Phonological outcomes of language contact
in the Palestinian Arabic dialect of Jaffa

Uri Horesh

A thesis submitted for the degree of Doctor of Philosophy
Department of Language and Linguistics
University of Essex
February 2014
## Contents

Abstract 3  
Acknowledgments 4  

Chapter 1: Background 6  
1.1 Diglossia and bilingualism 6  
1.2 An overview of the speech community 10  
1.2.1 Palestinian Arabic within the taxonomy of Arabic dialects 10  
1.2.2 Order of acquisition of languages 14  
1.2.3 The Status of Palestinian Arabic in Israel 15  
1.2.4 Terminological problems 16  
1.2.5 The site of the main study, Jaffa 17  

Chapter 2: The Envelope of Variation 23  
2.1 Overview 23  
2.1.1 Changes in Modern Hebrew Resulting in Divergence from Arabic 24  
2.1.2 Lenition in Palestinian Arabic 26  
2.2 Previous treatments 28  
2.2.1 General 28  
2.2.2 Previous work on the variables at hand 31  

Chapter 3: Methodology 35  
3.1 Preliminaries 35  
3.2 The Fieldwork 35  
3.3 The Sample 37  
3.3.1 The main sample 37  
3.3.2 The control group 40  
3.4 The interviews 41  
3.5 Treatment of the data 42  
Appendix A: Factor groups for multivariate analysis 44  

Chapter 4: Analysis and discussion: variable 1 – (ʕ) 51  
4.1 Description of the variable 51  
4.1.1 A complex variable 51  
4.1.2 Two types of vocalization 51  
4.1.3 Creaky voice 52  
4.1.4 A hierarchy of lenition 53  
4.2 Quantitative analysis 55  
4.2.1 Four facets of analysis 55  
4.2.2 Cross-tabulations of the data per speaker 56  
4.2.2.1 “Eyeballing the data” 56  

Abstract

This is a thesis in variationist sociolinguistics. It attempts to make a contribution to the study of a dialect of Arabic—Palestinian Arabic—spoken in a region where the population is gradually becoming engulfed in a language, which was once quite similar to Arabic, namely Hebrew, but has undergone drastic changes, particularly in its phonological structure, as a result of contact with European languages.

Now, Modern Hebrew is acting as a colonizing language vis-à-vis Palestinian Arabic, and in this study we are exploring the effects the contact between the two languages on the phonology of Arabic in the town of Jaffa, where Arabic-speaking Palestinians and Hebrew-speaking Israeli Jews reside, perhaps not in harmony, but nonetheless in the same urban space.

Employing quantitative methods for one linguistic variable and a sociohistorical analysis for another, we make the case that the two variables observed in this study are but a fragment of the entire complex. Examples from the data collected are provided and briefly analyzed, some of which are from other domains of the language, and these will be further explored at a later date.
Acknowledgments

I have resolved to keep this portion of my thesis brief.

In addition to the specific persons who have offered me professional advice and whom I have thanked in footnotes throughout this study, there are a handful of people I wish to thank for helping me through the rough years that I have gone through between the commencement of my work on this thesis as a graduate student at the University of Pennsylvania and the present day, as I end my studentdom at the University of Essex.

At the risk of my words sounding like a cliché, I cannot find adequate phrases in any of the three languages in which I am fluent to express enough gratitude, love and respect to my friend, mentor and—in the last year and a half—doctoral supervisor, Dr Enam Al-Wer. Enam and I have known one another for over 15 years now, and when I was just about to abandon it all, she would not let me and made every effort to enable me to complete my degree under her supervision. I may have not always appeared grateful, but I will always be humbled by her generosity.

Two dear friends in New York have been rooting for me since the beginning of time: Maryam Bakht and Reem Khamis-Dakwar. And while at Essex I made a new friend, colleague and like-minded ally on virtually all political issues. Bill Cotter is now, like me, a Chicagoan, though neither of us knows how long that will last.
Orit Bashkin has been a great friend and academic colleague since our days as undergraduates in Sheikh Muwannis in Palestine, via a year we spent together in Philadelphia, and through the present, which has coincidentally brought me to the same Chicago, where she has been teaching and researching for a decade.

Also in Philly, I have benefited greatly from the professional and compassionate services of Dr. Jeff Eiberson. Much of the intellectual foundation for this thesis is also owed to my days in the City of Brotherly Love, to the many friends, enemies and frenemies I have acquired there, and to my professors who taught me about language variation to begin with, particularly the late Ellen Prince, and—may then long live—Bill Labov and Gillian Sankoff of the University of Pennsylvania. I have since moved a quarter of the way around the globe and back (but into Northern Cities Shift territory). Philadelphia, however, is still my pseudo-home in America.

Finally, I would like to dedicate this work to my parents, Varda and Ruvik, who have been utterly patient and as supportive as parents can be.

And to the people of Palestine.
Chapter 1: Background

1.1 Diglossia and bilingualism

The original title of this study, when I first conceived of it, included two concepts that are as interesting as they are messy. Diglossia and bilingualism have each been at the core of many dissertations, books, articles and water cooler discussions. I have reservations as to the appropriateness of each of them to the situation prevalent in the speech community that I am describing in this study, and have therefore decided to purge them, at least from the very forefront of the thesis, its title. Furthermore, I believe that the former, at least, has been overused and over-emphasized in the particular context of the Arabic-speaking world. Some of my criticism – although not all – will be made explicit in what follows. Some will have to wait for a separate, theoretical discussion to be conducted elsewhere. What is, however, in the center of my thoughts on diglossia and Arabic, is that variationist theory can shed light on it and contextualize it as merely a case—surely a complex, interesting case—of language variation that is manifest in intra-speaker and inter-speaker domains, as well as in intra- and inter-community domains. This being said, I will be using both diglossia and bilingualism in this dissertation as mere tools, or “codenames” for the specific clusters of sociolinguistic phenomena that I have observed during my years studying Arabic and in particular during my year or so of fieldwork in Palestine\(^1\). I will cite several researchers, starting with Charles Ferguson in the very next section. As is customary in academic writing, my

\(^1\) The term “Palestine” refers to all of historic Palestine, between the River Jordan and the Mediterranean Sea.
citations will be for crediting such authors with positing a particular notion, in some case with my (full or partial) endorsement, in some cases in order to refute or argue against some or all of their ideas, and in many a case simply to create a common terminology on which to base the rest of my arguments. I shall try to clarify in each instance which of the above is the case, though I suspect that a good amount of ambiguity will still prevail.

Since Ferguson (1959), linguists and scholars in Middle Eastern studies in various other disciplines have been aware of the concept of diglossia, whereby two distinctly different varieties of a language are used to fulfill different communicative functions. The speakers of Palestinian Arabic are no exception. Their use of Modern Standard Arabic (MSA) and their native Palestinian Arabic vernacular is governed by social conventions parallel to those that dictate this alternation in other Arabic-speaking\(^2\) communities. Nevertheless, in the parts of Palestine occupied by Zionist forces in 1948 (known in some circles as “Israel”) Modern Hebrew (Sometimes referred to as “Israeli Hebrew” or ”Modern Israeli Hebrew”), has become the language of the majority. As a result, members of the various Arabic speech communities in that region are for the most part bilingual Arabic-Hebrew speakers, adding to the two main varieties of Arabic that they command at least one variety of an additional language, which fulfills certain communicative functions as well.

My own impression from over three decades of involvement with Arabic regarding a number of historically quintessential features of Arabic phonology

\(^2\) In this thesis, “Arabic” will serve as a blanket term, when a distinction between varieties of Arabic is neither clear nor useful. In all other cases, specific varieties will be mentioned explicitly.
giving way to less marked features has recently been corroborated by Rosenhouse (1991, 2002), though with neither a quantitative analysis of the variability nor a detailed account of the impact of language contact on the extent of the change. The two variables I investigate in this work both pertain to pharyngeality, one as primary place of articulation and the other as secondary place of articulation of coronal consonants. In both cases, the outcomes resemble those of similar processes that Hebrew underwent since its so-called revival in the late 19th century (See, e.g., Reshef 2014).

In what follows I will give a brief overview of the Jaffa speech community as an extreme case of language contact of the type experienced by many other speech communities in Palestine where one variety or another of Palestinian Arabic is the primary language of its members. A preliminary synchronic description of the features under investigation in Palestinian Arabic will be provided, alongside an analogical synchronic account of parallel features in Hebrew. In order to establish that Palestinian Arabic may be undergoing a change in progress with respect to these features, some diachrony will be covered as well for both Hebrew and Arabic.

Following the demographic, sociopolitical and linguistic background of the speech community, I posit a hypothesis regarding the role of Hebrew as a change-inducing contact language for Palestinian Arabic and lay out the methodology with which I examine the hypothesis. For this I rely on previous work (e.g., Nagy 1996) that has dealt with contact-induced change from a variationist perspective. Thanks to Nagy’s work, we have a solid theoretical basis and useful quantitative methods on which to base a study of situations similar to
that of the Francoprovençal dialect of Faetar, which she examined in her dissertation.

Thomason & Kaufman (1988:67) argue that “long-term contact with widespread bilingualism among borrowing-language speakers is a prerequisite for extensive structural borrowing.” Nagy (1996) provides us with a body of work, which has taken this notion and incorporated it within the theoretical thinking and methodology of variationist sociolinguistics. Nagy lists (1996:41) three groups of intensity of contact factors: amount of contact, cultural identity and linguistic factors. Following a survey of contact situations in various speech communities and relying more heavily on two specific studies (Nagy, Moisset & Sankoff 1994 on Anglophone Montreal French and Nagy 1996 on Faetar), Nagy (1996:48) points out that “universal metrics” for intensity of contact may not be possible to establish.

At this juncture, I am not prepared to propose a universal approach myself, but like Nagy, I wish to build upon previous studies of variation and change in a contact environment and adapt a model that will be community-specific, yet comparable to these previous studies and further adaptable for future studies on other languages and communities. Some of the factor groups that are a part of the multivariate analysis in this thesis are a reflection of both the cross-community and community-specific aspects of this issue.

In the case of the Palestinian Arabic-Modern Hebrew interaction observed in central Palestine, the current research is intended as a contribution to the sociolinguistic study of language contact in a situation where the second language within the bilingual component (in this case, Hebrew), is in effect an
“L2.5” of sorts. This is because Hebrew is in fact learned not immediately following the acquisition of the native language, but after (or in conjunction with) the acquisition of a distinct variety thereof. What constitutes L1 in this case is a local variety of Palestinian Arabic (e.g., Jaffa Arabic), and the distinct variety learned prior to the introduction of Hebrew is MSA, which may be viewed as “L1.5.”

1.2 An overview of the speech community

1.2.1 Palestinian Arabic within the taxonomy of Arabic dialects

The term "Palestinian Arabic" is widely used and in fact is suitable for various purposes of classifying the varieties of Arabic spoken by Palestinians, as long as one acknowledges that in virtually no case can a “national” label correspond precisely to a variety of Arabic. This is probably true of many other languages as well, as other factors play crucial roles in shaping the structure of a dialect, its vocabulary and the social value carried by using it in everyday life.

In Arabic specifically, it is common to classify speakers by their “ecolinguistic” background3. Cadora devotes a book-length study to the “three different ecological structures” (Cadora 1992:1) which have characterized the Arab World for over a thousand years. He stresses that there is a correlation between these structures and the linguistic systems used by their inhabitants. The three groups are:

3 The term “ecolinguistic” probably isn't as commonly used as the taxonomy itself. Cadora lays out the taxonomy quite neatly, and I therefore adopt his terminology for the sake of convenience.
Cadora also accounts for “[t]he development of one structure into another [...] attributed to contactual phenomena which are often facilitated by migrations”

As a result, he identifies such transitional stages as:

(4) Bedouin-rural

(5) rural-urban

Subsequently, Cadora introduces the linguistic varieties corresponding to the ecological structures adding the –ite suffix to the latter, coining names for the former. The list of varieties thus includes the following:

(6) Bedouinite → Bedouinite-Ruralite → Ruralite → Ruralite-Urbanite → Urbanite

(Cadora 1992:2)

We find a combination of this classification in the definitions of the dialects of Palestine. Thus, Kimary Shahin has authored the “Palestinian Arabic” entry for The Encyclopedia of Arabic Language and Linguistics (Shahin 2008), and also a grammar of a specific village dialect, the title of which is Rural Palestinian Arabic (Shahin 1995a; see below).

Another traditional taxonomy of Arabic dialects has to do with religious affiliation. It was probably Blanc's (1964) work on the “communal” dialects of
Baghdad that pioneered the notion that Muslims, Christians and Jews can have distinct dialects even though they not only live in the same country and share the same settlement patterns (in this case, urban), but actually dwell in the same city, interact with one another and otherwise could be rendered members of the same speech community. One may remark here that Blanc's descriptive work on Baghdad and Cadora's work on (mostly) Palestinian Arabic are in fact consistent with Eckert and McConnell-Ginet's (1992) contribution to sociolinguistic theory with regard to the role of “community of practice,” though it lacks the universal applicability of the Community of Practice theoretical approach.

The problem with Palestinian Arabic is that it includes all of the above, and an additional factor. While Palestinians are a group worthy of a uniform terminology to describe their national identity, their macro-political convictions, their cultural and familial backgrounds, “Palestinian” is not quite parallel to, e.g., “Egyptian” or “Iraqi” or “Saudi”. Palestinians, many of whom are speakers of what is typically known as Palestinian Arabic, do not have their own state and are most prominently residents of one of three places: the West Bank, the Gaza Strip and the so-called State of Israel. Of course, there are Palestinian refugees around the Arab World (e.g., in neighboring Lebanon and Syria, in Kuwait and Saudi Arabia, in Europe, North America and Australia) and a large Palestinian population in Jordan who are full-fledged citizens (some claim that they comprise 50% or more of Jordan’s entire population, or at least that of its capital, Amman; see Al-Wer 2003:60), but I will only focus here on those Palestinians who live in historical Palestine.
The variety of Arabic spoken in Gaza is an interesting and under-researched one. Like other varieties of Palestinian Arabic, it exhibits features of Levantine Arabic, shared by many dialects east of the Mediterranean Sea and west of the Arabian Peninsula, but also some features of Egyptian Arabic and some pan-Eastern Arabic Bedouin(ite) features. Gaza Arabic is therefore probably entitled to a class of its own in the taxonomy of Palestinian dialects. It is also worthy of research of the sort presented here, which incorporates contact factors into the analysis. Cotter (2013) has taken serious strides in this direction.

However, the dialects of the West Bank and of the Palestinian towns and villages in “1948 Palestine” (the territories claimed by Zionists to comprise the State of Israel) are very similar. And on many structural (mostly phonological and morphological) matters, the political boundary between the territories occupied in 1948 and the West Bank, (also occupied, but in 1967) have no bearing on the placing of isoglosses. A crucial difference does exist, however. Palestinians who live in the 1948 region are Israeli citizens and conduct much of their everyday life in a language other than Arabic, namely Hebrew. Palestinians in neighboring West Bank towns and villages have much less contact with Hebrew. And while some West Bank Palestinians can manage small talk in Hebrew from the days in which they used to commute to the 1948 region (predominantly prior to the 1987 intifada, ‘uprising’) for employment (mostly in construction and agriculture), and some have served in Israeli prisons as political prisoners and took the time to learn Hebrew, in many cases to a high degree of proficiency⁴.

---

⁴ One subject in my West Bank control group, who professed to being fluent in Hebrew, told me – off the record – that he had served time in an Israeli prison and therefore could speak the language. Since none of my subjects are identifiable, and since this portion of the interview was
many others are completely monolingual (insofar as the “regular” Arabic
diglossia contexts, e.g., in Egypt or Yemen) qualify as monolingual).

1.2.2 Order of acquisition of languages

Palestinian citizens of Israel attend, for the most part, public schools whose main
language of instruction is Arabic. Officially, the variety of Arabic used in the
school system is MSA. Pupils study Hebrew starting in the third grade of
elementary school (Amara 2001:160)\(^5\). However, scholars in educational
linguistics raise serious doubt regarding the degree of proficiency these students
typically achieve by the end of high school. In fact, studies by Amara & Mar’i
(2002 and 1999, also in Amara 2001) argue that language teaching in general in
Palestinian schools in the Israeli system is lagging in comparison with parallel
practices in Jewish schools in the country (where Hebrew is the main language
of instruction). A host of factors are cited, such as the general degradation of
anything Arab: the people, the language, the culture, in Israel\(^6\); the monopoly of
the Israeli Ministry of Education on devising curricula (often with no clear goals
or principles); but also the uniqueness of Arabic diglossia in a context where
MSA has even more limited uses than in countries where Arabs are the majority.

\(^5\) I have been informed \(\text{\textit{inter alia by}}\) Elana Shohamy of Tel Aviv University School of Education,
p.c., Feb. 2007) that this is changing and that due to pressure from parents and other
community factors, Hebrew is now entering the Arab schools as early as the second and first
grades in many instances.

\(^6\) It is important to clarify that when the terms “Israel” or “Israeli” are used in this thesis, it is not
meant to indicate the author’s recognition of Israel’s legitimacy, or indeed the legitimacy of the
Zionist occupation of Palestine. I have tried to minimize the use of these terms to refer to the
state institutions and people who identify as Israeli. The land itself is consistently referred to
throughout this thesis as “Palestine.”
After all, in Israel, the language of administration, politics, higher education, etc. is Hebrew\(^7\). In addition, Palestinian schoolchildren in the Israeli system, like their Jewish counterparts, must study English from fourth grade on, which for them is yet a third language; fourth if we consider that studying MSA, at least in the initial stages of primary education is (almost?) equivalent to learning a new language from scratch.

1.2.3 The Status of Palestinian Arabic in Israel

We therefore see a variety of conflicting facts about the intensity of contact between Arabic and Hebrew, and these facts intertwine with the already complex state of affairs regarding Arabic dialects in general and the local varieties in Palestine as a subset of the Arabic-speaking world. Yet I maintain that in spite of the complexity of the situation, a fundamental difference obtains between the varieties of Palestinian Arabic spoken in the West Bank and those spoken in the region occupied in 1948. The former are more like the dialects in other Arab countries, while the latter is a minority language (and see Talmon 2000 for an account of some of the basic facts in this regard), similar in a sense to the varieties of Arabic spoken in Afghanistan, Chad, Eritrea, Ethiopia, Nigeria, Somalia, Turkey, Uzbekistan, etc.\(^8\) In some countries minority Arabic is one of

\(^7\) Sadly, this sentiment, that “in Israel,” Hebrew is the main language, and Arabic must take a proverbial back seat, was also echoed by some of the subjects I have interviewed in Jaffa, including one man who claimed to be a native speaker of Hebrew, while he was clearly a native speaker of Palestinian Arabic.

\(^8\) Let us stress that these are all indigenous Arabic speaking communities. There are also “large immigrant communities in the USA, Latin America and Western Europe” (Spolsky & Shohamy 1999:116), not to mention Arabic-based pidgins and creoles, e.g. Juba Arabic in Sudan and Ki-Nubi in Kenya and Uganda, as well as the fascinating case of Maltese, the only vernacular variety of Arabic which has risen to the status of an official language with a uniform
two or more official languages (only in Chad, Israel and Somalia, according to Spolsky & Shohamy 1999:116).

Spolsky & Shohamy (1999:117) compare Arabic’s secondary role in Israel to “that of Swedish in Finland or of French in Canada [at large]” and not “like French in Québec or in Belgium”. It is my impressionistic view that both non-Swedish Finns and Anglophone non-Québécois Canadians are more tolerant toward their respective minority language (and ethnicity) groups than are most Jewish Israelis toward their Palestinian “cousins”.

1.2.4 Terminological problems

Given the analysis above, I had initially sought to coin a blanket term for the dialects of Palestinian Arabic spoken in 1948 Palestine, which are spoken by an ethnic minority, constantly exposed to, albeit not at all immersed in a community of speakers of Hebrew (and in fact many other immigrant languages, but only Hebrew, and to a much lesser extent Russian, as active participants). There is no universally accepted term for the population. Jews in Israel tend to call them ʔarab id-daːxiːl ‘the Arabs of Israel’ or ʔaravim ʔisraelim ‘Israeli Arabs’. In the Arab world they are either ʕarab tmaːnye w–ʕarbaːʔiːn ‘the Arabs of [19]48’ (the year the State of Israel declared its independence from the British Mandate). Many of them

---

(Romanized) orthography, divorcing itself from the diglossic nature of dialect vs. standard duality. We will refer more to points of similarity between Maltese and Palestinian Arabic in the chapters to come.

9 Jews and Arabs are often referred to as cousins because of the mythological ancestry of both peoples originating from the patriarch Abraham, then diverging through his sons Ismaːʔiːl (Ishmael) and Icxak (Isaac). The former is regarded an Arab patriarch, the latter a Jewish one.
have complex identities, such as “Palestinians who happen to be citizens of Israel” and so forth. Calling the dialects Israeli Arabic or Israeli Palestinian Arabic would be convenient, yet offensive and would disregard the problematic nature of identifying this group with Israel. This has not stopped some scholars from using the term (e.g., Dekel & Brosh 2012), but out of respect for the speakers of the dialect, I will refrain from doing so.

Because this is essentially a Palestinian Arabic dialect, and because there is a multitude of such dialects anyway—as explained above (rural, urban, Bedouin, etc.), I have resolved to forego coining a new term and am simply calling the dialect under investigation “Palestinian Arabic.” When qualification is warranted, for instance, to distinguish the Jaffa variety from a West Bank (e.g., Ramallah) variety, this can easily be done by adding specific modifiers: “Jaffa Palestinian Arabic.” Or, if we need to distinguish the urban type of dialect spoken in Jaffa, Jerusalem and Nazareth from the rural type spoken in Jaljulye, Umm il-Fahm and Deir il-Asad, that too can be done by adding “Urban” or “Rural” as appropriate.

1.2.5 The site of the main study, Jaffa

According to the most recent report by the Israeli Central Bureau of Statistics, “Arabs” (i.e., Palestinians) constitute 20.69% of Israel’s population of 8,114,000.10 Talmon (2000) reports that while most (ca. 65%) of the speakers of

Palestinian Arabic within Israel are concentrated in the Galilee and in Haifa, i.e., in the northern part of the country, some 100,000 live in the southern Negev region, and over 200,000 live in the central region, within the greater metropolitan area whose core is in Tel Aviv-Jaffa. The Palestinians living in the "muθallaθ 'Triangle' area north and northeast of Tel Aviv-Jaffa are not quite a part of the cluster of suburbs and semi-industrial towns of the metropolis. Those living in the mixed (i.e., Jewish-Arab) towns of Lydd and Ramle are closer to that status.

