

**CONTACT-INDUCED CHANGE IN THE LEVANTINE: EVIDENCE FROM
LEBANESE AND PALESTINIAN ARABIC**

YASMINE ABOU TAHA

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Doctorate in Philosophy degree in Linguistics

Department of Linguistics
Faculty of Arts
University of Ottawa

© Yasmine Abou Taha, Ottawa, Canada, 2022

ABSTRACT

In the Arabic-speaking world, sociopolitical upheaval, extended conflict and population displacement have triggered extensive contact between mutually intelligible varieties of the language. Notwithstanding these developments, Arabic sociolinguistic research on dialect contact settings remains limited to certain well-documented areas (e.g., Al-Wer 2020), with markedly less research targeting other locales believed to be highly propitious to convergent change, such as the long-term contact situation in Lebanon involving Lebanese and Palestinian Arabic (Fityan 1981; Hennessey 2011). Furthermore, few studies are embedded in a (comparative) variationist sociolinguistic framework (Owens 2013), and even fewer studies are articulated from a socio-historical perspective incorporating diachronic data sources with which to better understand the process of language change in Arabic (Owens 2013). Much previous research on Arabic dialects is also based on investigations of phonological variation (Al-Wer and de Jong 2018), with correspondingly less attention paid to (morpho-)syntactic variation (Choueiri 2019).

The present study aims to address existing lacunae in the research literature by investigating the outcomes of dialect contact in Beirut between Palestinian Arabic (PA), the minority variety, and Lebanese Arabic (LA), the majority variety. Drawing on the framework of comparative variationist sociolinguistics (Poplack and Tagliamonte 2001) as well as research on dialect contact (Britain and Trudgill 2005), this study combines synchronic and diachronic data sources to compare three variables in LA and PA: a phonological variable, involving the word-medial raising of /a:/ to [e:] (e.g., [ka:n] alternating with [ke:n] ‘he/it was’); and two morpho-syntactic variables: verbal negation and future temporal reference. The overarching aim of the research is to examine the extent to which PA shows evidence, as gauged from linguistic constraints on variant selection and variant repertoires, of becoming more structurally similar to LA in different linguistic components (Cheshire, Kerswill, and Williams 2005).

The synchronic data come from 45 hours of spontaneous speech recorded in Beirut from 39 Palestinian and 27 Lebanese speakers stratified by age, sex, and level of education, generating 7,671 tokens representing the three targeted variables. A further 15,381 tokens of these three variables come from two diachronic datasets. The first is a sub-set of speech recordings from the *Palestinian Oral History Archive*, an online compendium of interviews with first-generation (older) Palestinians in Lebanon, recorded between the 1990s and early 2000s. The second diachronic dataset is the *Lebanese Popular Theatre Corpus* (LPTC), based on 34 televised plays dating from the 1960s and performed in colloquial LA.

Results reveal that the [e:] variant, a stereotypical feature of LA, but not emblematic of PA spoken in Beirut (Hennessey 2011), is virtually absent from the speech of the older Palestinian generation in the synchronic and diachronic datasets, but it increases significantly in the speech of young (third-generation) Palestinian speakers, who replicate the linguistic conditioning of variant selection in LA. These results bolster the inference of contact-induced change in PA due to the influence of LA. With respect to verbal negation, the findings show that there is convergent change in terms of overall variant rates in this variable system in PA. Evidence suggests that this variable system is undergoing dialect levelling as a result of contact, with

socially marked minority variants diminishing over time in the speech of educated Palestinians. The future temporal reference system, however, seems to be less amenable to contact-induced change, despite similarities in surface forms between LA and PA. Results indicate that this variable system is undergoing an internal change in PA independent of contact with LA, which is led by young, educated speakers, in line with what has been observed in PA spoken outside Lebanon (AbuAmsha 2016).

