*. Edited by N. S. Eggen and R. Issa. Oslo: Novus Press. (pp. 57-88)

Vicente, Á. (2008). Personal pronoun (Arabic dialects). In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 3, (pp. 584-588). Leiden: Brill.

Walters, K. (1994). Diglossia, linguistic variation, and language change in Arabic. In M. Eid (Ed.), *Perspectives on Arabic Linguistics*, VIII (pp. 157-197). John Benjamins Publishing Company. Amsterdam/ Philadelphia.

Watson, J. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Watson, J. (2009). San'ani Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 4, pp. 106-115). Leiden: Brill.

Weinreich, U. (1953). *Languages in Contact*. The Hague: Mouton.

Wilmsen, D., & Woidich, M. (2007). Egypt. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, & Z. Andrzej (Eds.), *Encyclopedia of Arabic Language and Linguistics* (Vol. 2, pp. 1-12). Leiden: Brill.

Woidich, M. (2006). Cairo Arabic. In K. Versteegh, M. Woidich, M. Eid, A. Elgibali, &Z. Andrzej (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. 1, pp. 323-333). Leiden: Brill.

Wolfram, W. (1969). *A Sociolinguistic description of Detroit Negro speech.* Washington, D.C: Centre for Applied Linguistics.

Wolfram, W. and Schilling-Estes, N. (1996) Dialect change and maintenance in a post-insular community. In E. Schneider (ed.), *Focus on the USA*. Amsterdam: Benjamins, (pp. 103-48).