Those living in Jaffa (Arabic jaːfa; Hebrew ‏'jafו, in more formal registers ja'fo) formerly an autonomous municipal entity and since shortly after the formation of the State of Israel part of the city of Tel Aviv-Jaffa, are in many ways full participants in the urban experience, culturally and financially. There appears to be some controversy surrounding the number of Palestinians currently living in Jaffa. According to the demographic section of the 2012 statistical bulletin for the City of Tel Aviv—Jaffa,\(^\text{11}\) out of 404,800 people living in the city as a whole (Tel Aviv and Jaffa combined), at the end of 2011, 388,100 (95.87%) were "Jewish and other non-Arabs" and the remaining 16,700 (4.13%) were "Arabs" (i.e., Palestinians). However, only about 14,000 of these (83.83%) lived in Jaffa.

Map 1: British Mandate Palestine with the borders of the 1947 UN Partition Plan and the 1949 Armistice Agreements. Created by Chris Robinson based on UN maps. (Davis 2011)

Map 2: Locations and populations of Palestinians, 2009. Created by Chris Robinson based on UNRWA maps and data. (Davis 2011)
The numbers reported by the League of the Arabs of Jaffa, a local group that describes its goal as “to preserve the Arab presence in Jaffa and to protect the rights of the Palestinians in Jaffa as an Arab Palestinian Minority [sic],”\(^{12}\) are higher. The League devotes part of its web site\(^{13}\) to “Historical Background”—about Jaffa and its Palestinian identity, not about the League itself—and opens with the following:

“The number of Arabs in Jaffa today reaches about 23,000 inhabitants. Before Jaffa fell in Israeli hands in 1948, the Arab inhabitants of Jaffa counted more than 120,000. Most of


them were forced to leave their city. Only 3900 Arabs were able to stay in Jaffa. Today Jaffa is one of the six so called Palestinian mixed cities targeted by the Israeli authorities. Jewish new comers live in the homes of the Palestinians who were forced to leave in 1948. All six cities are targeted by the Israeli authorities to turn them into marginal insignificant minorities."

It is worth noting that this web page has a “© 2007” notation at its bottom, implying that the information in it may not have been updated in seven years. This actually brings the population figures more in line with the official ones from the municipal authorities (which, in turn, are drawn from the Central Bureau of Statistics). The municipal report (written in Hebrew), indicates that the non-Jewish population of Tel Aviv—Jaffa had peaked in 2003 at 28,200, having doubled from what it had been in 1991. But the current (i.e., 2011) statistic represents a fall, following a decrease in annual natural growth, currently at -0.3%.

An important distinction that is often made is one between Muslim and Christian speakers of Arabic (not just in Palestine; see Blanc 1964 for the case of Baghdad). The municipal bulletin reports that in Tel Aviv—Jaffa the division is approximately 80% Muslim, 19% Christian and half a percent Druze. For the purpose of this study, we have not taken religious affiliation into consideration. It appears as if for younger speakers—and in general for phonological features—religion is not a sociolinguistic factor in Jaffa, but in at least one instance of a morphosyntactic feature, that of the genitive exponent, there is some evidence that historically Christians and Muslims had differed in their lexical choices for this morpheme when analytic structures were used. This feature will be discussed briefly in Chapter 5.
Also relevant is the Urban/Rural distinction (see Cadora 1992). While Jaffa has retained a small percentage of its original urban population from the pre-1948 era, many of its current residents are refugees from surrounding villages, or descendants thereof. However, all of the Jaffa subjects interviewed for this study were born and raised in the city. Some had a parent from elsewhere in Palestine, but no one was excluded for this reason.
Chapter 2: The Envelope of Variation

2.1 Overview

Having been exposed to Arabic extensively since I first started studying it in high school in 1984, I have long suspected that some of the rigid guidelines for pronunciation of the language are not followed as rigidly by native speakers as they are taught to non-native speakers such as myself. While Arabic and Hebrew are both Semitic languages of the Central Semitic branch (see Figure 2), the phonemic inventory of Hebrew is impoverished in comparison with Arabic. Old

---

14 Arabic used to be classified as a South Semitic language, alongside the Ethiopian and South Arabian languages, but a newer classification has been proposed by Hetzron in 1972, and Rubin’s (2008) family tree (Figure 2) is something of a refinement of that tree. Rubin means for it to represent the “subgrouping of the Semitic family as it is best understood given the facts available to date” (Rubin 2008:61). Note that in a subsequent analysis, Huhnergard & Rubin (2011) consider various models of “phyla,” represented graphically by trees as in Figure 2, but also entertain the thought that language contact may have played such a significant role in antiquity as to render the genetic model of language classification inadequate. They rely in part on a statement by Labov (though of course, they could have cited many others) and conclude: “Only by integrating the two models of language change, the family tree model and the wave model, can we explain the relationships among the Semitic language [sic]” (Huhnergard & Rubin 2011:267).
Hebrew had already lost several Proto-Semitic (PS) consonantal phonemes. Modern Hebrew has lost several more (more on this in Chapter 5).

2.1.1 Changes in Modern Hebrew Resulting in Divergence from Arabic

Classical Arabic (CA), with its 28 consonantal phonemes, has all but one of the PS consonants. The 29th, a voiceless lateral fricative /ɬ/ (conventionally marked as /š/ by most Semitists), is found in Old Hebrew (at least judging by the Tiberian diacritics used in the orthography of Biblical Hebrew), but has merged in Modern Hebrew with the voiceless alveolar fricative /s/. CA has a set of emphatic (CA *mufaxxam) consonants, which are pharyngealized or velarized (or, according to Shahin 1995b, 1996, uvularized) counterparts of non-emphatic consonants: /ḍ/, /ṣ/, /ṭ/, /ð/. Old Hebrew merged the first and fourth of these with the second, and Modern Hebrew merged the third with /t/ and the merged Old Hebrew /š/ is pronounced as an affricate /c/ [ʦ]. In Palestinian Arabic as in virtually every contemporary vernacular of Arabic, Classical Arabic */ḍ/ and */ð/ are merged either as a stop or a fricative, depending on whether the dialect in general has retained the pronunciation of interdental fricatives (see the

---

15 I wish not to enter the debate on the precise phonetic nature of “emphasis” in Arabic. In Shahin 1996 it is strongly argued that pharyngealization and uvularization are two discrete processes. McCarthy 1994 is also of the view that: “The so-called pharyngealized consonants of Arabic should really be called uvularized.” On the other hand, both traditional groupings of Arabic consonants and modern acoustic accounts find that the emphatics share features with the pharyngeal and uvular consonants alike and that the emphatics do have constriction in the pharynx. I will therefore use the term pharyngealized for the emphatics and transcribe them with a superscriptˤ accordingly. I will expand upon this somewhat more in Chapter 5.

16 In this section only, I am transcribing the emphatics with an underdot, as is customary among Semitists. This is mostly because it is unclear what the exact nature of PS emphatics was. It is common to think that they were historically ejective. Cf. Bergsträßer (1983:4): “The oldest pronunciation of the emphatics was probably with following release of the glottal stop, as is still the case in modern Ethiopic; this is widely replaced by a weakened pronunciation with velarization – broader contact between tongue and palate, particularly the soft palate.” Starting in section 2.1.2, I will use a more IPA-compliant transcription, as is explained in fn. 15.
discussion in chapter 5 on whether or not this is a true merger, based in part on Al-Wer 2004). Dialects that merged /θ/ and /ð/ with /t/ and /d/, respectively, typically only have a voiced emphatic alveolar stop as a reflex of both /d/ and /ð/. Dialects that have retained the non-emphatic interdentals have a voiced emphatic interdental fricative as the merged fricative. In some dialects, a new variant, a voiced emphatic alveolar fricative /ż/ has emerged, usually by means of lexical diffusion and borrowing from CA or MSA into the vernacular. In the Jaffa dialect, a typical urban Mediterranean variety, all historical interdental fricatives have alveolar plosive reflexes.

Another difference between contemporary Hebrew and Arabic is that Arabic has distinctions of quantity: consonant gemination and vowel length. Biblical Hebrew (as far as the Tiberian “pointing” system for indication of vocalization can tell us) was beginning to lose some of the length distinctions for certain vowels, in some cases substituting different vowel qualities for a PS long vowel. Also in Biblical Hebrew, certain “guttural” consonants (pharyngeals, laryngeals and the liquid /r/) were not geminated, often with compensatory lengthening of a preceding vowel. Modern Hebrew is much more categorical: gemination and long vowels do not exist.

Finally, most speakers of Modern Hebrew do not have the PS (and Old Hebrew) pharyngeal fricatives that most varieties of Arabic have retained. In Modern Hebrew, /h/ has merged with /x/, and /ʕ/ has merged with the glottal stop /ʔ/ (both of which, as well as /h/, are often realized as a phonetic zero).
2.1.2 Lenition in Palestinian Arabic

The processes of sound change that I am grouping together as “lenition” include the following:

(7) Shortening of long vowels $V_1 \rightarrow V_1$

(8) Degemination of consonants $C_1C_1 \rightarrow C_1$

(9) Depharyngealization of the voiced pharyngeal fricative $\emptyset \rightarrow \emptyset \emptyset$

(10) Depharyngealization of secondary pharyngeal articulation of emphatic alveolar stops and fricative $\emptyset \emptyset$

I am using “lenition” as a categorization of both types of features (those involving loss of pharyngeal articulation and those involving loss of length distinction), mainly because the end result of each of these processes is a less complex system, insofar as it includes fewer features from which the speaker needs to choose, and the features that are taking over are in a sense of simpler articulatory nature. This is in line with Campbell’s (1998:41) definition: “Lenition is a reasonably loose notion applied to a variety of kinds of changes in

---

17 In the formal representation of these processes I am using the synchronic arrow ($\rightarrow$) rather than the diachronic angled bracket ($>$), as for now I am only treating these processes as variable rules, which do not necessarily affect the underlying phonological value of the features involved. Once the study is complete, it may be the case that some of these processes will turn out to be cases of regular sound change with little or no residual evidence of the old forms. If this proves to be the case, such changes will warrant an angled bracket rather than an arrow.

18 I am already taking into account that Jaffa Arabic, being an urban variety of Palestinian Arabic, has merged the two voiced pharyngealized alveolars: $\emptyset \emptyset$. 

which the resulting sound after the change is conceived of as somehow weaker
in articulation than the original sounds”. Curiously, in his list of examples of
changes that may fall under the larger category of lenition, Campbell lists neither
“degemination” nor “shortening,” each of which receives its own, separate
definition (1998:42-43). Hock, however, lists at least degemination as one of the
types of “changes which have been referred to as weakening” (1991:81; Hock
uses “lenition” and “weakening” as synonyms). More suitable perhaps is the
notion that lenition may be a by-product of the Principle of least effort. Labov
(2001:16-18) reexamines Bloomfield’s proposal that “…we speak as rapidly and
with little effort as possible, approaching always the limit where our
interlocutors ask us to repeat our utterance…” and posits three rephrased
versions of the principle. From Principle of least effort I, whereby effort
reduction in speech is restricted by the need to satisfy one’s addressees’ need to
understand, through Principle of least effort II, which recognizes some loss of
meaning, and culminating in Principle of least effort III:

“Under the influence of factors a₁, a₂ ... aₙ we reduce the phonetic information that we
convey to our addressees, sometimes to the point that they do not understand us.”

(Labov 2001:17)

The validity of Principle of least effort III with respect to variable rules (7)
through (10) is probably worth examining. One way to examine whether these
sound changes introduce ambiguity to an extent that may prevent interlocutors
from understanding one another may be by using matched guise tests. More will
be said about such tests in the section on methodology (Chapter 3).
2.2 Previous treatments

2.2.1 General

In this section I will provide a representative sketch of works on variation in Arabic dialects, on descriptive works on Palestinian Arabic and on Arabic-Hebrew interaction in Palestine. Perhaps at the forefront of the social and cultural evaluation of Arabic is the work of Haeri, who began her inquiry from the standpoint of a variationist sociolinguist, following work by Labov and his colleagues. Her 1991 University of Pennsylvania dissertation (published as Haeri 1996) studied phonological variation in Cairo using a quantitative approach to language variation, with qualitative insights on the role of gender in shaping language change. Much of her later work is devoted to contextualizing the linguistic situation in the Arab World within a broader culture-based approach. Her 2000 Annual Review of Anthropology paper points to the paucity of studies of urban/literate Arabic-speaking communities.

Two additional researchers whose work on Arabic is of sociolinguistic nature, are Holes and Walters. Holes has focused on dialects of the Persian Gulf, and his 1987 book examines questions of variation and change in Bahrain. Walters, mostly in papers published in the 1980s and 1990s, does the same for North Africa, with some insight into the general questions underlying variation in Arabic and its relation to language variation in general.

Linguistic research on Arabic dialects has been abundant in the past century or so. However, most of the work has been within dialectology proper. In other
words, it involved the meticulous charting of regional dialects and careful descriptions of grammatical features thereof. Most notable in the Palestinian case are works such as Bergsträßer’s (1915) linguistic atlas of Syria and Palestine; Blanc’s (1953) description of the Druze dialect of the Western Galilee and Mt. Carmel, as well as his later (1970) study of Negev Bedouin Arabic; Levin’s studies of Jerusalem Arabic, culminating in his 1994 grammar of the dialect (published in Hebrew); various works by Rosenhouse and Henkin on the Arabic of Bedouins in Israel, including Rosenhouse’s (1984) monograph on the Bedouins of the Galilee, and a series of papers by Henkin exploring various aspects of the Negev Bedouin dialect; Talmon (until his death in July 2004), Jastrow and Behnstedt and their collaborators have been leaders in the field of Palestinian Arabic dialectology in the past decade, spearheading a joint German-Israeli research project mapping dozens of regional varieties of Palestinian Arabic, mostly in the northern parts of Palestine (but more recently expanding to the central parts of the country as well).

Shahin’s work on the phonetics and theoretical phonological aspects of rural Palestinian Arabic, as well as her 1995 grammar of the Abu Shusha dialect (Abu Shusha was a village ruined in the 1948 War in central Palestine) are useful for our study, though the nature of the dialect she reports on is different, as it is rural and has not been in contact with Hebrew. Abdel-Jawad (1981) has written on variation in Amman. While the Jordanian capital has a considerable number of Palestinian refugees whose dialect is originally of the type that the Jaffa variety belongs to, they have fairly successfully been integrated into Jordanian society. More recent works – notably those by Al-Wer since the 1990s – have
shown that a Jordanian koiné alongside such sociolinguistic identities as “Ammanis” in the Jordanian capital have been forming. Shorrab’s dissertation (1982) on phonological and stylistic variation in Palestinian Arabic bears little relevance to the current study. The subjects he interviewed were from among Palestinians residing in Buffalo, NY, and his treatment of the one variable his study and my proposal have in common, the emphatic voiced alveolar stop /dˤ/, has to do with its alternation with its interdental counterpart /ðˤ/, a matter internal to Arabic diglossia.

Abu-Lughod’s review of the anthropological literature on the Middle East mentions studies on the so-called 1948 Palestinians only in passing, and with no reference to the linguistic issues at hand. Gulick has done extensive ethnographic work in both rural and urban communities in neighboring Lebanon (whose dialect of Arabic is also closely related to Palestinian Arabic). A University of Chicago doctoral dissertation in anthropology by Monterescu (2005) focuses on Jaffa and other mixed (Palestinian-Jewish) urban communities in Palestine. A dissertation on language and social identities in another mixed town, Haifa, was completed by Lefkowitz in 1995 (published as a monograph in 2004). Its linguistic accounts of the Palestinians living in Haifa are based on their Hebrew, not their Arabic, with the pharyngeals as a salient variable for asserting “Arab-ness”. While Palestinians use pharyngeals in Hebrew more than Mizrahi Jews19 (who have not lost the pharyngeal feature completely), they still show “a

---

19 Mizrahi (Modern Hebrew mizwa‘xi, literally ‘eastern’) refers to Jews of North African and Middle Eastern origin whose language background includes some regional dialect of Arabic, often a Jewish variation thereof. This term is nowadays preferred by many such Jews over the overarching Sephardi (Modern Hebrew sfasa‘di, literally ‘Spanish’), which was used until recently to denote virtually all Jews of non Ashkenazic origin.
wide range of variation.” My examination of their Arabic sheds some light on the manifestations of this variation in their Hebrew as well.

2.2.2 Previous work on the variables at hand

The four features (7) through (10), which are subject according to my hypothesis to variable rules, have received treatment in at least some dialect of Arabic, in at least one framework of descriptive, theoretical or variationist linguistics. They are renumbered here as (7’) through (10’) with brief summaries of the literature for each.

(7’) Shortening of long vowels

In his 1994 grammar of Jerusalem Arabic, a dialect closely related to Jaffa Arabic, Levin writes (in Hebrew; my translation – UH):

“In Jerusalem Arabic, long vowels cannot exist in unstressed syllables. Therefore, any vowel which is a long vowel in Literary (i.e., Standard/Classical – UH) Arabic changes to a short vowel in Jerusalem Arabic, when it is contained in an unstressed syllable. Examples: mafatīḥ > mafatīḥ (‘keys’ – UH) [...] safārna > safārna ‘we traveled.”


Raz (1996) is of the view that in pausal stressed syllables, historically long vowels in Jerusalem Arabic are only “potentially long vowels” (1996:196), unlike Damascus Arabic, in which vowel shortening does not occur. Raz questions the phonemic value of long vowels, but provides no further account of any factors which may govern variation other than stress, pause, and “vowel prominence”.

(8') Degemination of consonants

I have not found much about this phenomenon in the literature about Arabic. Rosenhouse (2002:601) cites two environments in “colloquial Arabic in Israel” in which “[w]eakening or complete loss of gemination” may occur: in cases where there is underlying cluster of the type $C_1C_1C_2$ (e.g., $mʕallme→mʕalme$ ‘teacher-F’); and in word final position (e.g., $sʕaff→sʕaf$). McCarthy (1994) mentions Semitic degemination in the known cases of Hebrew and Tigre (an Ethio-Semitic language), but only in the context of “guttural” consonants, and in any case, not in Arabic.

(9') Depharyngealization of the voiced pharyngeal fricative

Again, Rosenhouse (2002) shows some evidence of this phenomenon in Palestinian Arabic, which she co-classifies with the “weakening of the emphatics” (our “depharyngealization of secondary pharyngeal articulation of emphatic alveolar stops and fricative”; see (10) next).

McCarthy (1994) and Shahin (1996) make the case that the fricatives /h/ and /ʕ/, which have primary pharyngeal articulation, share the feature [PHAR] with the emphatics, whose primary articulation is coronal, despite their assertion that a more precise characterization of their phonetic nature is as uvularized, not pharyngealized.
Shahin (1995) provides some evidence from acquisition of Palestinian Arabic by her own son. While at first glance it seems as if the child, Hosam, acquired both the glottal stop /ʔ/ and the voiced pharyngeal fricative /ʕ/ by age 1;11, it is mentioned in a footnote (Shahin 1995:115) that the two phones “have an identical UR for Hosam”. Puzzled by that, I contacted the author via e-mail and her response (dated Nov. 25, 2003) was that “Hosam - in the corpus, which was from 1;11 - 2;8.5 - always produced a glottal stop for a target voiced pharyngeal /ʕ/ (except postvocally, where he omitted the target pharyngeal).” My understanding of this is that by the end of the data collection period he had not actually produced the voiced pharyngeal. This finding is consistent with Omar’s (1970) study of the acquisition of Egyptian Arabic. Omar shows that /ʕ/ is the fourth-to-last consonantal phoneme acquired by Egyptian children, at an average age of 4;6. (Omar 1970:158). Furthermore, it has not been found among Omar’s sample before the age of 4, and is “continued to be mispronounced as [ʔ] or ∅ in isolated cases long after its acquisition as a phoneme.” (1970:153). These data support the general hypothesis that the voiced pharyngeal is a phone prone to change or even elimination, but its vulnerability in dialects that cannot be suspect of being influenced by Hebrew seems, at least tentatively, to counter the hypothesis that contact with Hebrew is a necessary contributor to such a change. However, such general vulnerability might entail that situations such as contact with a language like Modern Hebrew could easily trigger or promote such a change.

---

20 This particular piece of evidence is not without problems. The reported child’s mother is a native speaker of Canadian English. His father is of Palestinian origin, and the report refers to the child’s first years of speech, which were predominantly in British Columbia.
(10’) Depharyngealization of secondary pharyngeal articulation of emphatic alveolar stops and fricative

As part of her study of palatalization of alveolar stops in Cairene Arabic, Haeri (1996: Ch. 3) has found that the probability (indicated by Varbrul weights) of palatalization for /tˤ/ (.53) is higher than that of /d/ (.43), even though the [+back] feature of pharyngealization is inconsistent with the [-back] feature of palatalization. Citing previous studies by Royal (1985) and Kahn (1975), Haeri concludes (1996:57) that the pharyngealized voiceless alveolar stop “loses its pharyngealization variably and becomes a plain [t] [...] Probably some of the pharyngeal[ized] phonemes are merging with the pharyngeal[ized] phonemes.” Once again, Egyptian Arabic exhibits processes similar to those observed in Palestinian Arabic. For this reason, the contact hypothesis must be scrutinized and tested using acoustic measures followed by multivariate analysis, with intensity of contact as a category of factor groups to be examined.

The lenition variables outlined above may be classified as belonging to two groups. The first pertains to segment length, and the second has to do with primary and secondary pharyngeal place of articulation. The detailed sociolinguistic study in this thesis will deal primarily with the latter group.
Chapter 3: Methodology

3.1 Preliminaries

This chapter will describe the type of fieldwork I had initially planned to conduct, it will discuss the rationale behind my sampling of the subjects, outline the basic structure of the interviews, and provide a list of factor groups. Subsequently, a more concrete description of the circumstances under which I worked will be provided, as well as a list of interviewees and their demographic and socio-economic backgrounds, and several examples demonstrating the structures of the actual interviews.

3.2 The Fieldwork

The main bulk of data needed for any study of sociolinguistic variation and change in a contemporary spoken language consists of extensive samples of spontaneously produced speech. The observer's paradox notwithstanding (Labov 1984:30), face-to-face sociolinguistic interviews have been the most fruitful means of obtaining a large amount of speech in the vernacular, “in which the minimum attention is paid to speech” (Labov 1984:29). It is the vernacular which is considered, following Labov, to be “the most systematic data for linguistic analysis” (Labov 1984:29). The vernacular is defined, for our purpose, as the variety of language acquired by the speaker in her/his pre-adolescent years, in which minimum attention is paid to speech. Stylistic variation within the interviews will be evaluated through comparisons with media speech (see
below) and with a small sample of recordings of family and peer group interactions.

Fieldwork for this project was initially planned for the first seven months of 2004, during which I would be physically present in the proverbial urban field in central coastal Israel for the purpose of data collection. This period was eventually extended to include most of the remainder of 2004, and an additional month in the spring of 2005.