Viewed in the aggregate, the results show that even though it is claimed that (morpho-) syntactic variables may be less susceptible to convergent change than phonological variables (Cheshire et al. 2005; Hinskens et al. 2005), we do not find a neat division between phonology and morpho-syntax. Word-medial *imāla* is overtly commented on and explicitly identified by the targeted Palestinian speech community as a marker of Lebanese speech. Its iconic association with Lebanese speech patterns renders it particularly susceptible to long-term dialect accommodation for some Palestinians. Verbal negation is also subject to social evaluation, as gauged from explicit speaker meta-commentary, and socially marked exponents appear vulnerable to attrition over time. By contrast, the expression of the future temporal reference appears less socially indexical than the other variables and is not subject to normative commentary or overt correction. These differences implicate the social salience of the targeted variables as a key factor influencing their susceptibility to convergence.

Situating the results in a wider perspective, the findings highlight the utility of the comparative variationist framework in elucidating the process of language change in spoken Arabic, especially in PA as spoken in Beirut, as well as in distinguishing contact-induced change from internally-motivated change. The results of this study indicate that the effects of dialect contact, and critically, the existence of contact-induced change cannot be fully understood without using a multi-faceted comparative approach incorporating horizontal and vertical comparisons. The results converge in demonstrating that an empirically accountable quantitative approach based on actual speech data is capable of transcending the limitations of alternative frameworks of analysis that have been used to investigate change in dialect contact scenarios in the Arabic-speaking world.

ACKNOWLEDGEMENTS

I am indebted to several people for their support and assistance during my Ph.D. program and the completion of this thesis.

I would first like to express my heartfelt gratitude and appreciation to my supervisor, Professor Stephen Levey, who has been immensely supportive, patient and motivating since I started my Ph.D. program in 2017 and throughout this project. Writing this thesis during the uncertainties of a global pandemic and in isolation has been made easier because of Professor Levey's understanding, kindness, empathy and words of encouragement. He has always been available to listen to and discuss my concerns, respond to my emails and questions, and connect me with the resources I need. He has been extremely generous with his time and excellent advice. Professor Levey's expertise and guidance were invaluable in every stage of this thesis, and his insightful feedback has inspired me to think critically and improved this thesis substantially. He has also been instrumental in the development of my research skills in variationist sociolinguistics. Thank you very much, Professor Levey, for your dedication and inspirational leadership. I cannot speak highly enough of you and all what you have done for me. You have become a true mentor to me, and I look up to you as a role model. I truly feel privileged to have had the opportunity to work with you, and I look forward to continuing to collaborate with you in the future.

I would also like to thank the members of my thesis evaluation committee for accepting to review this thesis and for their time and invaluable feedback: Professor Éric Mathieu, Professor Dennis Ott, Professor Shana Poplack, and Professor Lotfi Sayahi.

I extend my gratitude to all the professors in the Department of Linguistics at the University of Ottawa. Special thanks to Professor Stephen Levey and Professor Dennis Ott for

supervising me in my two qualifying papers. I am also grateful to friends, colleagues and administrative staff in the Department of Linguistics, as well as friends in Ottawa who have made being far from home much easier. I wish to also extend my gratitude to my friends in Beirut for their encouragement and for always staying in touch.

I would like to thank the University of Ottawa for providing me with an admission scholarship which has funded my studies. I am thankful for the Department of Linguistics at the University of Ottawa for giving me the opportunity to work as a Teaching Assistant and instructor, which has given me invaluable professional experience. Warmest thanks to Professor Stephen Levey for his unparalleled support and guidance during my first time teaching as a part-time instructor. My deepest thanks to him for also providing me with several opportunities to work as a research assistant, which has helped me immensely financially and professionally.

My sincere gratitude goes to all the participants in this study who generously gave me their time and speech data, without which this thesis would not have been possible. I really appreciate their willingness to help me and take part in the study. Special thanks to the Palestinian speakers who live in refugee camps and the individuals who introduced me to them and assisted me in locating informants for this study. These speakers, especially the first-generation Palestinians, made the refugee camp feel like home to me. I feel fortunate to have had the chance to get to know them and speak with them. I would like to express my deepest gratitude to one first-generation speaker who lived in the camps, and who has now sadly passed away. He welcomed me in his home in the camp and introduced me to other speakers, and was very keen to share his narratives of personal experience. He is one among many of the participants in this study who will always be etched in my memory.