I settled in a rented apartment in the demographically diverse section of Jaffa known most commonly as ˈlev ˈjafo, Hebrew for ‘Heart of Jaffa’, due to its central location on sdeˈisot  yesuʃaˈlaim  ‘Jerusalem Boulevard’\(^1\), the eastern of Jaffa’s two main north-south thoroughfares. It is an area that was historically heavily populated by Jewish immigrants from Bulgaria, as is apparent from the four Bulgarian or Balkan restaurants located within a two-block radius of my apartment building. The current demographics, however, have shifted. Word on the street is that “the Bulgarians” have been upwardly mobile and resettled in the all-Jewish town of Bat Yam (literally, in Hebrew, ‘daughter of the sea’\(^2\)), Jaffa’s immediate neighbor to the south, with the Palestinians from the predominantly Arab neighborhood of ʕaʒami spreading eastward to the proverbial “heart” of the city. More recently, particularly in the mid- to late

---

\(^1\) This is the Hebrew name of the street and the one most commonly used by Hebrew and Arabic speakers alike. Part of the Arabic portions of many of the interviews included a discussion of street names and their evolution under the Ottoman, British and Israeli governance of Jaffa.

\(^2\) Interestingly, and most probably not by accident, Jaffa still holds for many the romantic epithet ʕarūs al-bahr (Arabic for ‘bride of the sea’). Bat Yam adopted its current name in 1938, twelve years after it was originally founded as a “Hebrew settlement south of Jaffa” (according to the official home page of the City of Bat Yam, available, as of July 2007, in Hebrew only: [http://www.bat-yam.muni.il/show_item.asp?levelId=57658](http://www.bat-yam.muni.il/show_item.asp?levelId=57658)). The web site has since been updated (accessed 4 February 2014) to include English, French, Russian and Spanish, but the information cited above is still only available in Hebrew.
1990s and early 21st century, immigrants from Ukraine, Russia and other former Soviet republics have been mingling into the populace in the neighborhood, rendering it diverse again in a somewhat different fashion.

3.3 The Sample

3.3.1 The main sample

My main pool of subjects consisted of 24 people aged 16 and up, who have lived all or most of their life, since childhood, in Jaffa. The intent was to fill a grid of three age groups and three emulations of socioeconomic statuses (SES).

The rationale behind this sampling was as follows. I expected there to be some correlation between each age and SES category and the intensity of contact with Hebrew. “Blue collar” is defined here as workers in the auto businesses (mechanics, body shop workers, tire shop employees, etc.), whose clientele often includes many Jewish customers, and perhaps construction workers, whose bosses (contractors, etc.) tend to be Jewish. People in the food & services category include restaurant servers, grocery store workers, etc., who may encounter Hebrew speakers in their line of work as well. White-collar speakers include teachers, doctors, pharmacists, business owners, etc. these are typically the more educated and/or wealthier members of the community. They have typically been in contact with Hebrew speakers at least for some period of time, during their time as students of higher education, as virtually all post-secondary schooling in the Israeli educational system is conducted in Hebrew. Some of them, however, have studied abroad, in which case they have had a prolonged
exposure to some other languages (typically Russian, other Eastern European languages, or English; and in the case of some speakers in both the older and younger extremes of the spectrum – Arabic [through studies in Lebanon before 1948 or Jordan since the dawn of the 21st century, following the signing of the 1994 Jordan-Israel Peace Treaty]). Yet many of them are employed within the Palestinian community and may not have a lot of daily contact with Jews in the workplace. Unclear to me prior to embarking in fieldwork was the proper classification of high school and university students, who are not yet full members of the job market. One option would be to classify them by type of high school: vocational high schools (ones training their pupils for various trades, e.g., welding, auto mechanics) may be considered equivalent to “blue collar”; mainstream academic high schools leading to an Israeli Matriculation Diploma would be equivalent to “white collar”; high school dropouts (if any) would be classified according to their current occupation as either “blue collar” or “food & services”. Another option, more akin to common practice in sociolinguistics and the social sciences, would be to consider their parents’ SES and classify the offspring accordingly. Frankly, I believe that for children still in school, SES is not as relevant as other factors. One model to be considered is that of Community of Practice (Eckert & McConnell-Ginet 1992). However, as Eckert explicitly clarifies (2000:171): “A community of practice is not a unit, like a social category, that exists on one level and to which speakers can be assigned. This is specifically because speakers belong to multiple communities of practice on multiple levels.” One implication of this caveat is that variable rule analysis can be somewhat clouded with vagueness, yet not impossible. Eckert adds (2000:172): “But the necessities of statistical analysis cannot be confused with the mechanisms of
social reality.” Eventually, I created a separate, dummy SES category for these teenagers, simply calling them “teenagers.”

I encountered a similar dilemma with regard to educational level. On the one hand, a 15 year-old student studying in the 10th grade technically has the same formal education as a 60 year-old adult who dropped out of high school in the 10th grade. On the other hand, the 15 year-old is still in school, and in all likelihood will complete her or his education, and perhaps even continue and earn an academic degree. After much deliberation, I decided to recode high school students as “pupils.” This, obviously, created what statisticians call “confounding factors,” as being between the age of 16 and 18, having the education level “pupil” and the SES “teenager” essentially meant the same thing. Rbrul provides tools to control for these by running tests with “pairwise interactions.” This is not an ideal solution for this kind of case, but it was an ad hoc way to deal with this issue.

The age categories were chosen for the following reasons. All speakers up to age 60 were expected to have been taught Hebrew in school as a second language (or, as suggested above, as an “L2.5,” after MSA), as they have had all begun their primary schooling after the Israeli Ministry of Education had taken over in 1948. The older speakers (61 and up) may or may not have had a full curriculum of Hebrew, and for that reason I have chosen to include fewer of them in the sample. The 35/36 cutoff line between the two younger groups roughly corresponds to the 1966 ending of martial law for most 1948 Palestinians (see White 2012:73-76 on “military regime”). This had not affected Palestinian citizens of Israel in a mixed town like Jaffa as much as it had in other locales, but
it was expected to have some impact on attitude toward the state nonetheless. I chose to include speakers as young as 16 years old to enable a glimpse into the high school community, which is in itself quite diverse. Teenagers in Jaffa (or in many cases, their parents) have more choice nowadays. Some go to the local municipal Arab high school, where Arabic is the language of instruction; others go to Jewish schools (some of which are by now mixed Jewish/Arab), where Hebrew is used as the primary language; and some go to church-run schools, where interesting, and seemingly not carefully planned combinations of Arabic, French, English and even Russian are used in the curriculum and among the student populace.

3.3.2 The control group

A smaller sample of speakers (20-25) was scheduled to be interviewed from within a predominantly monolingual Palestinian speech community that is not in close contact with Hebrew speakers. The most likely candidate for this part of the study was Ramallah, a West Bank town some 60 km southeast of Jaffa. It is an urban setting within the same dialect region as Jaffa, but its speakers do not necessarily know any Hebrew, and in any case are not involved in interactions in Hebrew on a daily basis. Through this control group I was hoping to determine whether having no contact with Hebrew still yielded the same processes that otherwise seem to be very much Hebrew-like.

In practice, I interviewed 12 speakers, 6 of whom entered the quantitative analysis. The interviews were conducted on two occasions. The first was a
professional seminar held in Jerusalem, in which the participants were all middle class economists and members of their families. They were all natives of Jerusalem, Ramallah, and their environs. Some of these speakers attested to have originated in more rural locales in the vicinity, but they have all adopted an urbanite-type dialect. On a second occasion, I traveled to Ramallah and spent the day at a local government office and interviewed several employees. Some of them were middle-class professionals (mostly economists and lawyers), and others were clerical staff and lower-level employees. They, too, were all from the Ramallah-Jerusalem urban cluster. These interviews were conducted in late 2004 and early 2005.

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th>Jaffa</th>
<th>West Bank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>14-35</td>
<td>9</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>36-60</td>
<td>3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>61+</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Grand total</td>
<td>24</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Number of speakers sampled by age and socioeconomic status

3.4 The interviews

In addition to some of the standard urban topics of discussion that sociolinguists use to elicit vernacular forms (danger of death, premonitions, childhood games, etc.), I saw the need to construct a number of modules that would address questions of language contact and language attitude. Examples of questions of this sort can be found in the interview excerpts in a study of Anglophones in Quebec, by Nagy, Moisset & Sankoff (1996).
In the Jaffa case, similar modules were adapted to fit the local setting. Part of my strategy was to conduct the interview with a short Hebrew component, leading to a longer portion in Arabic. It had been my experience that as a non-Arab who happens to speak Arabic, I am often identified as an “other” (more specifically, a Jewish Israeli, regardless of my own personal views of my identity). Oftentimes, when I initiate a dialogue in Arabic with Palestinians, my interlocutors reply in Hebrew and impose a switching of the language of the interaction. Since I wanted to gather some information not only about the speakers’ own assessment of their Hebrew and their level of contact with Hebrew speakers, but also about their actual level of proficiency in Hebrew and the degree to which their Hebrew resembled that of native speakers, it seemed like a good idea to commence each interview with the Hebrew component, including, inter alia, an explicit language-centered module of questions and a short reading passage, and then introducing Arabic through an abrupt shift on my part in the form of “okay, now in Arabic!” (uttered in Arabic).

3.5 Treatment of the data

Acoustic analysis of the data collected has been done using the freeware application Praat. Quantitative analysis was done using Rbrul (Johnson 2009), following recent practice in variationist sociolinguistics. Chapter 4 of this thesis includes the lion’s share of the statistical analysis, and within this chapter I explain the specific advantages of using Rbrul for this particular study.
Appendix A provides a list of factor groups for the multivariate analysis. As will be explained in the data chapters (4 & 5), it was only feasible to code with any degree of precision for the first variable, (ʕ), and it is for this variable that this appendix pertains. The second variable, (EMPH), is analyzed qualitatively and within a sociohistorical framework. The remaining variable sketched above, as well as additional contact-induced features of the dialect, will also be discussed, though not thoroughly analyzed.

For (ʕ), each token examined is coded for one dependent variable reflecting its actual realization by the speaker and by a number of independent variables. Some independent variables are constant to all tokens per speaker, such as those relating to the speaker’s demographic information. In what follows, I will list the factor groups, and for those that are not self-explanatory, I will add a brief explanation.
Appendix A: Factor groups for multivariate analysis

**DEPENDENT VARIABLE**

For (ʕ)

<table>
<thead>
<tr>
<th>Code</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phar</td>
<td>[ʕ] (or creaky voice)</td>
</tr>
<tr>
<td>Glot</td>
<td>[ʔ]</td>
</tr>
<tr>
<td>Syllable</td>
<td>Addition of a syllable in lieu of consonantal realization</td>
</tr>
<tr>
<td>Length</td>
<td>Compensatory lengthening of adjacent vowel</td>
</tr>
<tr>
<td>0</td>
<td>Ø</td>
</tr>
</tbody>
</table>
**INDEPENDENT VARIABLES**

*SOCIAL FACTORS/SPEAKER INFO*

Age

<table>
<thead>
<tr>
<th>Code</th>
<th>Age range</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-35</td>
<td>16-35</td>
</tr>
<tr>
<td>36-60</td>
<td>36-60</td>
</tr>
<tr>
<td>61+</td>
<td>61+</td>
</tr>
</tbody>
</table>

Sex

<table>
<thead>
<tr>
<th>Code</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>f</td>
<td>female</td>
</tr>
<tr>
<td>m</td>
<td>male</td>
</tr>
</tbody>
</table>
### Occupational group

<table>
<thead>
<tr>
<th>Code</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue</td>
<td>“blue collar” (auto mechanics, construction, etc.)</td>
</tr>
<tr>
<td>serv</td>
<td>food &amp; services (waiters, shopkeepers, etc.)</td>
</tr>
<tr>
<td>white</td>
<td>“white collar” (teachers, doctors, pharmacists, business owners, etc.)</td>
</tr>
<tr>
<td>teenager</td>
<td>aged 16-18, not yet in the job market</td>
</tr>
</tbody>
</table>

### Education level

<table>
<thead>
<tr>
<th>Code</th>
<th>Schooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>elem</td>
<td>elementary school (up to 8 years)</td>
</tr>
<tr>
<td>high</td>
<td>high school (9-12 years)</td>
</tr>
<tr>
<td>uni</td>
<td>university/college (including grad/professional degrees)</td>
</tr>
<tr>
<td>pupil</td>
<td>still in (high) school</td>
</tr>
</tbody>
</table>
Language of primary/secondary schooling

<table>
<thead>
<tr>
<th>Code</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arab</td>
<td>Predominantly Arabic</td>
</tr>
<tr>
<td>Heb</td>
<td>Predominantly Hebrew</td>
</tr>
<tr>
<td>mix</td>
<td>Roughly even mix between Arabic &amp; Hebrew</td>
</tr>
</tbody>
</table>

Contact with Hebrew

<table>
<thead>
<tr>
<th>Code</th>
<th>Frequency of contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>no contact</td>
</tr>
<tr>
<td>1</td>
<td>occasional contact (1-2 times a week)</td>
</tr>
<tr>
<td>2</td>
<td>extensive contact (works/studies/lives with Hebrew speakers)</td>
</tr>
</tbody>
</table>

The values for this factor group has been determined by asking each speaker explicit questions about their frequency of contact with Hebrew.
Hebrew proficiency

<table>
<thead>
<tr>
<th>Code</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>none</td>
</tr>
<tr>
<td>1</td>
<td>intermediate</td>
</tr>
<tr>
<td>2</td>
<td>high</td>
</tr>
</tbody>
</table>

These values have been determined by my judgment of the Hebrew component of each interview (if there was one).

Hebrew phonology

<table>
<thead>
<tr>
<th>Code</th>
<th>Status of Hebrew pharyngeals</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>no pharyngeals in Hebrew</td>
</tr>
<tr>
<td>phar</td>
<td>pharyngeals in Hebrew intact</td>
</tr>
<tr>
<td>partial</td>
<td>pharyngeals in Hebrew variably intact</td>
</tr>
<tr>
<td>NoHeb</td>
<td>Speaker does not speak Hebrew, or there was no Hebrew component to the interview.</td>
</tr>
</tbody>
</table>

Ditto.
### Community

<table>
<thead>
<tr>
<th>Code</th>
<th>Community</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jaffa</td>
<td>Jaffa</td>
</tr>
<tr>
<td>WB</td>
<td>West Bank (Ramallah, Jerusalem)</td>
</tr>
</tbody>
</table>

### Linguistic Factors

**Position**

<table>
<thead>
<tr>
<th>Code</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>onset</td>
<td>onset</td>
</tr>
<tr>
<td>coda</td>
<td>coda</td>
</tr>
<tr>
<td>cluster</td>
<td>consonant cluster</td>
</tr>
</tbody>
</table>
Preceding segment

<table>
<thead>
<tr>
<th>Code</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Segments indicated “as is”

Following segment

<table>
<thead>
<tr>
<th>Code</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Segments indicated “as is”
Chapter 4: Analysis and discussion: variable 1 – (ʕ)

4.1 Description of the variable

4.1.1 A complex variable

Upon embarking on the fine-grained work of coding the corpus for this variable, it became clear that it is a more complex variable than I had originally envisaged it to be. I had suspected that there would be more to this variable than simply a pharyngeal realization – the traditional [ʕ] – and a mere deletion of the segment. What I had thought would be an intermediate variant, based on casual auditory observation of the data I had been collecting, was what impressionistically appeared to be a glottal stop [ʔ].

4.1.2 Two types of vocalization

However, when I began the fine-grained coding, I discovered the existence of at least two types of vocalic variants of (ʕ) in the Palestinian dialect. One of them I am calling ‘compensatory lengthening,’ as it involves simply the lengthening of the preceding otherwise short vowel, e.g., \(ba\de:n \rightarrow bax\de:n\) ‘later.’ The second type of vocalic variant is ‘syllabic vocalization.’ It consists not only of the addition of a vocalic mora, but also of resyllabification of the word. It often occurs at the final word boundary, though it is not limited to this position. A prime, recurring example of this phenomenon is found in the word \(us.bu:ʕ \rightarrow us.bu.a\) ‘week.’ What I find intriguing about this particular variant in the context of the Arabic-Hebrew language contact situation – and this specific lexical item
happens to illustrate this quite neatly – is that there is a similar phenomenon in Hebrew, traditionally known as *patah gənuva* ‘a stolen short [a] vowel,’ such as in the Hebrew equivalent of *us.buː* ~ *us.bu.a* ~ *fa.vu.a(ʕ)* – the [ʕ] only pronounced nowadays by Hebrew speakers whose dialects are influenced by Arabic (usually because of ancestry within the Middle East or North Africa).

4.1.3 Creaky voice

It is worth noting here that an additional type of variant was quite common in the data analyzed for this study. I had originally considered coding this variant, often referred to in the literature as ‘creaky voice,’ as separate from the traditional pharyngeal [ʕ], and in a subsequent study may in fact return to these tokens and study them in their own right. However, I have chosen, for the sake of both convenience and a certain degree of logic, which I shall defend henceforth, to code the creaky voiced tokens as ‘regular’ pharyngeals. The convenience argument has to do with the lack of certainty in distinguishing creaky tokens from pharyngeal ones in a good deal of the cases. On the one hand, many profoundly creaky tokens do show up on spectrographic images in quite recognizable fashions. On the other hand, I decidedly only used Praat and its spectrographic feature as an aid in determining the quality (and in the case of compensatory lengthening – quantity) of variants, secondary to my own auditory impressions of the phonations. For this reason, in those cases in which there was a discrepancy between Praat showing some degree of creakiness, but the auditory impression being that of an actual pharyngeal, I preferred to code the tokens as the latter.
4.1.4 A hierarchy of lenition

In terms of logic, my rational was as follows. The idea behind coding for multiple types of variants was that there was a hierarchy of lenition, possibly related to contact between Arabic and Hebrew, and that this hierarchy was gradient. This gradient nature of the variants would subsequently allow for a multivariate analysis (e.g., using Rbrul), treating the variables on a continuous scale. It was fairly clear where on this scale the four variants I eventually coded for would fall. Adding ‘creaky voice’ into the mix would have potentially jeopardized the analysis, as it would have been virtually impossible to assess where this variant belongs on the continuous scale.

It is worth noting here that ‘creaky voice’ (or ‘creaky phonation’ as a more technical, phonetic term) has been widely observed in a variety of dialects (see, e.g., Heselwood 2007:6, 13, 17). One variety in which it is common to see this feature is Maltese, which is of course tempting for drawing analogies with Palestinian Arabic because of the contact situations both varieties have been subject to. What is interesting about creaky voice as a reflex for /ʕ/ in Maltese, however, is that the literature seems to mostly attest to this feature being prevalent in the dialects of Gozo, the smaller inhabited island of Malta, which is considered to have a more conservative dialect, one that preserves pharyngeal consonants (see, e.g., Hume et al. 2009:15, fn 1). Agius (1992:130) informs us that “[a] ‘creaky voice’ to describe the alternation in the pronunciation of the voiced pharyngeal fricative […] occurs practically in all Gozitan dialects.” He further asserts that (similarly to the Palestinian case) “at times it is lost except
for a compensatory lengthening of adjacent vowels.” Compensatory lengthening for historical /ʕ/ is also the focus of Hume et al.’s 2009 study involving two speakers from the island of Malta itself (not Gozo). Their findings, however (e.g., p. 42) indicate that unlike the Palestinian data collected for the present study, vowels adjacent to a historical voiced pharyngeal are typically not lengthened to the same extent as phonemically long vowels.

In summation, the five variants eventually coded for with respect to the (ʕ) variable are as follows, in ascending order of lenition:

(1) Pharyngeal
(2) Glottal
(3) Compensatory lengthening
(4) Syllabic vocalization
(5) Deletion

It will later be explained how, for the purposes of multivariate analysis of (ʕ) as a continuous variable, the aforementioned variants have been recoded on a scalar continuum.
4.2 Quantitative analysis

4.2.1 Four facets of analysis

Among the many different ways one could have analyzed the data emerging from this corpus with respect to the variable [ʕ] from a quantitative perspective, it seemed to make sense to approach the data somewhat gradually, in what resulted in four facets of analysis:

(1) cross-tabulations of elements of the data per speaker.

(2) cross-tabulations and multivariate analysis of the data by factor groups, observing the pharyngeal variant [ʕ] of the variable vs. all of the lenited variants as an aggregate.

(3) cross-tabulations and multivariate analysis of the data by factor groups, observing the pharyngeal variant [ʕ] of the variable vs. total deletion only.

(4) multivariate analysis of the data by factor groups, observing the five different variants as representatives of a continuous variable.

In what follows, we will present a summary of these four types of observation and analysis, highlighting their potential contribution to our understanding of the role of language contact in the process of language variation and change.
4.2.2 Cross-tabulations of the data per speaker

4.2.2.1 "Eyeballing the data"

At the outset of the process of quantitative analysis, it seemed prudent to observe the realization of the (Ć) variable individually for each speaker in the sample. This was done because the data were not evenly distributed across the social factors traditionally taken into account in variationist sociolinguistic studies, viz., age, gender and socioeconomic status. An attempt was made to include representatives of the various groups for each of these factors, but at the time of the study, it was unknown how relevant these factors would be in this under-studied speech community. In addition, the sample size was relatively small, especially given that it actually comprised two sub-samples: the main sample from Jaffa and the control group from the West Bank. In the first two charts below, even though the main focus is on individual speakers, each sub-sample is color-coded differently, because of the initial hypothesis that the Jaffa speech community would, as an aggregate, pattern differently from the West Bank speakers, primarily because of the differences in levels of contact with Hebrew.

In my days as a graduate student at the University of Pennsylvania, we were often advised by Bill Labov and Gillian Sankoff to “eyeball the data.” This is a concept that is hard to come by in formal guides to quantitative analysis of linguistic variation, but is often practiced. Walt Wolfram mentions it explicitly in his interview to Natalie Schilling [then Schilling-Estes]: “What does it say when you tell a student, ‘Keep working with the statistics until you get a test that works,’ as opposed to eyeballing the data and saying, ‘Hmm. That seems to
indicate a pattern,’ even though it doesn’t quite meet the .05 probability threshold for statistical significance for one reason or another.” (Schilling-Estes 2008:7). Figure 1 and Figure 2 below are designed mostly to be “eyeballed,” not for significance tests to be run on their results. Similarly, Jennifer Hay advises students of sociophonetics to include a step she calls “[v]isualizing your data” (Hay 2011:201). In addition to Wolfram’s insight about the perils of probability testing prior to manual inspection of data, Hay adds: “Exploratory data analysis can reveal a lot about the patterns in your data set, and lead you to interpretations and hypotheses that were not in your initial plan.” One technique she strongly advocates for is inspecting a scatterplot. While this technique may be useful for the type of sociophonetic data described in her paper (e.g., English vowel formants), in our case, cross-tabulation can fulfill the same function of a preliminary quantitative summation of the dataset.
4.2.2.2 Starting from a binary opposition

Figure 1 assumes a very simple binary opposition: either a pharyngeal (including, as described above, creaky-voiced) [ʕ] was produced as the variant for ⟨i⟩ or it was deleted altogether. This is in line with older methods of variable rule analysis, which required the dependent variable in a sociolinguistic study to be a binary variable, due to the nature of the mathematical model with which the variable rule analysis programs available until just a few years ago had operated (See, e.g., Tagliamonte 2006:160, 2012:121, 127; and in particular Johnson 2009:359 – “A variable rule program evaluates the effects of multiple factors on a binary linguistic ‘choice’ – the presence or absence of an element, or any
phenomenon treated as an alternation between two variants.”). While the current statistical analysis was done using Rbrul, the very software package developed by Johnson in order to expand the statistical capabilities of variationist sociolinguists beyond such binomial conjunctures, its ability “to perform regression, with continuous responses” (Johnson 2009:362) is but one of its advantages over its predecessors, of the Varbrul/Goldvarb family. And in any case, some features, such as cross-tabulation, do not lend themselves very easily to representation of the data without its manipulation into binary variables. When we turn to the multivariate component of the analysis later on, we will explore both binary models and a continuous one, capitalizing on this novel feature offered to us by Rbrul.

What Figure 1 also affords us is a glance into a hypothetical distilled situation in which only two variants were possible. This is obviously not what was in fact observed, and additional analysis will be provided henceforth to complement the preliminary view offered in Figure 1, but what it displays is interesting nonetheless. In Figure 1, the green bars on the left represent the control group speakers from Ramallah and East Jerusalem, who are not in close everyday contact with Hebrew speakers and who do not possess proficiency in Hebrew in addition to their proficiency in Arabic (though many of them do speak English or French, which is why it would be misleading to label them ‘monolingual’ speakers of Arabic). The blue bars represent the speakers from the main sample.

---

23 Another helpful guide to quantitative analysis of language variation is Guy (1993). Gorman & Johnson (2013) offer a “state-of-the-art” of various techniques of statistics for sociolinguistics. My own quantitative analysis has also benefited from personal communication with Kyle Gorman and William Cotter. I am particularly grateful to Lauren Hall-Lew (University of Edinburgh) for sharing her teaching materials and for offering additional related valuable advice and resources.
of bilingual Arabic-Hebrew speakers from Jaffa (again, some of them have knowledge of other languages as well, but these are not germane to this study). The application value here was chosen to be [ʕ], simply to show a quick-and-dirty picture of the realization of the traditional voiced pharyngeal fricative across the urbanite Palestinian speech community.

Eyeballing the data on the bar graph in Figure 1 reveals that (ʕ) is variable not only in Jaffa, where Hebrew is hypothesized to be a force leading to contact-induced change, but also in other Palestinian speech communities. Furthermore, this variability is apparent in as extreme a variant as full deletion of the pharyngeal segment; the most linguistically conservative West Bank speaker in the sample, ‘husni,’ a 1965-born man, has a value of 0.876 for [ʕ], and the most advanced, ‘nuha,’ a 1966-born woman, has a value of just 0.581.

4.2.2.3 Individualized explanations for individual speakers

The range in Jaffa is much wider, however, and given the larger sample size and greater representability of age and social groups, we may be able to offer individualized explanations for at least some of the otherwise seemingly erratic behaviors within this group. The prime example for this is ‘nabiːl,’ who is not only the most conservative speaker linguistically, but is also the oldest in the sample (he was born in 1926). He also holds views about Palestinian nationalism and the decline of Jaffa and its culture since the advent of Zionism, which would make it reasonable for him to resist influences from Hebrew into his native Arabic. There are additional factors in his biography, which we will
later see factoring into the larger picture, that render him a lesser candidate for the lenition of (ʕ), such as the fact that he had been educated in Beirut prior to the 1948 *nakba*, and that his education had been predominantly administered in Arabic rather than in Hebrew. It is also notable that his interview was recorded in 1999, five to six years prior to the rest of the corpus.

Two of the youngest speakers, high school students ‘amal’ and ‘sawsan’ (both girls, born 1990), were also the most advanced linguistically, with values for [ʕ] at just barely above 0.3. This is not surprising, given their youth and the role girls and women play cross-linguistically in promoting linguistic change. Yet some of the more interesting speakers are the ones with less obvious demographic characteristics. One such speaker is ‘umm jaziːd,’ a woman estimated to have been born around 1948-1950 who works as a fortune-teller in Jaffa. Her [ʕ] value is a mere 0.472, placing her in the same range as the much younger ‘neviːn,’ (female teacher, born 1971, educated predominantly in Hebrew, high frequency of Arabic-Hebrew code switching, [ʕ] value 0.4) and ‘ʒamiːl,’ (gay male artist/student, born 1982, educated in Hebrew-speaking schools, socially immersed in the queer community of Tel Aviv, [ʕ] value 0.456).

While ‘neviːn’s’ and ‘ʒamiːl’s’ low frequencies of pharyngeal realization of the (ʕ) variable are easily described in cursory parentheticals appended to their pseudonyms, ‘umm jaziːd,’ the middle-aged fortune-teller, requires some

---

24 This is by now a well-documented sociolinguistic principle, support for which is abundant in a wide variety of publications. Suffice it here to cite two comprehensive studies that lay the foundations to a broader basis of knowledge surrounding this matter: Labov (1990, in particular pp. 212-213, where the author references the works of Abd-el-Jawad on Arabic and Farsi, and p. 215 and beyond, in which he discusses the innovative role of women in cases of “change from below;”); Eckert 1989, regarding the distinction between sex as a biological category and gender as a social construct in relation to language change – a precursor perhaps to the a series of publications that lay out a more elaborate theory of gender and social categories, viz., that of Communities of Practice.
elaboration for us to understand why her speech patterns with that of a
generation her junior, with a very different educational and sociopolitical
upbringing. Unlike most interviews I have conducted, which were variations on
the theme of the Labovian sociolinguistic interview, my interview with ‘umm
jaziːd’ began as a fortune-telling session. In fact, my desire to interview her
emerged when I was speaking with her son, ‘jaziːd,’ an old acquaintance of mine
from my undergraduate days at Tel Aviv University and a native of Jaffa himself.
I was making arrangements with ‘jaziːd’ to interview him, and he said that his
house would not be a very quiet environment for recording an interview,
because his mother has people coming in and out for her sessions reading her
clients’ coffee grounds (a traditional Middle Eastern method of fortune-telling).

I was fascinated by the idea, and asked whether ‘jaziːd’s’ mother would be
willing to be interviewed and recorded while performing the ritual of reading my
coffee grounds. She agreed, and I eventually interviewed both mother and son.
‘jaziːd’ was interviewed more conventionally in a local café, and ‘umm jaziːd’ in
her home, with the interview beginning as a fortune-telling session.

Early on, ‘umm jaziːd’ made it clear to me that she was not behaving –
linguistically – like her normal self. She told me that her son had “cautioned” her
that she “must speak Arabic” with me. I tried to soften the blow and explain that
there were no rigid rules, and that while I wanted to hear her speak Arabic, and
her local dialect at that, I also wanted to hear how she spoke Hebrew, hear what
her views on language were, and in general, make her feel natural and
comfortable rather than stressed and placed under artificial constraints.
Most of our conversation continued in Arabic, but at some point it was interrupted by a phone call she received, which was held in Hebrew. I took this opportunity, having heard her speak fluent, proficient Hebrew, to ask her how comfortable she felt speaking Hebrew and related questions about interacting in the two languages. At that point, ‘umm jazi:d’ gave me a bit of a motherly lecture.

The following excerpt is translated from Hebrew:

Uri: Even when you speak with Arabs, you insert Hebrew?

U.J.: Yes, with Arabs... I insert Hebrew, yes. Even with my kids I insert Hebrew. Well, it’s... Pay attention: the first language in this country is Hebrew, then Arabic. And I, all my life... all my friends and all my acquaintances, everyone’s Israeli.

Uri: So actually you end up speaking most of the day Hebrew, more than Arabic?

U.J.: Yes, yes, more than Arabic, I speak. Why, all my clients are Jewish. It’s like 20%... Let’s say 80% Jews, 20% Arabs.

It is unknown what ‘umm jazi:d’s’ political views are. Her son told me his views fluctuated “between Trotskyism and anarchism,” but this doesn't necessarily reflect on the entire family, of course. It would be extremely unlikely, however, to assume that she is a staunch supporter of the Zionist nationalist movement. Why, then, would she espouse such ideas about the importance of Hebrew being the dominant language in her community? If we keep the discussion within the framework of language and gender research, Eckert (1989:256) may provide a very convincing explanation: “An important part of the explanation for women's
innovative and conservative patterns lies, therefore, in their need to assert their membership in all of the communities in which they participate, since it is their authority, rather than their power in that community, that assures their membership."

We know that ‘umm jaziːd’ participates in more than one community, not only linguistically, but also socially, ethnically, professionally, and so on. There exists, however, a great deal of overlap between the linguistic aspect of these communal divisions and all the rest. She participates in some social and professional communities as a Hebrew speaker (an L2 Hebrew speaker, but this is not uncommon, even among Jewish Israelis, many of whom are immigrants), and in some – in the case of professional interactions, only a fifth of the time, by her account – in Arabic. It is therefore a logical conclusion that she will exhibit innovative forms in her native language to assure membership through authority, as Eckert asserts. Add to that the fact that the innovations she is exhibiting in her speech are akin to features already available in Hebrew, and the explanation is fortified.

An interesting speaker with whom to compare ‘umm jaziːd’ is ‘saːlim.’ They are about the same age (he was born in 1948, a fact he uses to make an interesting statement; see below). They are both roughly of the same socioeconomic background – some sort of lower middle class, if such a category applies in this community – ‘saːlim’ works as a night watchman for the City of Tel Aviv-Jaffa, but has managed to send his kids to private schools for several years. They also appear to share, to some extent, an attitude about the significance of Hebrew in
their community. ‘sa:lim’ and I had the following exchange (again, the excerpt below is translated from Hebrew):

Uri: After all, Hebrew is, is...

sa:lim: What, what, it’s the mother tongue, let’s put it this way: it’s the mother tongue.

Uri: Is Hebrew really your mother tongue?

sa:lim: Yes! I’m a native of this country.

Uri: Yes.

sa:lim: I was born in this country.

Uri: You were born when, what year?


Uri: ’48, together with the state [of Israel] you were born.

sa:lim: Together with the state.

The assertion that one is a native speaker of a language simply because one was born in the same year as the political entity where that language is spoken as a majority language, despite one clearly being part of a linguistic, ethnic and religious minority, seems naïve from a linguistic perspective. Eckert’s explanation cited above regarding the desire to assert membership in all communities to which one belongs applies here as well. Despite this, ‘sa:lim’s’
[ʕ] value is much more in congruence with that of other members of his generation: 0.659 (e.g., ‘sa:mi,’ 1934-born man, tire garage owner, at 0.667).

4.2.2.4 Beyond the binary

Figure 2, while still reporting results of cross-tabulations of the data, is somewhat more sophisticated. For each speaker, two values are reported. The left-hand side bar for each speaker represents a value for the application of a variable rule that assumes deletion of the pharyngeal only, ignoring all intermediate variants. Note that unlike in Figure 1, in Figure 2 the application value is a ∅ realization of the variable rather than [ʕ]. The bar on the right for each speaker aggregates all

**Figure 2:** Cross-tabulation of (ʕ) by speaker, community and 2 types of lenition
of the lenited variants: deletion (i.e., $\emptyset$), compensatory lengthening, and syllabic vocalization and treats them collectively as the application value, in opposition to the fully pharyngeal [ʕ] variant.

Again, the purpose of this cross-tabulation and its graphic representation is mostly for getting a general feel of the data, prior to submitting it to more rigorous “number crunching.” What I see here is a rather convincing confirmation of the general hypothesis at hand that correlates lenition of pharyngeal articulation with contact a speaker has with Hebrew. This is true across the sample. In other words, while the distinction between the bilingual Jaffa speakers and the non-bilingual West Bank speakers is helpful to maintain a perspective regarding speakers’ general proficiency with the superstrate language and its potential to induce change on a daily basis, high variability is apparent within the bilingual group as well.

Several Jaffa speakers have relatively low values for deletion (the blue bars in Figure 2), but when considering all forms of lenition (the red bars), their values surge. Consider such speakers as ‘tˤaːriq,’ who was an 18 year-old high school student at the time of the interview (born 1987). His value for $\emptyset$ realization of (ʕ) is 0.109, but his value for lenition in general is 0.562. ‘umm xali:'l’ is an even more striking case. Not only is she much older (born 1928), but her life trajectory is much more complex. She had left Jaffa in 1946 to marry a man in the village of abu kifk, where the dialect spoken is of a Bedouin type (I have met her husband, and he still speaks a Bedouin dialect), and after being banished from their village (which was eventually destroyed) in 1948 to Gaza and to
northern Sinai in Egypt, they returned to Palestine, but to a different village altogether, *ʒalʒuːlje*, where I ended up interviewing her in 2004.

Similarly to ‘tˤaːriq,’ ‘umm xaliː:l’s’ speech exhibits a very low rate of full deletion of the voiced pharyngeal fricative (0.111). Her value for lenition at-large is seven-fold at 0.752. I have no clear-cut explanation for this pattern, only a few educated guesses. We can begin with the gender hypothesis, which was discussed above. As a woman, albeit the oldest woman in the sample, it is not unreasonable for us to expect somewhat more advanced tokens from her than from her male counterparts. Also recall that at the time of the interview, ‘umm xaliː:l’ had been displaced from the original environment of her native dialect for six decades. She had acquired neither the rural, *fallaːhi*, dialect of her “new” domicile nor any of the dialects of the places in which she lived during her transition there (including her husband’s Bedouin dialect). She reports still visiting Jaffa from time to time (a mere 35-km drive from her current village). While her Hebrew proficiency is the lowest of all the Jaffa speakers in the sample, and her level of contact with Hebrew speakers the lowest as well, she is still exposed to Hebrew through mass media and other sources.

All of the above amount to a rather nebulous set of factors, which admittedly could have led to either a more conservative linguistic behavior or, as indeed we see in this case, a more advanced one. I tend to think that her rather special position in her family and community have led her to assert her urban origins in some ways that she is aware of (e.g., the non-pronunciation of interdentals and the glottal realization of historical /q/, cf. Shahin 2008:527) and in others – such
as the lenition in the case of the (ʕ) variable – that are below her level of consciousness.

4.2.2.5 Sounding “native”

Before turning to a description of the results of the multivariate analysis, I wish to address one more issue arising from the less-accurate cross-tabulated data, particularly as they are represented in Figure 2. What emerges is some sort of correlation between the ratio of the “blue” and “red” bars for each speaker and the subjective impression they may leave regarding their Arabic “nativeness.”

A speaker like ‘ʒamiiːl,’ for example, is often mistaken for a non-native speaker of Arabic by colleagues of mine to whom I’ve played his speech. Furthermore, his Hebrew, to me as a native speaker of that language, sounds as close to native as it can be, and I would even characterize it as a middle-class, ‘hip’ sort of Hebrew, which does not fit the stereotypical image of the Hebrew spoken by a Palestinian. His overall lenition value is high, but not as high as I had expected it to be: 0.687 (lower than the much older ‘umm xaliːl’ – 0.752 – and, e.g., ‘jaziːd,’ who is also 10 years his senior, heterosexual, and whose educational trajectory included more Arabic than ‘ʒamiiːl’s,’ with 0.812).

‘umm xaliːl’ and ‘jaziːd’ “sound more Arab” than ‘ʒamiiːl,’ even though their absolute values for overall lenition are higher than his. But they both have much lower total deletion (i.e., the Ø variant) values than he does: 0.111 and 0.182 for ‘umm xaliːl’ and ‘jaziːd’ respectively and 0.373 for ‘ʒamiiːl.’
Table 2 introduces a new metric – the ratio per speaker between the value for the $\emptyset$ variant and the value for all lenited variants. It will take further investigation beyond the scope of this study to unravel the precise parameters of this metric, but looking at the speakers above, a plausible working hypothesis may be formulated as follows:

(1) The lower the metric, the more native the speaker will be perceived.

Needless to say, such a hypothesis cannot be confirmed with only three speakers. Adding a few more speakers is somewhat useful, but also complicates the matter, as Table 3 illustrates.

<table>
<thead>
<tr>
<th></th>
<th>$\emptyset$ only</th>
<th>All lenited variants</th>
<th>$\emptyset$ : lenited ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>umm xaliːl</td>
<td>0.111</td>
<td>0.752</td>
<td>0.148</td>
</tr>
<tr>
<td>jaziːd</td>
<td>0.182</td>
<td>0.812</td>
<td>0.224</td>
</tr>
<tr>
<td>ʒamiːl</td>
<td>0.373</td>
<td>0.687</td>
<td>0.543</td>
</tr>
</tbody>
</table>

Table 2: Deletion to lenition ratio in 3 speakers
<table>
<thead>
<tr>
<th></th>
<th>Ø only</th>
<th>All lenited variants</th>
<th>Ø : lenited ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>nabiːl</td>
<td>0.053</td>
<td>0.325</td>
<td>0.163</td>
</tr>
<tr>
<td>nuha</td>
<td>0.212</td>
<td>0.706</td>
<td>0.300</td>
</tr>
<tr>
<td>neviːn</td>
<td>0.379</td>
<td>0.747</td>
<td>0.507</td>
</tr>
<tr>
<td>ħusni</td>
<td>0.104</td>
<td>0.261</td>
<td>0.398</td>
</tr>
<tr>
<td>bijaŋka</td>
<td>0.117</td>
<td>0.539</td>
<td>0.217</td>
</tr>
</tbody>
</table>

Table 3: Deletion to lenition ratio in 5 additional speakers

It almost appears as if the working hypothesis holds very neatly, even when adding the five speakers represented in Table 3. I would argue that a speaker like ‘neviːn’ definitely “deserves” a high score, if that score means that her Arabic sounds less native than many of her peers. She was born in 1971, educated mostly in Hebrew, from the middle of her primary school career into her PhD studies, which she was engaged in at the time of the interview. She actually hesitated to be interviewed as all, claiming that she didn’t even speak Arabic. It was only after I had caught her speaking on the phone with a friend in fluent Arabic that I persuaded her that she was capable of conducting a conversation in the language. Yet the task was indeed difficult for her, and her Arabic was
consistently interlaced with Hebrew expressions, to the point where at times, I knew she was code-switching between the two languages, but I was not quite sure which of the two was the matrix language (see Myers-Scotton 1993:66).

I can also understand the values for ‘nabi:l,’ who is described in an earlier section above as an older, linguistically conservative Jaffa speaker; for ‘nuha,’ a linguistically advanced West Bank speaker – her score is high for a non-bilingual, but still well below 0.5; and for ‘bijanja,’ a semi-retired seamstress from Jaffa (born 1935), who had worked for many years alongside Hebrew-speaking colleagues, though as her year of birth indicates, had lived for over a decade before Hebrew became the language of the majority in her surrounding region.

The outlier in this table is ‘husni.’ This isn’t only because he is a fairly typical West Bank speaker, who has no particular contact with Hebrew in his daily life, but also because of the specific numbers that formulate the relatively high ratio for him in the right-hand column. Here is where Wolfram et al.’s advice to “eyeball” the data comes in handy. While I’m not yet prepared to abandon the metric I have proposed (though, arguably, it needs some refinement), it may only be valid for certain values. It works nicely for speakers like ‘jami:l,’ whose values for both deletion and overall lenition are already fairly high. But for ‘husni,’ both of these values are very low to begin with. Why then, would we even bother looking at the ratio of the two? Imagine a case where a speaker had both a deletion and an overall lenition value of exactly 0.1 each. Her ratio would then be 1.0. Would that mean that she sounds non-native? Not at all, given the initial values we used to calculate that very ratio. At this juncture, I’ll leave it at
that, and hope to further develop this metric (or a better substitute for it) in a more large-scale subsequent study.

4.2.3 Multivariate analysis

When I set out to conduct this study, I knew that there would be hurdles when I’d reach the stage of attempting a multivariate analysis. As mentioned earlier, the most robust work after which to model my work, one that dealt with language contact within a variationist framework, was Naomi Nagy's PhD dissertation on Faetar. Nagy was cognizant of the problematic nature of this complex type of analysis, or in her words: “Numerous problems arise when one attempts to construct a model for comparing the effects of these factors. First, it is clear that many of these factors interact” (Nagy 1996:27). That this is a problem may not be obvious at first glance, but the literature on multivariate analysis is quite clear on this matter. Consider the following: “Multivariate statistical methods assume that the predictors are "orthogonal," that is, fully independent of each other” (Gorman & Johnson 2013:217). Programs such as Rbrul have options to somehow account for “pairwise interactions,” but to the best of my understanding, even when such an option is used, the analysis may be skewed, except that the interaction will be reported in the analysis output.

Some of the social variables in the current study, in particular several of the variables examined in order to assess speakers’ degree of contact with Hebrew, are not at all independent of one another. For instance, it is fairly obvious that the factor group “language of primary/secondary schooling” is at least partially
interrelated with the factor group “age,” because members of the oldest age group did not have the option to be schooled in Hebrew prior to 1948, when they were growing up. Similarly, the factor group “contact with Hebrew” is not totally independent of the factor group “occupational group,” because certain occupations lend themselves more readily to everyday interaction with speakers of Hebrew.25

Nevertheless, since I am using multivariate analysis – and statistics in general – as a mere tool and not as a be all and end all, I am satisfied in presenting the output of such an analysis along with this disclaimer and the accompanying qualitative commentary, both in this chapter and in Chapter 6, devoted to the discussion of the results of the entire study.

It seemed to make sense to attempt to run the analysis in a number of different ways: with the data from both Jaffa and the West Bank control group and with just the Jaffa data; as two different binomial logistic regressions – one with ∅ as the sole application value and one with all non-pharyngeal variants as the application value; and as linear regressions, i.e., treating the dependent variable as a continuous variable on a scale from 0 to 4, 0 representing the ∅ variant (deletion) and 4 representing the [ʕ] variant, and 1-3 as the intermediate variants.

25 I intentionally refrained here from referring explicitly to native speakers of Hebrew, because many of the Jewish Israelis with whom the Palestinians from Jaffa (and elsewhere in 1948 Palestine) interact are immigrants for whom Hebrew is not their native language, but within the Israeli context use Hebrew as the main medium of communication. Myhill (2004:104-105) addresses this issue when he discusses what he calls “[l]ingua franca uses” of Hebrew, problematic as his analysis may be.
4.2.3.1 Jaffa & control group (West Bank) with binary variable

Table 4 (note that some of the following tables to follow span across several pages and include a number of sub-tables within) introduces the first set of results from an Rbrul run, drawing from data from both Jaffa and West Bank speech communities, observing the variable as binary, with the application value being the $\emptyset$ variant. As customary in multivariate analyses, a step-up/step-down run was conducted (this is true of all of the runs in this study; this detail will not be repeated henceforth). As often happens, the step-up and step-down runs produced a slight mismatch. Reported below in Table 4 are the main results from the step-down run. These were chosen because it had a slightly higher $R^2$ value (defined as “the proportion of variance ‘explained’ by the regression

Figure 3: Organization of multivariate analysis
model,” Nagelkerke 1991:691) than the step-up run. In each instance henceforth, I will report in table form the results from the run with the highest $R^2$ (also known as the coefficient of determination, or the multiple correlation coefficient, Nagelkerke 1991:691). Additional findings, including those from other runs, will only be reported verbally.