I am also grateful to my family, who has been very encouraging and enthusiastic since I got accepted to pursue my Ph.D. at the University of Ottawa. I am especially thankful to my aunt for all her support and for always being there to listen to me. Last, but by no means least, I wish to express my eternal gratitude to my mother and father for their constant love and support. I am especially indebted to my mother, who has sacrificed a lot for me to reach this stage of my education. Words cannot express my appreciation for her and all what she has done for me. I thank her for believing in me even when I doubted myself, and for her endless support and encouragement.

Table of Contents

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iv
List of Tables	xi
List of Figures	xiii
1 Introduction.....	1
1.1 Rationale.....	1
1.2 Context.....	3
1.2.1 Social factors.....	3
1.2.2 Linguistic factors.....	7
1.3 Research questions.....	12
1.4 Contribution of the data sources to assessing contact-induced change in PA.....	13
1.5 Contribution of the theoretical framework to assessing contact-induced change in PA...16	
1.6 Road-map of the thesis.....	18
2 Literature Review.....	20
2.1 Raising of /a:/ to [e:]	20
2.1.1 <i>Imala</i> in Old Arabic.....	22
2.1.2 <i>Imala</i> in modern dialects of Arabic.....	25
2.1.3 <i>Imala</i> in Palestinian and Lebanese Arabic.....	28
2.1.4 Social conditioning of <i>imala</i> in PA spoken in Beirut.....	30
2.2 Verbal negation.....	33
2.2.1 Overview of verbal negation in Standard Arabic.....	34
2.2.2 Overview of verbal negation in Arabic dialects.....	35
2.2.3 Verbal negation in Lebanese and Palestinian Arabic	40
2.3 Future temporal reference	43
2.3.1 Overview of the expression of futurity in Standard Arabic.....	44
2.3.2 Future temporal reference in some Arabic dialects.....	46
2.3.3 Linguistic constraints on the use of future variants in spoken Arabic	50

2.3.4	Grammaticalization of future tense markers.....	55
2.3.5	Linguistic constraints and measurements of grammaticalization	61
2.4	Critical summary	66
2.4.1	Methodological and analytical gaps in the literature.....	66
2.4.2	Building on the literature.....	70
3	Theoretical Framework and Data.....	73
3.1	Theoretical framework.....	73
3.1.1	Data collection in the variationist framework.....	74
3.1.2	Data analysis in the variationist framework	76
3.1.3	The comparative variationist method.....	77
3.2	Data.....	79
3.2.1	Synchronic speech data	79
3.2.2	Diachronic data	86
3.3	Summary	101
4	<i>Imala</i> in Word-Medial Position.....	102
4.1	Circumscribing the variable context	102
4.2	Factor groups	104
4.2.1	Social factor groups.....	104
4.2.2	Linguistic factor groups.....	113
4.3	Results.....	116
4.3.1	Overall distribution of variants	116
4.3.2	Distributional analysis of social factors	118
4.3.3	Multivariate Analysis	130
4.4	Excursus on word-final <i>imala</i>	139
4.5	Discussion	142
5	Verbal Negation.....	147
5.1	Circumscribing the variable context	147
5.2	Factor groups	149
5.2.1	Social and stylistic factor groups	149
5.2.2	Linguistic factor groups.....	159
5.3	Results.....	163
5.3.1	Overall distribution of variants	163

5.3.2	Social and stylistic factor groups	165
5.3.3	Linguistic predictors of variant choice	180
5.4	Discussion	201
6	Future Temporal Reference.....	206
6.1	Circumscribing the variable context	206
6.2	Factor groups	209
6.2.1	Social factor groups.....	209
6.2.2	Linguistic factor groups.....	211
6.3	Results	214
6.3.1	Overall distribution of variants	214
6.3.2	Multivariate analysis of social factors	217
6.3.3	Multivariate analysis of linguistic factors.....	222
6.4	Discussion	241
7	Conclusion	244
7.1	Key motivations for the study	244
7.2	Objectives of the study	246
7.3	Utility of the data sources for addressing the research objectives	247
7.4	Contributions of the comparative variationist framework to addressing the research objectives.....	247
7.5	Convergence and divergence in different grammatical components	248
7.6	Limitations of the study and future lines of inquiry	257
7.7	Epilogue	260
	References.....	262
	Appendix A: Distribution of negative constructions in northern Levantine dialects	313
	Appendix B: Distributional analysis of the contribution of linguistic factors to the selection of word-medial raised [e:].....	314
	Appendix C: Overall distribution of the variants of verbal negation according to linguistic factors across different age groups in LA and PA.....	316