Also reported verbally will be results pertaining to preceding and/or following segments. These were coded as the actual discrete segments preceding and following the independent variable for each token. While in some runs this factor group was found to be significant, the multitude of factors within the factor groups, many of which have very few tokens, make it quite difficult to tease out patterns for these factor groups. In future research, these factor groups will have to be reevaluated and probably recoded in some phonologically meaningful way with only a handful of options.

Also note that Table 4 includes two metrics, which essentially express the same relation: log-odds and centered factor weights. This will be the only table in which factor weights will be reported, as indeed reporting both log-odds and factor weights is redundant. They are only mentioned here at all because the use of log-odds, one of the novelties of the Rbrul program for conducting multivariate analysis of linguistic variation (see Johnson 2009) is only a few years old, and may still be unfamiliar to many readers, especially those whose specialty is Arabic (socio-)linguistics.

As a reminder, factor weights, a relic from the days of Varbrul/GoldVarb (e.g., D. Sankoff et al. 2005 in its latest version), "is a number between zero and one that indicates to what extent and in what direction the factor affects the rate of
application of the rule. These values pattern as follows: a value above .5 is a factor which favors the application of the rule, while a value below .5 indicates a factor which disfavors the rule” (Guy 1993:244). Log-odds are numbers between \(-\infty\) and \(\infty\), such that any positive numbered log-odd indicates favoring application of the variable rule, and negative numbered log-odds indicate disfavoring. A log-odd of 0 is equivalent to a GoldVarb factor weight of 0.5 indicating neutrality (Johnson 2009: 361; also see Fig. 1 on that page for a conversion chart between log-odds and factor weights).

<table>
<thead>
<tr>
<th>Age group (p&lt;0.001)</th>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>36–60</td>
<td>0.349</td>
<td>762</td>
<td>0.586</td>
<td></td>
</tr>
<tr>
<td>14–35</td>
<td>0.314</td>
<td>798</td>
<td>0.578</td>
<td></td>
</tr>
<tr>
<td>61+</td>
<td>-0.662</td>
<td>332</td>
<td>0.34</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex (p&lt;0.0005)</th>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.233</td>
<td>923</td>
<td>0.558</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.233</td>
<td>969</td>
<td>0.442</td>
<td></td>
</tr>
</tbody>
</table>

\[R^2=0.304\]
### Occupational group (p<10^-6)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teenager</td>
<td>0.562</td>
<td>336</td>
<td>0.637</td>
</tr>
<tr>
<td>Blue collar</td>
<td>0.463</td>
<td>307</td>
<td>0.614</td>
</tr>
<tr>
<td>Service</td>
<td>-0.374</td>
<td>128</td>
<td>0.408</td>
</tr>
<tr>
<td>White collar</td>
<td>-0.651</td>
<td>1121</td>
<td>0.343</td>
</tr>
</tbody>
</table>

### Language of primary/secondary schooling (p<10^-11)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hebrew</td>
<td>0.726</td>
<td>277</td>
<td>0.674</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.340</td>
<td>224</td>
<td>0.584</td>
</tr>
<tr>
<td>Arabic</td>
<td>-1.066</td>
<td>1391</td>
<td>0.256</td>
</tr>
</tbody>
</table>

### Level of regular contact with Hebrew speakers (p<0.05)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
<th>Centered factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.608</td>
<td>1207</td>
<td>0.648</td>
</tr>
<tr>
<td>1</td>
<td>0.145</td>
<td>213</td>
<td>0.536</td>
</tr>
<tr>
<td>0</td>
<td>-0.753</td>
<td>472</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Most of the results shown in Table 4 confirm the hypothesis correlating language contact and the change in progress in Palestinian Arabic with reference to the weakening of pharyngeals. The favoring of the younger two age groups, which include speakers born after 1948, and most importantly educated – at least in the Jaffa speech community – under the auspices of the Israeli Ministry of Education is in line with this rational, as is, of course the factor group relating to language of schooling (note the extremely high level of statistical

---

26 While some Jaffa speakers in the sample were educated in private Christian schools, they have all reported that Hebrew was part of their school curriculum and general upbringing.
significance). The self-reported level of contact with Hebrew speakers factor group is also consistent with this pattern.

Favoring of women versus men in the multivariate analysis is consistent with preliminary observations we have noted in the cross-tabulations above. As for occupational group, the separation of teenagers from the rest of the sample may have skewed the results somewhat. We already know, from the results regarding the age group factor group, that young speakers favor deletion. But blue-collar workers also favor deletion. To some this may seem surprising, but this is actually what I had hypothesized from the start, as I had anticipated that blue-collar workers (and some service workers) would have more daily contact with Hebrew speakers than white-collar workers. Since this run included both Jaffa and West Bank speakers, and the West Bank speakers were all either service or white-collar workers, it makes sense that only blue collar workers made the cut here.

Linguistic factors: two factors for position of the variable were found to favor deletion: consonant clusters (e.g., {tallamet}, ‘I studied’ → [tallamet]) and onsets (e.g., {aruːs}, ‘bride’ → [aruːs]). While, on the one hand, neither of these positions as deletion-favoring ones should surprise us – the former because of its phonological complexity and the latter because of a diminished functional load in the onset compared to the coda,\textsuperscript{27} we shouldn’t lose sight of the disparity between the paucity of tokens for clusters versus the large quantity of tokens for the other two factors in the group. This is not an ideal sample for any statistical analysis.

\textsuperscript{27} Also recall that in coda position, a more likely form of lenition is that of syllabic vocalization, e.g., {maw.dˤuːʕ}, ‘subject’ → [maw.du.a].
Also note (not in the table) that preceding segment and following segment were found to be significant factor groups at significance levels of $p<10^{-17}$ and $p<0.0005$, respectively. Analysis from the step-up run (also not shown in the table) suggests that speakers for whom we have no recorded data in Hebrew (most of these are West Bank speakers) and Jaffa speakers who pronounce historical Hebrew pharyngeals as actual pharyngeal some (but not all) of the time, are also favored for deletion. Again, this is consistent with some of the cross-tabulation observation, namely that West Bank speakers may tend to delete but not weaken their pharyngeals in other ways.

The next Rbrul test – examining, still, both speech communities, but expanding the application value to include not only deletion of the pharyngeal but all intermediate variants as well – produced quite similar results, though slightly nuanced. These are presented in Table 5.
### Sex (p<0.001)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>0.155</td>
<td>1615</td>
</tr>
<tr>
<td>M</td>
<td>-0.155</td>
<td>1554</td>
</tr>
</tbody>
</table>

### Education (p<0.0005)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>0.954</td>
<td>153</td>
</tr>
<tr>
<td>Current pupil</td>
<td>-0.011</td>
<td>601</td>
</tr>
<tr>
<td>Secondary</td>
<td>-0.364</td>
<td>572</td>
</tr>
<tr>
<td>University</td>
<td>-0.579</td>
<td>1843</td>
</tr>
</tbody>
</table>

### Language of primary/secondary schooling (p<0.05)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hebrew</td>
<td>0.306</td>
<td>434</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.188</td>
<td>450</td>
</tr>
<tr>
<td>Arabic</td>
<td>-0.494</td>
<td>2285</td>
</tr>
</tbody>
</table>
### Level of regular contact with Hebrew speakers (p<0.05)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.749</td>
<td>2060</td>
</tr>
<tr>
<td>1</td>
<td>0.211</td>
<td>442</td>
</tr>
<tr>
<td>0</td>
<td>-0.961</td>
<td>667</td>
</tr>
</tbody>
</table>

### Hebrew proficiency as assessed by researcher (p<0.005)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.737</td>
<td>667</td>
</tr>
<tr>
<td>1</td>
<td>-0.737</td>
<td>374</td>
</tr>
<tr>
<td>2</td>
<td>-0.737</td>
<td>2128</td>
</tr>
</tbody>
</table>

### Realization of pharyngeals in Hebrew speech (p<0.01)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial</td>
<td>0.540</td>
<td>732</td>
</tr>
<tr>
<td>Pharyngeal</td>
<td>0.540</td>
<td>476</td>
</tr>
<tr>
<td>0</td>
<td>-0.027</td>
<td>1294</td>
</tr>
<tr>
<td>No Hebrew data</td>
<td>-0.513</td>
<td>667</td>
</tr>
</tbody>
</table>

Table 5: Rbrul results for Jaffa & West Bank (binary: all lenition variants)
The main contributions of this test are that when variants other than $\emptyset$ are considered, education is picked by Rbrul as a significant factor group. Caution is advised here, however, because the relative low number of tokens (153) is the typical N for many of the Jaffa speakers in the sample. This has alerted me to revert to my original spreadsheet and check my suspicion that there was only one speaker, from Jaffa, whose education level has not surpassed the primary school level. And indeed, there is only one such speaker, ‘umm xaliːl,’ whose high level of lenition (but not deletion) was discussed earlier in conjunction with the data presented in Figure 2. We can therefore discount this finding in the multivariate analysis, as it does not add anything meaningful to our overall knowledge of the linguistic patterns of the speech community as a whole.

It is interesting – and at first glance counterintuitive – that bilingual speakers who realize (all or some) of the historical Hebrew pharyngeal consonants as pharyngeals, tend to lenite them in their native Arabic. What this test actually tells us may be interpreted somewhat differently, especially in light of the overall results we’d seen earlier in Figure 2. We have established that everyone has lenition in their speech. The question now is what variants of lenition do they use more frequently. In the first test (Table 4) this was not consistently found to be a significant factor group, because only full deletion, the most extreme form of lenition, was considered. And in the one run in which it was found to be significant, it pertained to speakers with partial – but not full – pharyngeal realization. Following this logic, it is therefore much less surprising that when adding the three intermediate lenited variants, the more linguistic
conservative speakers are also found to be favored for lenition (again, because everyone has some kind of lenition in their speech).

Of the internal (linguistic) factor groups, only adjacent segments – both preceding and following – were found to be significant (p<10^{-7} for both factor groups).

In the step-up run, which once again mismatched slightly with the step-down run, language of schooling was not chosen as a significant factor group. However, community was found to be highly significant (p<10^{-19}), but with a rather odd outcome: both Jaffa and West Bank resulted in log-odds of 0. The formal explanation for this is that while the factor group “community” is significant for lenition, whether a speaker is from Jaffa or the West Bank is neutral regarding their tendency to have lenition in their speech. This may just be one of these cases in which quantitative analysis and human common sense simply do not interact favorably with one another.

4.2.3.2 Jaffa only with binary variable

The next set of results, presented in Table 6, reverts to a simplistic treatment of the variable (application=deletion only), but it focuses on the Jaffa speech community only. This enables us to analyze the role of language contact within this community rather than across the two communities as has been done in the previous two tables. The main results here are consistent with previous ones, with a few interesting nuances. For age, the middle group only is favored for

---

28 Both runs resulted in identical coefficients of determination (R^2).
deletion. But note that the inclusion of “Teenager” as an occupational group and its favoring of deletion – alongside “Blue collar” – suggests a sub-division of age that we had not previously considered. A reanalysis of age in light of the results spanning both factor groups reveals that in fact the youngest speakers (current secondary school pupils) as well as speakers in their mid-thirties up to age sixty, are in the pharyngeal-deleting favoring group. We see the youngest age group emerging as a favored group once again when pairwise interactions are reported.

<table>
<thead>
<tr>
<th>Age group (p&lt;0.05)</th>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>36-60</td>
<td>0.365</td>
<td>290</td>
<td></td>
</tr>
<tr>
<td>14-35</td>
<td>-0.024</td>
<td>798</td>
<td></td>
</tr>
<tr>
<td>61+</td>
<td>-0.341</td>
<td>332</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupational group (p&lt;10^{-10})</th>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teenager</td>
<td>0.858</td>
<td>336</td>
<td></td>
</tr>
<tr>
<td>Blue collar</td>
<td>0.076</td>
<td>307</td>
<td></td>
</tr>
<tr>
<td>Service</td>
<td>-0.144</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>White collar</td>
<td>-0.791</td>
<td>705</td>
<td></td>
</tr>
</tbody>
</table>

R^2=0.272
### Language of primary/secondary schooling (p<10^{-16})

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hebrew</td>
<td>0.646</td>
<td>277</td>
</tr>
<tr>
<td>Mixed</td>
<td>0.314</td>
<td>224</td>
</tr>
<tr>
<td>Arabic</td>
<td>-0.960</td>
<td>919</td>
</tr>
</tbody>
</table>

### Realization of pharyngeals in Hebrew speech (p<0.001)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.226</td>
<td>786</td>
</tr>
<tr>
<td>Partial</td>
<td>0.218</td>
<td>400</td>
</tr>
<tr>
<td>Pharyngeal</td>
<td>-0.444</td>
<td>234</td>
</tr>
</tbody>
</table>

### Position of (ʕ) in word (p<0.005)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster</td>
<td>0.752</td>
<td>61</td>
</tr>
<tr>
<td>Onset</td>
<td>-0.190</td>
<td>1012</td>
</tr>
<tr>
<td>Coda</td>
<td>-0.563</td>
<td>347</td>
</tr>
</tbody>
</table>
### Pairwise interaction – Age group : Occupational Group

(p<0.05)

<table>
<thead>
<tr>
<th>Factor:Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-35:White collar</td>
<td>0.647</td>
<td>462</td>
</tr>
<tr>
<td>36-60:Blue collar</td>
<td>0.647</td>
<td>85</td>
</tr>
<tr>
<td>61+:Blue collar</td>
<td>0.647</td>
<td>222</td>
</tr>
<tr>
<td>14-35:Service</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>14-35:Teenager</td>
<td>0.000</td>
<td>336</td>
</tr>
<tr>
<td>36-60:Service</td>
<td>0.000</td>
<td>72</td>
</tr>
<tr>
<td>36-60:Teenager</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>61+:Service</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>61+:Teenager</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>14-35:Blue collar</td>
<td>-0.647</td>
<td>0</td>
</tr>
<tr>
<td>36-60:White collar</td>
<td>-0.647</td>
<td>133</td>
</tr>
<tr>
<td>61+:White collar</td>
<td>-0.647</td>
<td>110</td>
</tr>
</tbody>
</table>

Table 6: Rbrul results for Jaffa only (binary: deletion)

The following results, in Table 7, are, as well, from Jaffa only, and they include in the application value all forms of lenition. In this instance, again, the youngest group does not initially appear to favor lenition, and in fact the oldest speaker do favor lenition. However, the “Current pupil” educational group does favor
lenition (but not in its pairwise interaction with age, which, admittedly, is an odd result; though at least in that case it is reported to be neutral, not disfavoring lenition).

<table>
<thead>
<tr>
<th>Age group (p&lt;0.0005)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>36-60</td>
</tr>
<tr>
<td>61+</td>
</tr>
<tr>
<td>14-35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education (p&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>Current pupil</td>
</tr>
<tr>
<td>Primary</td>
</tr>
<tr>
<td>Secondary</td>
</tr>
<tr>
<td>University</td>
</tr>
<tr>
<td>Language of primary/secondary schooling (p&lt;10^-9)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>Mixed</td>
</tr>
<tr>
<td>Hebrew</td>
</tr>
<tr>
<td>Arabic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of regular contact with Hebrew speakers (p&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Position of (ʕ) in word (p&lt;10^-17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Cluster</td>
</tr>
<tr>
<td>Onset</td>
</tr>
<tr>
<td>Coda</td>
</tr>
</tbody>
</table>
Table 7: Rbrul results for Jaffa only (binary: all lenition variants)

<table>
<thead>
<tr>
<th>Factor:Factor</th>
<th>Log-odds</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>14–35:University</td>
<td>1.037</td>
<td>787</td>
</tr>
<tr>
<td>36–60:Primary</td>
<td>1.037</td>
<td>0</td>
</tr>
<tr>
<td>61+:Primary</td>
<td>1.037</td>
<td>153</td>
</tr>
<tr>
<td>14–35:Secondary</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>14–35:Current pupil</td>
<td>0.000</td>
<td>601</td>
</tr>
<tr>
<td>36–60:Secondary</td>
<td>0.000</td>
<td>297</td>
</tr>
<tr>
<td>36–60:Current pupil</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>61+:Secondary</td>
<td>0.000</td>
<td>275</td>
</tr>
<tr>
<td>61+:Current pupil</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>14–35:Primary</td>
<td>-1.037</td>
<td>0</td>
</tr>
<tr>
<td>36–60:University</td>
<td>-1.037</td>
<td>238</td>
</tr>
<tr>
<td>61+:University</td>
<td>-1.037</td>
<td>151</td>
</tr>
</tbody>
</table>

4.2.3.3 Jaffa only with continuous variable

The results in the final two tables (Table 8 & Table 9) are to be taken with a grain of salt. First of all, we have much less experience dealing with the results of linear regression (i.e., the multivariate analysis of continuous variables) than of
logistic regression (analysis of binary variables) in sociolinguistics, particularly in Arabic.

Secondly, the $R^2$ values in both of these cases (just over 5%) indicates that these models only account for less than a fourth of the variance than that accounted for in the previous models (where $R^2$ was equal to or greater than 20%). It is, however, a novel first step, which needs to be refined in future studies.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.098</td>
<td>1554</td>
<td>2.651</td>
</tr>
<tr>
<td>Female</td>
<td>-0.098</td>
<td>1615</td>
<td>2.410</td>
</tr>
</tbody>
</table>

Occupational group ($p=1.05\times10^{-13}$)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service</td>
<td>0.300</td>
<td>215</td>
<td>2.372</td>
</tr>
<tr>
<td>White collar</td>
<td>0.298</td>
<td>1773</td>
<td>2.570</td>
</tr>
<tr>
<td>Blue collar</td>
<td>-0.082</td>
<td>580</td>
<td>2.697</td>
</tr>
<tr>
<td>Teenager</td>
<td>-0.516</td>
<td>601</td>
<td>2.300</td>
</tr>
</tbody>
</table>
### Language of primary/secondary schooling (p=2.29e-22)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>0.575</td>
<td>2285</td>
<td>2.680</td>
</tr>
<tr>
<td>Mixed</td>
<td>-0.089</td>
<td>450</td>
<td>2.378</td>
</tr>
<tr>
<td>Hebrew</td>
<td>-0.485</td>
<td>434</td>
<td>1.885</td>
</tr>
</tbody>
</table>

### Realization of pharyngeals in Hebrew speech (p<0.005)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharyngeal</td>
<td>0.243</td>
<td>476</td>
<td>2.668</td>
</tr>
<tr>
<td>Partial</td>
<td>0.020</td>
<td>732</td>
<td>2.527</td>
</tr>
<tr>
<td>0</td>
<td>-0.041</td>
<td>1294</td>
<td>2.325</td>
</tr>
<tr>
<td>No Hebrew data</td>
<td>-0.222</td>
<td>667</td>
<td>2.823</td>
</tr>
</tbody>
</table>

Table 8: Rbrul results for Jaffa & West Bank (continuous variable, social factors only)
<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service</td>
<td>0.250</td>
<td>145</td>
<td>2.248</td>
</tr>
<tr>
<td>Blue collar</td>
<td>-0.067</td>
<td>580</td>
<td>2.697</td>
</tr>
<tr>
<td>Teenager</td>
<td>-0.468</td>
<td>601</td>
<td>2.300</td>
</tr>
</tbody>
</table>

Language of primary/secondary schooling ($p<10^{-22}$)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>0.582</td>
<td>1618</td>
<td>2.621</td>
</tr>
<tr>
<td>Mixed</td>
<td>-0.092</td>
<td>450</td>
<td>2.378</td>
</tr>
<tr>
<td>Hebrew</td>
<td>-0.491</td>
<td>434</td>
<td>1.885</td>
</tr>
</tbody>
</table>

Realization of pharyngeals in Hebrew speech ($p<0.005$)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Coefficient</th>
<th>Tokens</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharyngeal</td>
<td>0.200</td>
<td>476</td>
<td>2.668</td>
</tr>
<tr>
<td>Partial</td>
<td>-0.091</td>
<td>732</td>
<td>2.527</td>
</tr>
<tr>
<td>0</td>
<td>-0.109</td>
<td>1294</td>
<td>2.325</td>
</tr>
</tbody>
</table>

Table 9: Rbrul results for Jaffa only (continuous variable, social factors only)
While Rbrul itself may be somewhat limited in offering us tools for interpreting this type of outcome, the following preliminary conclusions may be drawn:

1. On a scale of 0 to 4, women/girls tend to have lower value for pharyngeality than men.

2. Occupational group is a significant factor group, with white-collar speakers favoring pharyngeality and blue-collar speakers favoring lenition/deletion.

3. The role of primary and secondary school language of instruction in determining one’s propensity for lenition is reinforced in the linear regression, in the same direction as in the logistic regression.

4. There appears to emerge a correlation between speakers’ realization of pharyngeal segments in Hebrew (if they are proficient in that language) and their tendency to retain the original pharyngeals in their native Arabic. This, I believe, is an important finding, as it really is only clearly apparent when the variable was recoded as a continuous one. The methodology for determining speakers’ Hebrew phonology/phonetics must, however, be refined, as in this instance it was done as an impressionistic, almost subjective, determination by the interviewer/researcher.

29 I have consulted Kyle Gorman (of Gorman & Johnson 2013) on this matter, who has assured me that the greater R programming environment, in which Rbrul is written, is capable of dealing with statistical predictions based on the coefficients in the results presented above. I myself have begun undertaking basic training in R at the University of Essex, thanks to a grant from the Graduate School and the Department of Mathematical Sciences. I intend to implement this knowledge and collaborate with Gorman, who in turn has offered to co-author a future study with me based on these data.

30 Note that in these two tables, coefficients should be read inversely compared to log-odds in all other tables. In other words: a positive coefficient indicates a tendency to pronounce a variant closer to the historical pharyngeal (4 on the 0-4 scale); a negative coefficient indicates favoring of variants closer to deletion (0 on that scale).
Chapter 5: Analysis and discussion: variable 2 – (EMPH)

5.1 Description of the variable

The variable dealt with in this chapter comprises three phonemes of Palestinian Arabic, which traditionally have a primary articulation in the alveolar ridge and a secondary articulation, which has been described by phoneticians and phonologists as “velarized” or “pharyngealized.” See Davis (2009) for a succinct overview of this phenomenon. Note also that other scholars prefer to characterize these consonants as “uvularized” – see Shahin (2002, especially pp. 24-28), who reviews literature on this very matter, as well as studies such as Al-Tamimi & Heselwood (2011), who continue to review the literature, as well as conduct instrumental experiments and find that uvularization and pharyngealization occur interchangeably in, e.g., Jordanian Arabic. Unfortunately, in most cases the imaging technologies they have used have not been very useful in allowing them to visualize the uvulae of their speakers. However, acoustic analysis has allowed them to account for uvularization in an environment of an adjacent vowel /uː/, which was ascertained through observation of the third formant (F3).