Appendix D: Contribution of linguistic factors to variant selection in verbal negation in the POHA	318
Appendix E: Contribution of social factors to variant selection in future temporal reference in LA and PA.....	319
Appendix F: Contribution of linguistic factors to variant selection in future temporal reference across age groups in LA and PA	321
Appendix G: Contribution of linguistic factors to variant selection in future temporal reference in the POHA	326

List of Tables

Table 2.1: Consonants reported to disfavour vowel raising in the preceding phonological context ...	25
Table 2.2: Default <i>imala</i> variants in word-medial and word-final positions across some Arabic varieties.....	26
Table 2.3: Major exponents of verbal negation in some Arabic varieties	40
Table 2.4: Reported linguistic constraints on variant choice in verbal negation	43
Table 2.5: The distribution of <i>rah</i> , <i>ha-</i> , <i>bidd-</i> and <i>b-</i> across some vernacular varieties of Arabic	49
Table 2.6: Reported linguistic constraints on variant choice in future temporal reference in spoken Arabic	54
Table 2.7: Summary of predictions for the grammaticalization of <i>go</i> -futures	65
Table 3.1: Characteristics of the Palestinian and Lebanese sample populations	81
Table 3.2: Characteristics of the speakers in the POHA sampled for this study	91
Table 3.3: List of plays in the Lebanese Popular Theatre Corpus (LPTC)	93
Table 4.1: Number of tokens of raised and non-raised variants of word-medial <i>imala</i> extracted from the diachronic and synchronic datasets	104
Table 4.2: Overall distribution of the variants of word-medial <i>imala</i> in the diachronic and synchronic datasets	117
Table 4.3: Individual rates of word-medial <i>imala</i> among 14 speakers in the POHA	118
Table 4.4: Individual rates of word-medial <i>imala</i> among Palestinian second-generation (middle-aged) speakers and their socio-demographic characteristics.....	127
Table 4.5: Contribution of social and linguistic factors to the selection of word-medial raised [e:] in the LPTC and the Lebanese synchronic dataset (all age groups).....	133
Table 4.6: Characteristics of the younger Palestinian and Lebanese sample populations.....	136
Table 4.7: Overall distribution of the variants of word-medial <i>imala</i> in the PA and LA synchronic datasets (younger speakers)	136
Table 4.8: Contribution of social and linguistic factors to the selection of word-medial raised [e:] in younger Lebanese and third-generation Palestinians	137
Table 4.9: Overall distribution of the variants of word-final <i>imala</i> in the Palestinian synchronic dataset.....	140
Table 4.10: Contribution of the preceding phonological segment to the selection of word-final raised [e] in the PA synchronic dataset (all age groups)	141

Table 5.1: Number of tokens of verbal negation extracted from the synchronic and diachronic datasets	149
Table 5.2: Overall distribution of the variants of verbal negation in the synchronic and diachronic datasets	163
Table 5.3: Overall distribution of variants according to social and stylistic factors in the LA and PA synchronic datasets (all age groups).....	167
Table 5.4: Contribution of social and stylistic factors to variant selection in the POHA.....	172
Table 5.5: Overall distribution of variants according to social and stylistic factors in the LPTC	176
Table 5.6: Overall distribution of variants according to linguistic factors across different age groups in the LA and PA synchronic datasets.....	182
Table 5.7: Contribution of linguistic factors to variant selection in the POHA and young PA speech	190
Table 5.8: Overall distribution of variants according to linguistic factors across post-secondary educated speakers in the POHA and younger Palestinians in the synchronic dataset.....	193
Table 5.9: Overall distribution of variants according to linguistic factors in the LPTC and the LA synchronic dataset	197
Table 6.1: Number of tokens of future temporal reference extracted from the synchronic and diachronic datasets	208
Table 6.2: Overall distribution of the exponents of futurity in the synchronic and diachronic datasets	214
Table 6.3: Contribution of social factors to variant selection in the POHA	217
Table 6.4: Contribution of social factors to variant selection in the LPTC	218
Table 6.5: Contribution of social factors to variant selection in LA and PA synchronic datasets.....	220
Table 6.6: Contribution of linguistic factors to variant selection across age groups in LA and PA synchronic datasets.....	224
Table 6.7: Summary of the comparisons across middle-aged LA and middle-aged PA speech	228
Table 6.8: Summary of the comparisons across young LA and PA speech	230
Table 6.9: Summary of the comparisons across middle-aged and young LA speakers	232
Table 6.10: Summary of the comparisons across older, middle-aged and young speakers in PA.....	235
Table 6.11: Contribution of linguistic factors to variant selection in the POHA and in young PA speech	237
Table 6.12: Summary of the comparisons across the POHA and young PA speech	240