5.1.1 Main research questions

The current study is less concerned with the particular articulation of the emphatic variants of the historical emphatics of Palestinian Arabic. What does interest us is very similar to what we tried to observe with respect to variable 1. The questions we pose are therefore as follows:
1. Are the historical “emphatics” of Arabic undergoing a change in progress similar to the voiced pharyngeal?

2. If so, to what extent can this change be attributed to contact with a Hebrew superstrate?

For these two questions, it seems phonologically useful, even if not always phonetically accurate, to call the emphatic consonants “pharyngealized,” as it permits us to treat them as analogous to the full-fledged pharyngeal consonants, the variation in which we have addressed in Chapter 4. Shahin (2002) follows Bessell & Czaykowska-Higgins (1991) and speaks of a natural class of consonant called “postvelars.” These are defined as “sounds articulated wholly or partly in the postvelar region of the vocal tract” (Shahin 2002:18), and as such include the “gutturals” (Shahin’s terminology) –/ʕ h g q ʔ/ (and perhaps /ʔ h/) – wholly articulated postvelarly, and the emphatics, which have “primary non-postvelar, secondary postvelar articulation” (Shahin 2002:19).

As mentioned above, Palestinian Arabic has three traditional emphatic consonants. Some Palestinian dialects have two emphatic stops and a fricative, others have two emphatic fricatives and a stop. Jaffa, being an urban dialect, is of the former type. The emphatic phonemic inventory is therefore as follows:

1. /dˤ/
2. /sˤ/
3. /tˤ/

The observed pronunciation of these traditional phonemes in some cases, however, is what makes them, as a group, to be considered a variable. The
alternate (or in Labovian terms, ‘advanced’) realizations of (1)-(3), respectively, are as seen below in (1′)-(3′):

1′. [d]  
2′. [s]  
3′. [t]

If we forget for a moment that this is still merely a variable, and that for each of the three phonemes both variants exist in the data, we may wish to ask an additional question, namely:

3. Are the emphatics merging with their non-emphatic counterparts?

While it may seem premature to ask such a question prior to even establishing that what we have in front of us is a change in progress, let alone a “done deal,” the diachrony of emphatics – and indeed other postvelar phonemes – both within Arabic and across Semitic languages (specifically the Semitic languages of the Levant: Arabic, Hebrew and Aramaic), virtually begs of us to ask this question at some point or another. And unlike the other instances of phonological processes in these languages, most of which had been signed, sealed and delivered between decades ago (in the case of Modern Hebrew) and over a millennium ago (in the case of Tiberian Hebrew), here we have the opportunity to document and analyze such a change while it is in its variable state. And yes, its very status as an eventual merger is questionable, not only because the change is not yet complete – and for all we know may never reach completion – but also for reasons that will be explained below grounded in theoretical sociohistorical linguistics.
5.1.2 A sociohistorical linguistic approach

The nature of the analysis in this chapter therefore will differ from that of the previous chapter in that it will not be as precise from a quantitative point of view. Rather, it will draw from data collected for this study and contextualize it using methodologies and theoretical notions of both historical linguistics and sociolinguistics. There are a number of variants to the approach known as “sociohistorical (or socio-historical) linguistics” – sometimes referred to as “historical sociolinguistics.” One of the most recent treatises undertaking a linguistic study employing such an approach, Trudgill (2010), devotes very little time to even defining what sociohistorical linguistics is. In the prologue to the book, Trudgill (2010:xii) alludes to a certain “historical-sociolinguistic puzzle” and “historical-sociolinguistic tales of detection.” Labov (1982:21) reflects on an older classic he had co-authored with Weinreich and Herzog (Weinreich, Labov & Herzog 1968), arguing that their proposal was to form an alliance “between dialectology, sociolinguistics and historical linguistics [which] is oriented towards a type of theory that would redress the balance between historical and synchronic explanation.”

For Nevalainen (1996:5) it was not until 1982, when Romaine published a monograph whose main title is Socio-historical linguistics, that a “discipline” by this name was developed, with James Milroy following suit a decade later. It is difficult to dispute this assertion, as neither the term “sociohistorical linguistics” nor the variant preferred by Milroy, “historical sociolinguistics” appear to be used by Labov and his collaborators. But in all fairness, Labov has always
envisaged his work as an overarching intellectual examination of the human language faculty, very similar to the approach taken by generativist linguists. As early as 1972, he wrote: “I have resisted the term sociolinguistics for many years, since it implies that there can be a successful theory of practice which is not social” (Labov 1972:xiii).

5.1.3 What, if anything, is merging?

In Jaffa, there appears to be a lenition of the emphatic consonants so prominent as to raise the question in (3) above as to whether they have merged with their

![Figure 4: Example of emphatic misspelled as non-emphatic by native speaker](image)
non-emphatic counterparts. The degree to which this phenomenon has permeated the speech community can be seen anecdotally through such things as spelling errors. In Modern Hebrew, which has undergone a loss of the emphatics, but has retained the orthographic representation for those historical emphatics that Biblical has retained (more on that later), it is not uncommon for many speakers, be they immigrants, or even native speakers – especially children (see Ravid 2001 for experimental data), but some adults as well – to spell words that traditionally or prescriptively are spelled with a grapheme denoting a historical emphatic consonant (e.g., ̀ for /tˤ/) with the grapheme denoting the non-emphatic counterpart of that historical emphatic consonant (in this example, ֶ for /t/), as they have undergone a complete merger in Modern Hebrew (see Bolozky 1997:287).

Figure 4 illustrates a similar error made by an adult bilingual (Arabic L1, Hebrew L2) in his twenties, who spends his time between Jerusalem where he studies and Jaffa, where he is a political activist. It was posted on the online social network Facebook in the spring of 2013. It reads: {миʃʕаːриːф кіːф kunt аqді wаqт-і w-аnа fі tаріːq-і bi l-бaːs bi-duːn l-аjфoːn!!} ‘dunno how I woulda spent my time while I’m on the bus without the iPhone!!’ The word аqді ‘I spend’ (subjunctive) is spelled in this example with the grapheme ַ, which represents the non-emphatic voiced alveolar stop /d/, whereas in normative Arabic it is pronounced with the emphatic counterpart /dˡ/, represented orthographically asض

The questions of mergers in Arabic – and in Central Semitic – is not a new one. As promised above, we shall return to the fate of emphatics in earlier varieties of Hebrew later in this chapter. Our main goal in this chapter, however, is to
ascertain what sort of process or processes are under way in Palestinian Arabic with respect to these variables. Al-Wer (2004) has analyzed a closely related question in a number of Arabic dialects: whether the two voiced emphatic phonemes attested in Classical Arabic – interdental fricative /ðˤ/ and alveolar stop /dˤ/ – had merged.

Both in that study and subsequently in the encyclopedia entry on phonological mergers (Al-Wer 2008) she concludes that while the plain interdental fricatives /θ/ and /ð/ had merged with their alveolar plosive counterparts /t/ and /d/, respectively, in a good number of urban dialects in North Africa and the Levant, this assertion cannot be made for the emphatic pair, despite the lack of contemporary dialects which distinguish between the two phonetically.

One of the most widely accepted principles applied to phonological mergers is that they are typically irreversible. Al-Wer (2008:605) explains this principle in the following manner: “What this means in practice is that it is conceptually impossible for native speakers to unmerge a merged word class.” The example given thereafter is that of the words [taːni] ‘second’ and [tamir] ‘date (the fruit).’ Al-Wer asserts, quite convincingly, there that is “no rule” (i.e., no linguistically internal mechanism) by which a speaker, let alone a speech community at-large, would be able to group words like [taːni] into one etymological group and words like [tamir] into a different such group, thus reversing the merger back to the historical *θaːni and *tamir.

This, I argue, is not the be all and end all of the argumentation, however. In one of Labov’s first discussions of phonological mergers (Labov 1982) his conclusion that these mergers are irreversible is both non-categorical and accompanied by
a number of interesting questions, which may shed particular light on the case of Arabic. In addition to this, most of Labov’s phonological principles regarding mergers, splits and chain shifts (see Labov 1994) are based on such phenomena involving vowels. One of the reasons Labov himself encourages the study of variation in non-Indo-European languages, such as Arabic, is that they are likely to provide data that would test the principles which he has posited through variables of a different nature (Labov, p.c.). Clearly, the pharyngeal and pharyngealized consonants discussed in this thesis are fairly dissimilar from the vowels of English, Spanish, Yiddish and other IE languages upon which most of these principles are based.

Recall the case of Modern Hebrew with its divergent spelling and convergent phonology. Labov explicitly asks: “can mergers be reversed under the influence of spelling?” (Labov 1982:29). Arguably, Palestinian Arabic differs from Modern Hebrew in that it is not quite the same variety of Arabic people use in literacy, whereas in the Hebrew case, speakers tend to read and write a variety much closer to the vernacular.

A number of arguments can be made in favor of Arabic speakers being able, cognitively, to distinguish, between plain interdentals and their plosive counterparts. Having posited Labov's question regarding the potential influence of spelling on the reversibility of mergers, a question, which he partially answers by citing Garde (1961) who in turn provides examples from Russian to this effect, we should ask ourselves whether there is any spelling convention that differentiates between interdental fricatives and alveolar stops that is relevant to a dialect such as Urban Palestinian Arabic. After all, the Arabic orthography is
primarily a vehicle for reading and writing Standard Arabic (or, as Niloofar Haeri refers to it in her research, “Classical Arabic”).

There is no doubt that in a speech community such as the Palestinian Jaffa community, the prevalent native language is Urban Palestinian Arabic. As such, it is a dialect that lacks interdentals, be they emphatic or plain. Yet, as Haeri (2000:64) quite astutely assesses:

“If we define ‘mother tongue’ as a language that is learned at home without instruction, there is no community of native speakers of Classical Arabic. At the same time, it is the language of Islam, of the state, and of pan-Arab nationalism, and it is explicitly foregrounded as a central marker of ‘Arab’ identity.”

For a community like Jaffa, this statement may need some modification, as “the state” has *de jure* two official languages – Hebrew and Arabic – and *de facto* one primary language, Hebrew, and one secondary language, Arabic (as unfortunate as this may be). It is important to bring this sort of insight into the discussion, while appreciating that we wish not to turn back the wheels of time and regress in our perceptions of language. It is easy to read these lines and accuse their writer of slipping back into the days of reactionary thinking, denying the existence of the vernacular altogether. This is by no means the intention here. Rather, this is an acknowledgment of a subtlety about certain speech communities, which is that even when they use markedly unwritten forms of language, the written forms are not necessarily absent from their awareness.

Haeri tackles this issue from a slightly different angle as well, when she deals with the notion of literacy in the Arab World. She notes that Arabs (as well as other Muslims) may or may not be literate (as individuals, if I read her correctly), but they “conceive of their tradition as a literate and textual one”
This collective literacy is linked, Haeri argues, to the reading of the Qur’an. She cites Wagner & Lotfi (1983), who remind us that Islam’s two monotheistic predecessors, Judaism and Christianity, are also “‘literate’ religion[s],” due to the prominence of their respective sacred texts.

Add to this the following artifact about Palestinian speech communities: only a small subset of these are urban. Many other speakers throughout Palestine speak rural or Bedouin dialects, in which interdental fricatives are retained as separate phonemes, distinct from the alveolar stops (see Table I in Shahin 2008:527). This gives speakers in cities like Jaffa, Jerusalem, Nazareth and Haifa access to interlocutors throughout the country, whose phonemic inventory is richer than their own. And unlike some of the well-studied variables of North American English, for instance, the differences in phonemic inventories among speakers of Palestinian Arabic are well above the level of consciousness. Here are several reasons for this:

1. The country is small, and the number of Arabic speakers within it is limited.
2. The variables in question are consonantal, and therefore quite easily noticeable.
3. Within the Mediterranean Basin, being urban carries a certain degree of prestige, mostly due do the glory associated with such large cities in the region as Cairo, Beirut and Damascus. Rural life, specifically rural speech, sounding falla:hi, ‘rural’, is often ridiculed.
4. The expanded phonemic inventory, which includes interdentals, resembles (but is not identical to) that of Classical Arabic, which, as noted
above, is a variety that the speech community as an aggregate has access to, such that deviations from it are likely to be readily apparent.

Contrast, therefore, the distinction between [d] and [ð] in Arabic to the low-back vowel merger in North American English. The former is much more likely to be in a speaker’s – indeed in a speech community’s – cognition than the latter. Al-Wer, in fact, makes quite a similar statement regarding the Arabic interdentals:

“[I]t is noticeable that the speakers have no problems in re-splitting /t/ into [t] and [θ], and /d/ into [d] and [ð]. The difficulty they show is confined to the distinction between the emphatic variants. I attribute this difference to the fact that in the case of the plain consonants, the phonemic distinction is available in the linguistic experience of the speakers, even if, sometimes, they do not make this distinction phonetically, whereas in the case of the emphatic consonants, there is no such evidence on which speakers can make a phonemic split when required, e.g. speaking or even reading the Standard variety.”

(Al-Wer 2004:28)

5.1.4 From Proto-Semitic to contemporary Palestinian Arabic and Hebrew

In order to illustrate the complexity of the question of mergers in the case of the Jaffa Palestinian Arabic emphatics, I have constructed the two tables below. A complete analysis of the question of contact with Hebrew and its role in the changes in progress discussed in this thesis may, in fact, warrant augmenting these tables with additional rows for pharyngeal, uvular, glottal and velar consonants, as well as additional statistical outputs. I hope to accomplish these tasks in the near future.
In a previous account of the Jaffa Arabic phonemic inventory (Horesh 2000:17), I included the alveolar fricatives: [z], [s], [zˤ] (in the original paper I wrote “z’y”) as possible reflexes for the historical interdentals, commenting that they occur “only in MSA loanwords.” Holes (2004:71) asserts that in urban dialects of Syria, Jordan and Egypt, “educated speakers” in particular use these allophones nowadays, not only in “neologisms imported from MSA,” but “more generally in ‘dialectal’ words also,” in order to sound more “educated” or convey a sense of “correctness”. Al-Batal (2002:101) goes as far as labeling examples with sibilants read on the Lebanese TV channel LBCI as “Lebanese Colloquial,” alongside other examples with alveolar stops. We should have no qualms adding this Palestinian dialect into the mix (see Palva 2006:607-608). What is important to remember is that the productivity of the sibilants as reflexes of the interdentals is limited in urban dialects (such as Jaffa), and is typically confined to specific lexical sets.

Table 10: Emphatic consonants and related phonemes across languages, eras and varieties

<table>
<thead>
<tr>
<th>Proto-Semitic</th>
<th>Classical Arabic</th>
<th>Jaffa Arabic</th>
<th>Biblical Hebrew</th>
<th>Modern Hebrew</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>d</td>
<td>d</td>
<td>/d/ ~ [ð]</td>
<td>d</td>
</tr>
<tr>
<td>δ</td>
<td>δ</td>
<td>d³¹</td>
<td>/d/</td>
<td>/z/</td>
</tr>
<tr>
<td>δ'</td>
<td>δˤ</td>
<td>/dˤ/ ~ [d]</td>
<td>sˤ</td>
<td>ts</td>
</tr>
<tr>
<td>Ʒ</td>
<td>dˤ</td>
<td>/dˤ/ ~ [d]</td>
<td>sˤ</td>
<td>ts</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>ɟ</td>
<td>/s/</td>
<td>/ʃ/</td>
<td>/s/</td>
<td>/ʃ/</td>
</tr>
<tr>
<td>ʃ</td>
<td>s</td>
<td>s</td>
<td>ʃ</td>
<td>ʃ</td>
</tr>
<tr>
<td>θ</td>
<td>θ</td>
<td>t</td>
<td>ʃ</td>
<td>ʃ</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>/t/ ~ [θ]</td>
<td>t</td>
</tr>
<tr>
<td>t'</td>
<td>tˤ</td>
<td>/tˤ/ ~ [t]</td>
<td>tˤ</td>
<td>t</td>
</tr>
</tbody>
</table>

3¹ In a previous account of the Jaffa Arabic phonemic inventory (Horesh 2000:17), I included the alveolar fricatives: [z], [s], [zˤ] (in the original paper I wrote “z’y”) as possible reflexes for the historical interdentals, commenting that they occur “only in MSA loanwords.” Holes (2004:71) asserts that in urban dialects of Syria, Jordan and Egypt, “educated speakers” in particular use these allophones nowadays, not only in “neologisms imported from MSA,” but “more generally in ‘dialectal’ words also,” in order to sound more “educated” or convey a sense of “correctness”. Al-Batal (2002:101) goes as far as labeling examples with sibilants read on the Lebanese TV channel LBCI as “Lebanese Colloquial,” alongside other examples with alveolar stops. We should have no qualms adding this Palestinian dialect into the mix (see Palva 2006:607-608). What is important to remember is that the productivity of the sibilants as reflexes of the interdentals is limited in urban dialects (such as Jaffa), and is typically confined to specific lexical sets.
### Phonetic realization | Possible correlates | phonemic | Hebrew equivalents
--- | --- | --- | ---
[d] | /d/ | /d/ | /d/  
 | /ð/ > /d/ | /d/ | /zd/  
 | /dˤ/ | /ts/ | [d]  
 [s] | /s/ | /ʃ/ | [s]  
 | /sˤ/ | /ts/ | [s]  
 [t] | /t/ | /t/ | /t/  
 | /θ/ > /t/ | /ʃ/ | [t]  
 | /tˤ/ | /t/ | /t/  

**Table 11: From Jaffa Arabic phonetics to Arabic phonemics and Hebrew equivalencies**

The columns in Table 10 are fairly self-explanatory. Unless otherwise indicated, the sounds in the table are understood to be phonemes. If a column is split, this is an indication that a phoneme in one language or variety has split in one or more languages or varieties into two phonemes. In cases of allophonic variation, this is indicated as: /x/ ~ [y], with the understanding that [x] is also an allophone of /x/.

There is a difference between the kind of allophonic variation referred to in the Jaffa Arabic column as opposed to the Biblical Hebrew column. The variation in Jaffa is of the type described throughout this thesis, and is subject to a series of variable rules, not all of which we have managed to figure out quite yet. In the case of Biblical Hebrew, the interdental variants of /d/ and /t/ are the well-attested “spirantized” allophones thereof, which are in the vast majority of cases phonologically conditioned (see, e.g., Khan 1997:86; 89-90).

Note that Modern Hebrew is treated here, unfairly, as a monolith. Bolozky (1997:287) and others rightfully distinguish between the phonologies of
Standard (or General) Israeli (or Modern) Hebrew and various, more conservative dialects of contemporary Hebrew, which preserve a larger portion of the original Semitic phonemic inventory. Bolozky calls this “‘Arabicized’ Hebrew,” and with good reason – it is typically spoken by native speakers of Arabic (including Jews who had immigrated to Palestine from elsewhere in the Middle East or North Africa) and their native Hebrew-speaking offspring. Yet given the paucity of variationist sociolinguistic studies (a handful are currently underway) on dialects of Modern Hebrew, I am basing this analogy on the variety of contemporary Hebrew which is best documented, and probably has the most influence on the speech of Palestinian speakers of Arabic anyway.

What I find to be interesting – and relevant to the question at hand – is that within this analogy across varieties of Arabic and Hebrew – while keeping the history of Semitic languages at-large in mind – there is an incongruence between synchrony and diachrony. In a way, it may very well resemble the conundrum raised by Al-Wer in her critique of the supposition that the voiced emphatic alveolar stop and voiced emphatic interdental fricative of Old Arabic have merged in virtually all of the contemporary dialects known to us, such that in any given dialect, there is either a /ðˤ/-type phoneme or a /dˤ/-type phoneme, but never both (Al-Wer 2004).

Al-Wer contrasts the surface synchronic data regarding the lack of one phonetic realization or another in the case of the voiced emphatics with that of the plain phonemes in Jordanian and Moroccan Arabic; and she also introduces historical input from Steiner (1976) and others to the effect suggesting that it is quite
plausible that in Arabic, such a distinction never really existed, therefore any attempt to discuss a “merger” of the voiced emphatics would be moot.

The current case has some similarities to Al-Wer’s thesis and some attributes for which it is unique. What I wish to focus on is the following. On the one hand, phonetically, the Jaffa Arabic emphatic consonants are variably gaining allophones that are non-emphatic, and which exist in both Palestinian Arabic and Modern Hebrew. However, the historical development of these phonemes in Hebrew has not been precisely parallel to that of their counterparts in Arabic. What the various degrees of shading in the two tables above intend to illustrate are the cases where Arabic and Hebrew phonemes do in fact have similar realizations in contemporary Palestine – even if dissimilar allophones exist as well. Table 11 distills these similar phonetic outcomes and tracks their etymological origins in both languages. What we learn is that only in the case of the voiceless emphatic alveolar stop /tˤ/ is there both a diachronic and a synchronic rationale for merging the emphatic stop with the plain stop, assuming the motivation is contact with Hebrew, especially given that in both languages this merger would further complicate a merger already attested to of an interdental fricative with an alveolar stop (though in Hebrew, this is more complicated).
5.1.5 Synchronic or diachronic? Internal or external?

In both other cases, the Hebrew factor would only be viable as a synchronic, phonetic, on-the-surface justification for merging the Arabic emphatics. At this juncture, it would be premature to insist that this change is a direct result of contact with Hebrew. Thomason & Kaufman (1988:57-64) discuss at length the question: “When Is an External Explanation Appropriate?” They note that traditionally, historical linguistics prefer to explain changes language-internally before turning to “external causation,” e.g., contact-induced change. However, they reject the rather dogmatic approaches invoked by such scholars as Martinet, Polomé and Ohala (cited in Thomason & Kaufman 1988:57-58), who only permit resorting to external explanations (Ohala 1974:268, for instance, mentions “social, psychological, or historical facts”) if internal explanations fail. They agree with Ohala that when explaining sound change phonetic explanations should take precedence, but they add that the analysis should be “as complete as possible” (Thomason & Kaufman 1988:58).

I wish I had a better way to convey the conclusion drawn in the following paragraph than to quote it en masse. But it is so succinct and elegant, that I feel as if I have no choice:

"We need a methodological criterion that matches better with theoretical considerations. Here it is. As with the establishment of genetic relationship, a successful criterion for establishing external causation is possible only when we consider a language as a complex whole—a system of systems, of interrelated lexical, phonological, morphosyntactic, and semantic structures. Instead of looking at each subsystem separately, we need to look at the whole language. If a language has undergone structural interference in one subsystem, then it will have undergone structural interference in others as well, from the same source. Not necessarily in all other subsystems: as we have argued above, lexical interference may be negligible in cases of
interference through shift; and considerable structural interference may occur without including externally motivated changes in the inflectional morphology. But we have found no cases of completely isolated structural interference in just one linguistic subsystem.”

(Thomason & Kaufman 1988:60)

The implication of Thomason & Kaufman’s generalization to the Arabic case is quite clear. We should first seek explanations for the (variable) loss of pharyngealization of the emphatic consonants, which are internal to Arabic; since this is a sound change, if we can find explanations grounded in phonetics, such explanations would be preferable; however, given our overall knowledge of the sociohistorical context in which the speech community under investigation has been evolving, as well as other linguistic changes, for which there is evidence— with varying degrees of robustness—of contact being the prime motivation, it is my view that the initial hypothesis (i.e., contact with Modern Hebrew, which lacks emphatics, has induced loss of emphatics in Palestinian Arabic) must be entertained, despite the partial phonemic incongruence from a historical perspective.

5.2 A diversion: Other contact-induced changes in Palestinian Arabic

What are some of these additional linguistic changes, which Palestinian Arabic in speech communities such as Jaffa is undergoing, and which contact with Modern Hebrew is a factor in initiating and/or promoting?
1. **In Phonology:**

   a. The (ʕ) variable: This has been shown in Chapter 4 above to be a change in progress in more than one variety of Palestinian Arabic, including among speakers who have little or no contact with Hebrew. However, quantitative analysis has revealed that both within the Jaffa community and across the Jaffa—West Bank communities, contact with Hebrew (in its various instantiations) is statistically significant as a factor favoring the various non-pharyngeal variants of the /ʕ/ phoneme of Arabic.

   b. Loss of phonemic length: This feature of Palestinian Arabic has yet to be studied from a variationist perspective. We have sporadic evidence from the corpus collected for this thesis that in many cases, vowels that are contrastively long in other varieties of Arabic (including other Levantine dialects) are significantly shortened in this Palestinian dialect that is in contact with Hebrew. While Ancient Hebrew (defined by Rendsburg 1997 as the Hebrew attested between ca. 1100 BCE and ca. 250 CE) is presumed to have distinguished between long and short—and even extra-short (schwa-like)—vowels (see Rendsburg 1997:76-79), the variety of Hebrew which had recorded in writing the Jewish tradition of orating the Old Testament, known as Tiberian Hebrew (7th-9th centuries CE) appears to have only maintained long vowels as allophones of short vowel phonemes, which in turn had developed distinctions that were in vowel quality rather than quantity. See Khan (1997:91-100) for a detailed account of
Tiberian Hebrew vowels. Modern Hebrew has no phonemic vowel length distinctions (Bolozy 1997:288). There is some evidence that phonologically conditioned long allophones of vowels are once again emerging in Modern Hebrew. Bolozy (2006) attributes this phenomenon mostly to younger speakers of North African descent, which, if true, would call for an interesting study of the mutual influences of Arabic and Hebrew on one another. Such studies, however, must not fall into the trap of a false symmetry, as the two languages play very different roles both psycholinguistically and socio-politically in their speakers’ lives.

c. Loss of gemination: This is yet another feature of Palestinian Arabic, which is emerging, and is a prime candidate for being a product, perhaps partially, of contact with Hebrew. Modern Hebrew “no longer maintains gemination, except phonetically, across morpheme boundary” (Bolozy 1997:288; one example he provides is jafan+nu, ‘we slept’). Note that both Ancient and Tiberian Hebrew varieties maintained gemination, except in the case of laryngeal and pharyngeal consonants and the phoneme /r/. Both Rendsburg (1997:74) and Khan (1997:90) describe these exceptions for each of these historic varieties, respectively, as an inability to geminate. Khan writes: “Some consonants could not be geminated. These included the laryngeals (/ʕ/, /h/) and pharyngeals (/ʕ/, /h/) and also /r/, except in a few isolated cases.”

2. Morphosyntax—the genitive exponent fe:ṭ: It is well known that Arabic dialects have developed a structure largely unavailable in Classical Arabic
to express possession in an analytic structure. While the synthetic “construct state” is used as well, genitive exponents (sometimes referred to as “possessive markers,” “possessive pronouns,” “prepositions,” or simply, as by Holes 2004:208, “particles”) often replace them in most Arabic vernaculars.

In Palestinian Arabic, the most widespread such particle is probably tabaf (see Shahin 2008:535, 537; Rosenhouse 2007:489), but šeːt is also attested in Jerusalem (Levin 1994:209-210), and in the corpus I have collected is the most frequently used such particle for most speakers.32 That this analytic structure is rampant in virtually all contemporary dialects and that this particular morpheme is documented in at least one other Palestinian community with significantly less contact with Hebrew appear to run against any hypothesis that this feature would have anything at all to do with contact. But there is much more.

We will soon consider the parallel structures in Hebrew, but prior to doing so, it would be useful to explore whether there are any parameters or constraints for preferring analytic possession over synthetic possession—in Palestinian Arabic and in Arabic dialects in general.

Holes (2004:208) first posits a semantic preference—and this is for Arabic dialects at-large—namely that “[g]enerally speaking, this construction [i.e., the analytic one] is only used to express alienable possession; inalienable possession is expressed by the construct state, as

32 Incidentally, Versteegh (2001:107) notes that the etymologically cognate fajt is the genitive exponent typical of Cypriot Arabic.
in MSA.” Shahin (2008:537) makes a cursory assertion to that effect regarding *tabaṭ* in Palestinian Arabic as well.

Holes (2004:209) also adds that analytic dialectal structures may be used in dialectal forms as alternatives to “noun + enclitic pronoun structure.” Following this introduction, Holes enumerates the “types of noun phrase” that appear to favor such analytic construction over the traditional synthetic ones. Rosenhouse (2007:489) and Levin (2004:209-210) offer somewhat similar phonological and prosodic explanations for triggering the use of *tabaṭ* in Jerusalem Arabic.

What precisely constitutes alienable and inalienable possession is by no means a clear-cut distinction. From a formal semantic perspective, it may also be the case, as in some Austronesian languages, that alienability is not a binary category (see Chappell & McGregor 1989:27). There is interesting work in formal semantics analyzing synthetic and analytic (or periphrastic) possessive constructions in Maltese (e.g., Gatt 2004, Fabri 1996, Koptjevskaja-Tamm 1996). Studies in two North African dialects: Boumans (2006) on Moroccan Arabic in Morocco and the Netherlands and Sayahi (2011a; 2011b) on Tunisian Arabic, offer a new dimension, that of language contact, into the analysis of possessive constructions and the distributions of synthetic versus analytic possessive phrases. Boumans has shown that statistically, Moroccan speakers who were raised in the Netherlands in a bilingual Dutch-Arabic environment had a higher propensity for using genitive exponents (e.g., *djal*) than the construct state. At least in some constructions, this could have been
attributed to a calquing of the Dutch particle *van*. Within Morocco, speakers in Tangier differed significantly from speakers in Casablanca, Rabat and Oujda. While it is not certain, one hypothesis Boumans raises for this is that the circumstances that led to the formation of the Tangier dialect in early Medieval times involved much more contact with Amazigh and perhaps Romance speakers than in the other three cities. In Tunisia, Sayahi notes, “the use of the genitive exponent *mtɛːʕ* to express possession and the use of pre-verbal markers to indicate the future, as part of an overall move towards more analytical forms” (Sayahi 2011a:2) are distinctive features of the local dialect. Furthermore, when Arabic-French code-switching is considered, the analytic construction with *mtɛːʕ* has been found to occur 64% “with a French word either as the possessor or the possessee” (Sayahi 2011b:131).

Back to Modern Hebrew: In the first book-length complete grammar of Modern Hebrew to be published for a general linguistics readership in English (there had been a number of similar publications in Hebrew beforehand), Berman (1978:231-276) devotes a hefty chapter to “construct state genitives” and recognizes “three surface structures,” all of which, in her view, “are essentially synonymous, and that the choice between them is a stylistic matter alone, with certain important exceptions” (Berman 1978:232). Berman concludes that “everyday colloquial Hebrew usage will today prefer one of the two forms that use the genitive particle *fel*, and the bound form with no particle will be
confined to more formal, literary style, with certain quite limited exceptions.”

Subsequent Modern Hebrew grammars, by Glinert (1989:24-49), and Coffin & Bolozy (2005: e.g., 169) are in general agreement that “formal” Modern Hebrew registers tend to use synthetic possessive phrases, whereas “casual” (Glinert) or “common/colloquial” (Coffin & Bolozy) prefer the analytic forms with _fel_. It is worth mentioning that an even newer form has emerged in recent years in Modern Hebrew, for _fel_+pronominal suffix in a phonologically reduced form, such that, e.g., _fe′li_, ‘my’ is sometimes realized as _fli_; _fe′lax_, ‘your (f.sg.)’ as _flax_; _fel′xa_, ‘your (m.sg.)’ as _fxa_, etc. Cohen (2003) treats these reduced forms as “clitics.” Whether this is a regression towards a synthetic construction is unclear, but may be worth investigating.

At any rate, what we have in Jaffa are speakers whose native Arabic dialect has a genitive exponent _fe:t_ and whose second language, Modern Hebrew, has a genitive exponent _fel_. And while it is clear that the two are not etymologically related (see Pat-El 2010, 2013 for the history of analytic possessives in Hebrew), the phonetic similarity, and the much more widespread use of the analytic forms in Modern Hebrew than in any documented dialect of Arabic—with the exception perhaps of Maltese and the other cases mentioned above, where contact was a clear factor driving out the synthetic construct state—there is sufficient evidence to hypothesize that this, too, is a contact-induced change. The next step would have to be a quantitative one.
3. Code-switching: Almost inevitably, some speakers in a place like Jaffa begin their sentences in Arabic, *y terminan en hebreo*, to paraphrase the title of Poplack’s 1980 article on English—Spanish code-switching among Puerto Ricans in New York City. One speaker, ‘nevi:n,’ was actually so prone to code-switching, that at times it was difficult to ascertain whether she was speaking Arabic as the matrix language and Hebrew as the embedded language or vice versa. On this matter it can be useful (though at times it is equally confusing) to consult such debates on these issues as between Myers-Scotton (2001), who is the originator of the Matrix Language Frame model (Myers-Scotton 1993), and Bentahila & Davies (1998). The latter present data in which Moroccan Arabic and French are used interchangeably, often in such a manner that renders it quite difficult not only to identify which is the matrix language and which is the embedded language, but also *whether* these categories even apply to this brand of “mixed discourse.”
Consider the following sentence uttered by ‘nevi:n’ in Jaffa:

\[
\text{fi: ktir illi b?ulu:-li ana lo tsodeket, ze lo}
\]

EX many REL IND-tell-me I NEG right, DEM NEG

\[
naxon lax?ov kaxa, aval hada al-?ani ha-pnimi jej-t-i.
\]

right to-think this-way but DEM DEF-I DEF-internal POSS-1S

‘There are many [people] who tell me I’m not right, it’s not right to think like that, but that’s (the) my internal “me”.

In order to capture the complexity of the mixed discourse, I have presented the Hebrew words above in italics, be it in the transcription, in the word-for-word glosses, or in the final translation. The genitive exponent (discussed earlier in this chapter) is presented in boldface.

Interestingly, part of the debate between Myers-Scotton and Bentahila & Davies revolves around the function of the Moroccan Arabic genitive exponent *djal* in the discourse.

Since it is not the primary goal of this section to analyze cases of code-switching, but rather to contextualize the sound changes that we are analyzing and provide support for there being a plethora of features in Palestinian Arabic influenced by contact with Hebrew, I will leave the detailed analysis of the code-switching phenomena for a separate study.

### 5.3 Back to emphatics

Let us return to the emphatic consonants and provide some examples from the data of their depharyngelization. The following passage is transcribed from an interview with ‘ʒami:l’ (gay male artist/student, born 1982). In the transcriptions henceforth, historically emphatic consonants that are
depharyngealized in the subject’s speech are shaded, and words uttered in Hebrew are printed in italics (the italics are replicated in the translations as well).

ba.araʃ-j bi-zabet lej jani bas hij wiʃlat il-muʃtamaʃ il-jafawi wiʃil la- la-daraʃe nno baʃal ji?amin bi-t-tayji:r. jani:. fiʃ fi jani: e:m. kaman il-e kul il-mumasil-in ʃet-nu-na bi-l-e-, b-il-bal-, b-il-baladija w-ʔaʃja min ha-n-noːʃ illi hume bisaːʔidu jxl il-wadeʔ hek had. iza minʃuf matlan inno kaʔilu bisawu baʃ-jafa: haj minsawi la-jafa e fi firʔat kaduregel ʃet ʔarab u-jahud, fi ʕanna firʔet baʃraf-eʃ e ʃu. sawu ʒanajen iʒdiːde, fi anna maraʒiːh hoːn u-ɣard. ana bitʃuf ئa haj nno mafkiʃim b-jaʃfa. minzʔabbit, minsawi ʃawaːriʃ, min- min-minsawi-ha jaʃni ahsan. bass e tisʔal e nafs-ak haj il-e ʒanajin w-il-e kaduregel bitʃib ʃuʃel la-ḥada? e: bitsʔalleh il- il-wade ʃet is-sakan u-t-taʃliːm? kul il-maʃakil il-ʔasasije bisibu-ha ʃala ʃaʔʔa ubisaːwu feʃing jaʃni.

‘I don’t exactly know why, like, but it’s come to the point, the Jaffa community has reached a point where it no longer believes in change. Like, there’s nothing, like, um. Also, the, all of our delegates in the, in the Municipality and things like that, which help keep the situation as it is. If we look, for example, as if they do [things] in Jaffa. We do for Jaffa, um, there’s a football team of Arabs and Jews, we have a team of I dunno what. They’ve made new gardens, we have new seesaws here and there. I, you see that like, they invest in Jaffa. We fix, we pave roads, we, we, we make it, like, better. But, um, ask yourself: those new gardens and the football – do they provide jobs for anyone? Um, do they fix the situation of housing and education? All of the fundamental problems – they dump them on one apartment [building] and, like, create a façade.’

While ‘ʒami:l’s’ speech was rather clear and crisp and conducive to dichotomous distinctions between phonetically plain and emphatic consonants, this was not always the case. ‘sa:lim,’ a municipal worker, born 1948, had something of a lisp (as did ‘nevi:n’ and several other speakers), rendering some auditory coding of this variable—especially for the sibilants—quite challenging. To illustrate this,
following is a transcribed passage from ‘sa:lim’s’ interview, followed by its translation and seven figures produced by Praat (Figures 2 through 8), depicting the sound waves for the sibilant and its intensity in dB. (Short utterances in parentheses are my interjections and questions.)


‘Let’s start from the south. From the south, we call [that area] “the Bat-Yam boundary”, “the Bat-Yam boundary”.

(bass bass bitsammu:- jaʃni ʔyul bat jəm?)

(‘But, but, you call it, I mean, “the Bat-Yam boundary?”’)

ʔa: binsmami: ʔyul bat jəm. hecka.

‘Yep, we call it “the Bat-Yam boundary”. Just like that.’

(bi-l-ʃablən?)

(‘In Hebrew?’)

bi-l-ʃablən, ʔa; bi-l-ʃabrən. xamis mit miter joter hala, ʒəmb bijaaret dakke, ʒəmb bija:ret dakke. ze jaʃni hada l-haki hada ʔahl il-balad kull-ha, bidd-ak tiwṣ ʔaʃi, btiwṣ ʔaf hecka. xamis mit miter ʔəbel, haju ʔə-ʒəbalijje. xamis mit miter ʔəbel, haju ʔan təməɾ. diɾ bacl-ak. ʔa, diɾ bacl-ak. xamis mit miter ʔəbel, bass, jaʃni, ḥkajt bass xamis mit; teritorija xamis mit miter, ʔəbel, haju ʔə-ʒəμi. haju ʔə-ʒəμi haju ʔəkbar ḥajj. ḥada ḥajj il-ʔaʃəμi jabtadi? min ʕənd il-ʔan təməɾ, li-ya::jət, ja ʃəd-i ʔaʃi:z, il-balad il-ʔədəɾəm. ḥada kull-o, la-ʃənd il-balad il-ʔədəɾəm, fəɾ a-ʃaṭika jəməni, biʔul il-ʃəɾ a-ʃaṭika. bi-b-balad il-ʔədəɾəm. ʒəmb is-ʃəfə, ḥajj is-ʃəfə. ʒəɾət irufalajim, ḥajj in-nizha. […] ḥajj in-nuzha, kull-ha, min ʕənd, min ʕənd geulim, min ʕənd kupat xalim geulim, il-arbaʔesre komot, be-bat jəm, baʔed-ha minʔul bat jəm. min ʕənd-ha la-ʔe::n busʔi:rəs. La-ʃənd e: ʒəɾo:-, la-ʃənd kolnoʔa noga. ḥajj

33 Bat Yam is the Israeli town located immediately south of Jaffa, on the Mediterranean coast. It is populated predominantly by Jews, and is considered to have a lower-middle class character. While Jaffa is officially part of Tel Aviv-Jaffa, a socioeconomically diverse city, usually associated in the public opinion with its more affluent neighborhoods, Jaffa residents often share experiences, such as shopping and medical services with Bat Yam, due to its proximity and their shared underprivileged populations. See Cohen (2013) for a brief history of Bat Yam and a sociological analysis of its role as a stigmatized town, including both tensions with Jaffa and similar trajectories in recent decades.
In Hebrew, yeah, in Hebrew. 500 meters *further ahead*, near the Dakke citrus grove, near the Dakke citrus grove. *That’s*, like, that’s the way we talk, that’s all the people in town, you wanna describe something, that’s how you describe it. 500 meters beforehand, the Jabaliyye Neighborhood. 500 meters beforehand, the *Tamar Garden* Neighborhood. Watch out! Yeah, watch out! Just 500 meters beforehand. I mean, it’s a matter of just 500 [meters], a territory of 500 meters, beforehand, the ‘Ajami Neighborhood. The ‘Ajami Neighborhood is the largest neighborhood. This neighborhood of ‘Ajami begins at *Tamar Garden*, all the way, my dear sir, to the Old City, the Old City that is, we say the *Old (the) City*. In the Old City. Near the Clock [Tower]—the Clock Neighborhood, *Jerusalem Boulevard*—the Nuzha [the Promenade] Neighborhood. [...] the Nuzha Neighborhood, all of it, from, from *Geulim*, from *Geulim Medical Center*, the fourteen stories, in Bat-Yam, after that we say Bat-Yam. From there to Bustrus, to, um, Bouleva-, to Cinema Noga. The Nuzha Neighborhood. After Cinema Noga, al-Manshiyya, we call it. Um, north, north of Cinema Noga, the Shlush Neighborhood, the Shlush Neighborhood. *Neve Tsedek*. To this day, all these definitions, a native of the town, a native of the town will recognize. Just tell him the neighborhood, [tell him] what’s nearby, and he’ll tell you where. Um, by *Maccabee Jaffa* [football] pitch, beyond al-Manshiyya, you say “by Maccabee Jaffa pitch”. That’s the name of that area. There used to be a big citrus grove there, a big grove, back in the day when there were groves here. It was owned by someone named Barakat. We, the older ones, the older generation, we call it “by Barak[at’s] grove”, [but] we know [it also as] “by Maccabee Jaffa”. You get it?

---

34 Maccabee Jaffa is a now-obsolete football (soccer) team, which used to be popular especially among the town’s Jewish Bulgarian immigrant community. The team played for several decades in the Israeli Premier League, but it dissolved in 2000 due to financial difficulties. The stadium mentioned in the interview was demolished in 2009, some five years after the interview took place (http://www.nrg.co.il/online/54/ART2/349/273.html – accessed 26 January 2014).
5.3.1 Phonetic cues (or the lack thereof)

What I had hoped to discover by generating the figures in Appendix B was whether there were any acoustic distinctions and/or similarities, which I was not able to detect auditorily. A survey of the literature on emphatics, as well as personal communication with a number of phoneticians and phonologists, some of whom are more intimately knowledgeable about the sounds of Arabic and Semitic languages, others have worked extensively on language variation\(^{35}\) have taught me two things I actually thought I already knew. The first was that acoustic studies of pharyngealized consonants rarely, if ever, focused on the consonants themselves, but rather on the formants of the vowels adjacent to them. For example, Laufer & Baer (1988:195) conclude: “Observed constriction of the pharynx was always accompanied by lowering of F2 and raising of F1, which were clearly evident in the spectrograms.” Secondly, if one really wishes to determine the place and manner of articulation of these complex segments, acoustic phonetics—be it auditory or instrumental (i.e., using a computer program such as Praat)—is typically not considered to be adequate. Laufer & Baer's (1988) use of spectrographic evidence was merely complementary to images from x-rays (nowadays no longer used due to ethical considerations), fiberscope and endoscopic images. Similarly, Al-Tamimi & Heselwood (2011) use Nasoendoscopic and videofluoroscopic imaging to analyze the precise quality of emphatic (and plain) coronals in Jordanian Arabic; Hassan & Esling (2011) use a fiber-optic nasal laryngoscope for emphatics in Iraqi Arabic – again

\(^{35}\) Special thanks go to Jalal-eddin Al-Tamimi, Paul Foulkes, Ghada Khattab, Nancy Kula and Bert Vaux. I do, of course, take responsibility for any conclusions I have drawn from their advice.
the latter two study employ acoustic measures as a means to supplement the experimental imaging.

It may very well be that in order to answer the question posed toward the beginning of this chapter, namely – are the emphatics of Palestinian Arabic merging with their plain counterparts – we would have to resort to at least some of these techniques used by the scholars cited above. While on the one hand it is extremely appealing to try to isolate, almost distill, the individual segment and compare various instances of it either auditorily or numerically using images as in Figures 2-8, a number of problems arise when doing so. First of all, we are not quite sure what to look for. It seems as if in Figures 2-3 the wave form for the sibilant (in the center of the figure, between the two vertical dotted lines) is somewhat more dense than, for instance, in Figures 4, 7 and 8, consistent with my own auditory impression that these first two figures represent phonetic emphatics, whereas the other three do not. However, note that Figure 6, which represents a geminate plain sibilant, is also quite “dense.” Does this mean that emphatic sibilants and plain geminate sibilants are merging? Or that they are phonetically equivalent to begin with? Or perhaps that we should be looking at something else altogether? And how should we deal with amplitude (or intensity, as Praat labels it)?

Add to this the fact that we know that “emphasis” in Arabic spreads across the phonological word (see, e.g., Davis 1993, Davis 1995, Watson 1999, Davis 2009). It is therefore plausible that some contrasts may be preserved suprasegmentally, even if the segmental distinctions have been lost, especially if these losses are variable and irregular.
5.4 An alternative (albeit interim) conclusion

Before we decide that we cannot decide whether a merger is occurring in Jaffa Palestinian Arabic with respect to the emphatic and their plain counterparts, let us consider one more possibility. Labov (1982:56-59; 1994:349-370) discusses the concept of "near-mergers" in quite some detail. *Prima facie*, (one of) Labov’s definition(s) of this phenomenon appears to be precisely what we have been grappling with in this chapter: "It arises when, as the result of sound change, two word classes that are quite distinct in some dialects come into close approximation in a given dialect" (Labov 1994:350). Much of the rest of Labov’s (1994) Chapter 12 on near-mergers is devoted—in addition to providing examples and some theoretical foundations—to offering some tools for empirical testing of the existence of such near-mergers. These tests are very different from the kinds of tests cited above, which were employed by the phoneticians who have attempted to determine the exact place of secondary articulation of the emphatics. Rather these are diagnostics for phonemic contrast between minimal pairs, psycholinguistic intuition quizzes, and so forth, designed to “set aside” (p. 351) much of the phonetic matter in favor of a deeper phonological analysis. This sounds almost counterintuitive coming from someone like Labov, for whom precision is of utmost importance. But we must remember that scholars such as Labov, who rely heavily on precise measurements, do so in relatively large numbers, which add up to formulate generalizations about language and the speech communities who use it. In this light, we should not be surprised that Labov finds himself confronted by die-
hard phoneticians, who find his near-merger proposition too imprecise (pp. 367-370).

I would like to end this chapter with a statement, a conclusion of sorts, as I have thus far asked many a question, raised many a doubt, introduced bits and pieces of data, and I would not like to sum it all up by beating around the bush. The truth of the matter is that I do not know for a fact whether the emphatics in the Palestinian dialect of Arabic of Jaffa are heading toward a merger with their plain counterparts. The data really are inconclusive. However, there is a good deal of evidence—some of it may be considered by legal standards circumstantial evidence—that younger speakers in particular are going precisely in the direction outlined in Labov’s definition of near-mergers. In fact, his definition is ever so elastic, that I am almost willing to sign my name to it.

Here is an annotated version of the definition: two word classes that are quite distinct in some dialects [e.g., emphatics vs. plain coronals in dialects of Arabic that are not in contact with Hebrew] come into close approximation [i.e., speakers do not always distinguish between them phonetically and are unsure regarding their phonemic contrast] in a given dialect [e.g., Palestinian Arabic in bilingual speech communities that are in close contact with speakers of Hebrew].

This does need some substantiation, and within the variationist paradigm this means having the kind of robust quantitative support for this argument as we had for the lenition of the voiced pharyngeal fricative. But the efforts in future research should probably be placed on determining speakers’ perceptions rather than fine-tuning our understanding of their phonetic realizations of these segments.
Appendix B: Praat images for plain and emphatic sibilants

Figure 5: Sound wave and intensity contour for tiw[sʰ]af (1)

Figure 6: Sound wave and intensity contour for tiw[sʰ]af (2)
Figure 7: Sound wave and intensity contour for [s]:di

Figure 8: Sound wave and intensity contour for [s]:a (1)
Figure 9: Sound wave and intensity contour for [ss]eːa (2)

Figure 10: Sound wave and intensity contour for min[s]ammiː-ha
Figure 11: Sound wave and intensity contour for [s]:hib-ha
Chapter 6: Conclusions, implications and future directions

6.1 When things get complicated

Upon initially embarking on this research project, I thought things would be simple: I had a linguistic hypothesis, which was grounded in some social and historical considerations; I would go to the “field,” which was my town of birth (though I had never actually lived there, but rather in its suburbs, where I had grown up speaking a language other than the main language I was to investigate); I would record interviews; I’d run some statistics; my hypothesis would be hopefully validated. Anyone but a novice can realize how naïve so much of this “planning” was.

6.2 Form and content

I began my fieldwork shortly after Niloofar Haeri published her second book, Sacred language, ordinary people (Haeri 2003). It was a great departure from her first, The sociolinguistic market of Cairo (Haeri 1996), which was based on her University of Pennsylvania doctoral dissertation that she had written under the supervision of Bill Labov. I am mentioning this, because Haeri (1996) was my initial role model for a sociolinguistic study: pick a select group of “interesting” variables; run them through the statistical tests available at the time of your writing; make sure to include both “conventional” factor groups and ones that are unique to the speech community that you are investigating; and augment your quantitative analysis with insightful comments based on your intellectual and personal familiarity with the community. In a way, this description does not
do justice to Haeri's work, because it makes it sound dull and formulaic, which it is not. Yet Haeri herself, in the years subsequent to publishing *Sociolinguistic market* has changed gears. She has not only begun teaching and conducting research under the auspices of an anthropology department (at Johns Hopkins University, which she now chairs), but has truly become an anthropologist.

I have not become an anthropologist, nor is my work that of an anthropologist. In one of my first classes in sociolinguistics at the University of Pennsylvania, Gillian Sankoff told us that there were many definitions of sociolinguistics, and that “here at Penn we teach sociolinguistics within linguistics.” Sankoff herself, incidentally, has a doctorate in anthropology from McGill University. Despite my consistent groundedness in sociolinguistics as a sub-discipline of linguistics itself, I have found myself in the latter stages of writing this thesis thinking about issues that are beyond the rather narrow scope of language. While much of my analysis may resemble that of Haeri (1996), my thought process has often led me to drift across into the mindset of Haeri (2003).

2014 marks the thirtieth anniversary of my commencement of studying Arabic. In recent years I have been learning the language not by taking classes, but through research and as a teacher of the language in American institutions of higher education. As a result of my continuous enhanced proficiency in Arabic, each time I listen to the interviews I had conducted, now nearly a decade ago, for the purpose of this thesis, I understand more of their content. I also care more about their content. I wish not to make some sort of unsubstantiated statement regarding the reasons for my relatively new interest in the content of the interviews as opposed to my initial interest predominantly in the linguistic
forms embedded in them. Let us just accept it as a fact. I do feel, however, that this added interest has changed—for the better—my analysis of the data, even the portions that included quantitative analysis. It will also enable me to further my research using the same corpus and create additional, comparable corpora, with social, political and historical sensibilities in mind.

6.3 So what?

When I speak about my research, be it in public fora or in private settings, many people quite readily identify the same sensibilities and implications outside of linguistics for this kind of research. Many others ask me, “so what?” There are various, I think quite important, ways in which sociolinguistic research can have an impact on tangible matters within the community. I will not delve into these here, but merely mention that two of the pioneers of American sociolinguistics, cited elsewhere in this thesis, William Labov and Walt Wolfram, have contributed immensely to the field of language education, through creating and increasing awareness in the communities in which they’ve conducted their research (and indeed beyond) that language variation is a natural feature of the human condition.

Labov (2010:21) illustrates both verbally and graphically how poverty and dialect (in his example, African-American Vernacular English), compounded with residential segregation and institutionalized racism, lead to a vicious circle of more racism and decreased levels of literacy among black children in American cities. While most linguists are aware of Labov’s monumental
contributions in theory and methodology pertaining to language variation and change, fewer are aware of his engagement in the community, where he and his collaborators have capitulated on years of analysis of data to create reading manuals and other curricular materials for inner-city children in American primary schools. Labov stresses that it is not only the linguistics structures that need to be adapted to these pupils, who speak non-standard English, but some of the contents as well: “Many of the narratives I have written for our Individualized Reading Program deal with conflict between students and the school, and the injustice that children see in the world around them. In contrast, most of the standard school-reading materials deal with a happy, anodyne, and irrelevant world in which children take their sand buckets to the beach and dip their toes in the water. By the time they reach the fourth grade, most of our students are alienated from the reading process as they have known it and from the institution of education as a whole” (Labov 2010:22).

Wolfram describes a somewhat different task, one of educating towards dialects awareness and preservation. He describes this as the most “difficult and elusive” challenge of his career (Wolfram 2000:279) because “principals, teachers, and parents [have been] worried that the programs might teach students to speak ‘bad language’ and spread dangerous linguistic propaganda” (Wolfram 2000:279). However, eventually, once access is gained, Wolfram reports that the North Carolina Language and Life Project has been met with great enthusiasm and has spread beyond the state’s schools in the private sector as well, and with the production of documentary films and publication of books and pamphlets, has reached many homes across the community.
In London as well, it appears as if the Multicultural London English project (see, e.g., Cheshire et al. 2011, Kerswill 2013, Kerswill et al. 2013), while predominantly devoted to sociolinguistic studies within the discipline, is beginning to branch out into the domain of linguistic policy in the educational system. A prime example of such work in the London setting are Levey (2012), in which the author’s own research on London schoolchildren’s discourse (specifically, their use of the verb be) is leveraged in a preliminary attempt to offer advice to professionals outside of linguistics: “Practical applications of sociolinguistic research are not only of educational value to teachers and children but are also equally important to school-based professionals, such as speech and language therapists. Detailed information about dialect structure and the evolution of children’s variable speech patterns can inform therapeutic interventions, helping to reduce the incidence of inappropriate remediation strategies resulting from the erroneous association of non-standard usage with language impairment” (Levey 2012:418).

Also consider Trudgill (1978), Trudgill (1984, an edited volume, with several relevant papers), Milroy & Milroy (1998), Cheshire (2007), to name but a few studies by prominent variationists in Britain incorporating the study of sociolinguistics into elements of bilingual, bidialectal or other forms of multicultural education. Cheshire’s paper, which is framed as “responses from sociolinguists” to the kinds of issues educators often encounter in a bidialectal classroom, concludes as follows: “research on dialect, both ‘pure’ and applied, is

---

36 I am only referring here to works done in London in recent years within the framework of variationist sociolinguistics. There is a separate body of work led by such scholars as Itesh Sachdev (e.g., Young & Sachdev 2011) concerning teachers’ and pupils’ attitudes in London and elsewhere in Europe, but these are seldom based on empirical studies of linguistic features in the dialects of the speakers about whom they report.
essential, because it is only on the basis of knowledge of the linguistic features of the dialect and the standard, together with the sociocultural aspects of the situations in which dialect and standard are used, that realistic and effective policies can be developed. Furthermore, whatever the situation in which dialect and standard coexist, there is a great deal of evidence to show that children reach higher levels of educational achievement when their dialect has a recognised and explicitly valued place within the educational system” (Cheshire 2007:31).

Academic works on “language policy in Israel” are not hard to come by, including such that deal explicitly with the Palestinian community, who is bilingual in both Arabic and Hebrew. The main actors in this line of research are Muhammad Amara, Elana Shohamy and Bernard Spolsky. Of these three, only Amara, a native speaker of Palestinian Arabic, actually knows the language. None of these scholars derive their research from a variationist perspective. This, of course, does not invalidate it; it simply amplifies the need for the kind of research embodied in this thesis to be intertwined with research in linguistic education and related fields. I, personally, find it disconcerting that a state-of-the-art—type article on “Language rights in the multilingual society of Israel” (Shohamy 2006) treats Arabic, the predominant indigenous language of Palestine roughly on the same plane as Russian, a language of immigrants, who have been given automatic citizenship by the state, while millions of native Palestinians remain refugees either within Palestine or in the diaspora. It is also telling that the author not once mentions “Palestine” or “Palestinians.” The country is always referred to as “Israel,” and the Palestinians are referred to generically as “Arabs.”
Amara (2007, 2013) does refer to Palestinians as Palestinians, offers a more critical analysis of school curricula and distinguishes between indigenous Palestinians and immigrants in linguistic policies, but mostly insofar as the purpose of Hebrew instruction for the former is to augment the pupils’ native language, whereas for the latter it is to replace it (Amara 2007:245).

A flurry of studies that investigate the linguistic patterns of the 1948 Palestinian community from a variety of angles, in which they are contextualized beyond the narrow scope of Israel and the Zionist project, have begun appearing in recent years. Such are Uhlmann (2012), regarding the teaching (and learning) of Arabic grammar by Palestinians in Israeli universities; and Isleem (2012, 2013), which are combinations of the kinds of studies we've seen from Shohamy and her associates on linguistic landscape and language policy, and a dash of sociolinguistic examination, particularly in the domain of code-switching, of several Druze speech communities in Palestine.

6.4 Desiderata

What I have found striking in reviewing the literature on Hebrew-Arabic contact is the absence of some key terminology with respect to the role of Hebrew in this duet. I am specifically referring to the lack of any mention—as far as I could detect—of Hebrew as a colonial or colonizing language, at least not within a purely linguistic discussion. It was refreshing, however, to encounter a paper by literary theorist Lital Levy, in which she analyzes two novelists who wrote in Arabic in 1948 Palestine: Palestinian Emile Habiby and Iraqi Jewish Samir
Naqqash, who she argues was “brought against his will to Israel” (Levy 2003:110). In this article she unequivocally describes Israel as a “neo-colonial power” (Levy 2003:109), and makes the following observation about the two writers: “while Habiby’s Arabic narrative portrays Hebrew as the dominant (colonizing) tongue, in Naqqash’s work it is the colloquial dialect of the Iraqi Jews that is dominated by standard/literary Arabic (and, to a lesser degree, by the Muslim colloquial dialect and by Israeli Hebrew)” (Levy 2003:108).

That Israel is a colonialist power and Zionism, the national movement that propelled its founding, has been a colonialist ideology, are now accepted truisms across much of the Middle East and in many other parts of the world, though, as historian Ilan Pappé notes, in some circles, there is still much resistance to this assertion: “Zionist settlers—indeed, Zionist thought and praxis—were motivated by a national impulse but acted as pure colonialists. Such an argument has not been easily accepted until today in the United States and Israel” (Pappé 2008:612). The Zionist settlement in Palestine, Pappé continues, “happened in a century when French settlers colonized Algeria, claiming an atavist and emotional link to the Algerian soil no less profound than the one professed by the early Zionists with regard to Eretz Yisrael” (Pappé 2008:612-613).

Pappé—at least rhetorically—claims to be “bewildered” by this lack of consensus (e.g., in the US and Israel) regarding the colonial nature of Zionism, and I should have probably framed my own sense of astonishment upon failing to find references to Hebrew as a colonizing language in the linguistic literature as a hyperbole of sorts. But I, too, was thinking of cases like the North African one, in which it would be trivial to refer to French as a colonizing language.
Hoffman (2008) lays out the case for Morocco, which was technically a protectorate, not a colony, but notes: "In reality, the distinction between protectorate and colony was not clearly applied. It is unsurprising that today Moroccan laypeople typically refer to the French period as *al isti’mar* [sic] and only rarely as *al himaya* [sic] (the protectorate)” (Hoffman 2008:728).

Why is this important? I would argue that framing the Arabic-Hebrew situation in Palestine as a colonized-colonizer language opposition bears significance beyond a mere stubborn insistence on historical accuracy. There are two points that are to be made in this regard:

4. **Context is important.** If the historical, political, legal, literary, cultural, educational (and so forth) environments in which the languages are used with relation to one another are comparable to such environments in other colonial settings, we must treat this setting on equal footing, especially if there is linguistic evidence indicating that one language is acting as a superstrate to the other.

1. **The linguistic literature, particularly in sociolinguistics and historical linguistics, is rife with examples of the effects of colonization on language contact.** We are well aware of its role in the formation of pidgins and creoles (see, e.g., Mesthrie 2009:271-294). However, the role of colonization is far from limited to pidginization and creolization. We have already mentioned Hoffman’s analysis of colonization in Morocco and elsewhere in North Africa, where not only did French emerge as a prestige language, but also the indigenous Amazigh (Berber) languages were pushed aside in favor of Moroccan Arabic,
partly because most French officials who resided in Morocco during the occupation had only learned Arabic but never Tamazight or Tashelhit (Hoffman 2008:727).

Consider, too, the following point made by Trudgill, albeit with respect to English: “Many colonial varieties, too, like North American English, also experienced language contact—with indigenous languages and other colonial languages in the new location. Urban dialects have also more often than not resulted from dialect contact, brought about by immigration from surrounding areas” (Trudgill 2011:67). It is tempting to take this statement and superimpose it on the Palestinian case. Yet one wonders: what are the parallels here? Perhaps the outcome that are the varieties of North American English are akin to Modern Hebrew, although unlike English, Hebrew had not been a productively spoken language prior to colonization. And Trudgill’s statement doesn’t tell us what happened to the indigenous languages of North America (which would be the parallels of Palestinian Arabic in our current study). Of course we know from many other studies (as well as from elsewhere in Trudgill’s own research, e.g., Trudgill 1991:63) that dozens of native North American languages “have been killed off since European contact.”

6.5 Universalist implications

Much of the linguistic literature about Arabic tends to be confined to discussing linguistic principles that are only pertinent to Arabic itself. It
should be obvious to the reader that I reject this rationale. Clearly, it makes sense to compare certain aspects of one’s analysis to other analyses conducted with regard to the same language, but one of our goals in linguistics is to discover general principles about language (rather than about a language). Maltese, of course, is something of a hybrid. Some scholars tend to classify it a highly standardized Arabic variety, while others prefer to view it as a closely related, yet distinct language (Brincat 2008:141; Mifsud 2008:146). Borg (1997:271) sums it up succinctly when he calls Maltese a “highly deviant offshoot of vernacular Arabic.” I have made reference to Maltese in Chapter 4, because some of the reflexes of the (ʕ) variable, specifically compensatory lengthening of the preceding vowel, and creaky voice (which, for this study, we refrained from coding as a discrete variant), were also documented in Maltese. For the purpose of that analogy, it was almost assumed that what Palestinian Arabic in Jaffa and Maltese had in common was that they were Arabic vernaculars which were prone to sound change (and other linguistic changes as well) as a result of language contact.

In her survey of the diachrony of Maltese, Mori makes a vague statement regarding colonization in Malta, followed by another, somewhat controversial assertion implying that Arabic had been the colonizing language: “Thanks to its strategic position in the middle of the Mediterranean Sea, some of the most important cultural and political forces have left their influence on the island of Malta: Phoenicians,
Romans, Arabs, Knights of the Order of St. John, and more recently the French and English domination. From the VIIIth century until the second half of the XXth century, Malta experienced colonization which brought different populations, speaking various languages, to the island, the end result being a diglossic condition which affected the Maltese language” (Mori 2009:292). And later: “Considering Arabic colonization as *termine a quo*, in Malta the Maghrebine variety of Arabic has progressively differentiated itself from other Western Arabic dialects” (Mori 2009:292).

Borg is more specific as to what he considers Malta’s colonial period: “Throughout Malta’s colonial period under British rule (1800–1964), knowledge of English was the key to social preferment and professional advancement. Many Maltese are bilingual in M and English” (Borg 1997:264).

It is interesting therefore that in the Maltese case, while there is virtually no dispute\(^{37}\) that colonization is a part of the country’s history, and that this has contributed to the shaping of the language in its current form, the precise nature of this colonization, and even the identity or identities of the colonizing language or languages are not agreed upon by all scholars.

The bottom line is that we should begin—or rather continue—to contextualize our sociolinguistic analyses, both within linguistics and across related fields. The data collected for the purpose of this thesis, the corpus of sociolinguistic interviews, was only used for a fraction of its value. There are many other

---

\(^{37}\) I hedge this statement with the adverb “virtually,” not because I know of any specific dispute, but because I find it curious that neither the “Malta” entry nor the “Maltese” entry in the *Encyclopedia of Arabic Language and Linguistics* mention any term related to colonization.
linguistic variables to be analyzed (some of them were mentioned in Chapter 5), and many other tangential issues to discuss in collaboration with colleagues in the social sciences and the humanities. In the remainder of this conclusion chapter I will outline additional research outcomes that are either already underway or envisioned as a result of my work on the data thus far.

6.6 What more can I do?

One of the most recent contributions to the sociolinguistic study of Palestinian Arabic came in the form of a 2013 University of Essex Master of Arts dissertation by William Cotter, who investigated two variables in the dialect of the city of Gaza: the historic voiceless uvular fricative (q), realized traditionally in Gaza as a voiced velar stop [g], but in other urban Palestinian dialects as a glottal stop [ʔ]; and the feminine suffix morpheme (ah), which in the traditional Gaza dialect has been pronounced [a] in all phonological environments, but in most other Palestinian dialects (except most Bedouin Palestinian varieties) is raised to [e] in non-back environments.

Cotter’s corpus is unique in that it is the first series of sociolinguistic interviews to have been carried out in Gaza City, it is the first comprehensive linguistic survey of Gaza in nearly a century (since Bergstäßer 1915), and notably, a subset of the speakers in this corpus (7 out of 22) comprises refugees who were either born in Jaffa or are offspring of Jaffa natives (Cotter 2013:7-8).

Two collaborative projects are currently in progress:
1. A joint paper pooling the speakers of Jaffa origins from Cotter (2013) with an equal number of demographically comparable speakers from the corpus collected for this thesis. We are complementing the analysis done in Cotter (2013) by adding the (ʕ) variable to the speakers in that sample and the analysis in this thesis by adding the (q) and (ah) variables (although, in reality, our prediction is that they will prove not to be variables at all within the Jaffa speech community). We are currently preparing this paper for presentation at the upcoming Arabic Linguistics Symposium and are hoping to submit it subsequently for publication.

2. An entry provisionally entitled “Sociolinguistics of Palestinian Arabic” to be published in the online edition of *Encyclopedia of Arabic Language and Linguistics*.

There are various other research projects that I envisage deriving undertaking based on the corpus I have already collected, including analyzing the remaining phonological variables outlined in Chapter 5, providing a more detailed account of the various forms of synthetic versus analytic possession (also see Chapter 5), and a discussion of code-switching among Palestinians who are also fluent in Hebrew (also see Henkin-Roitfarb 2011). There is also room, I believe for more detailed microanalysis of subsets of the data, in cases where individual linguistic behaviors and/or life trajectories are somewhat atypical.

I have begun, for instance, investigating the differing patterns of pharyngeal usage among three generations of Jaffa women (Horesh 2013). I intend to expand this study on both the quantitative and qualitative planes, as I believe
that the individual “stories” each of these woman had to tell: one was a woman in her 30s juggling a job in a public school with attaining a postgraduate degree and attempting to become a single mother through artificial insemination, all the while battling a host of stigmas both within the Palestinian community in which she lived and in the Jewish community among whom she spent much of her time leisure and academic pastime; another was a fortune-teller, renowned not only in Jaffa, but throughout the country among Palestinians and Jewish Israelis alike, who was kind enough to actually simulate a session of coffee grounds reading with me prior to conducting a more conventional sociolinguistic interview; and the third was an elderly woman, who had not been living in Jaffa since childhood, as both family life and the ethnic cleansing of 1948 (the Nakba) have taken her in roundabout ways to various places in the region, until she had finally settled with her husband in the village of Jaljulye.

The scope of this thesis has only enabled me to briefly touch upon some aspects of the speech patterns of ‘neviːn,’ ‘ʔumm jaziːd’ and ‘ʔumm xaliːl,’ but I look forward to providing a more fine-grained account of their speech in a dedicated paper. I see this as a potential contribution to the ethnography of the community, as well as to the sociolinguistic of Palestinian Arabic. Gillian Sankoff in recent years has probably been the chief proponent of the study of language change across the lifespan. While I may not have access to the types of longitudinal studies she and her collaborators have conducted in Montréal (e.g., Sankoff & Blondeau 2007), I can no doubt attest to some of the issues brought forth in their argumentation, e.g., upward mobility as a factor (see Sankoff & Blondeau 2007:572), but also contribute data from a community, which has
known upheavals quite dissimilar to most of the Western speech communities that have been the objects of so many sociolinguistic studies thus far.
References


Sociolinguistic and Pedagogical Dimensions of Dialects in Education. 14-33.


doi:10.1017/S1366728906002598


Isleem, Martin A. 2012. Language attitude and change among the Druze in Israel. PhD dissertation, University of Texas at Austin.


References


First appeared online: 2013. First Print Edition: 9789004176423, 20130809