List of Figures

Figure 3.1: Refugee camps targeted in this study	80
Figure 4.1: Arab settlements abandoned in 1948-1949 (Institute for Palestine Studies 1988; see also Morris 2003:xi)	120
Figure 4.2: Distribution of the [e:] variant in word-medial position by age and sex of the characters in the LPTC.....	122
Figure 4.3: Distribution of the [e:] variant in word-medial position by social status of the characters in the LPTC.....	122
Figure 4.4: Distribution of the [e:] variant in word-medial position by age and sex in the Lebanese synchronic dataset	123
Figure 4.5: Distribution of the [e:] variant in word-medial position by age/generation and sex in the Palestinian synchronic dataset	123
Figure 4.6: Distribution of the [e:] variant in word-medial position according to age and level of education in the Lebanese synchronic dataset	125
Figure 4.7: Distribution of the [e:] variant in word-medial position according to age and level of education in the Palestinian synchronic dataset	125
Figure 4.8: Distribution of the [e:] variant in word-medial position according to age and place of residence in the Palestinian synchronic dataset	126
Figure 5.1: Location of Cornet Chahwan (and Beirut).....	156
Figure 6.1: Overall variant distributions by age in the LA and PA synchronic datasets.....	216

1 Introduction

1.1 Rationale

The Arabic-speaking world has witnessed drastic social and demographic shifts in recent decades, triggered by socio-political upheaval, protracted conflict and large-scale migration (Horesh and Cotter 2016:372). As a result, Arabic varieties have been involved in heterogeneous contact situations (Lucas and Manfredi 2020:1). This affords valuable opportunities for exploring the sociolinguistic outcomes of contact between speakers of different but mutually intelligible dialects of Arabic (see e.g., Al-Essa 2009; Al-Wer 2007b, 2020; Cotter 2013, 2016; Cotter and Horesh 2015; Hennessey 2011; Horesh 2014; Miller et al. 2007).

The sociolinguistic complexity of the Arabic-speaking world has led to a growing body of research on dialect contact, especially in the Levant (Cotter and de Jong 2019:51–54; Haeri and Cotter 2019:244; see e.g., Al-Wer 2007b, 2020; Cotter 2013, 2016; Cotter and Horesh 2015; Hennessey 2011; Horesh 2014). Nevertheless, only certain areas have been extensively documented (Horesh and Cotter 2015), such as the capital city of Jordan, Amman (involving contact between Jordanian and Palestinian Arabic) (see Al-Wer 2002, 2007b, 2011, 2020), while other contact situations remain insufficiently investigated. A case in point is Lebanon, a context believed to be particularly hospitable to contact-induced change due to long-term contact between the local Palestinian refugee population and Lebanese Arabic speakers (Fityan 1981; Hennessey 2011). In much previous work addressing dialect contact in Arabic-speaking communities, there is also a dearth of systematic quantitative treatments of linguistic features articulated from a comparative variationist perspective (Owens 2013:11). This area of research additionally suffers from the lack of a socio-historical component (Owens 2013) to assess

linguistic change, as many Arabic scholars have tended to infer language change without always supplying the requisite diachronic evidence without which change cannot be ratified (see Lucas and Manfredi 2020:22). Arabic sociolinguistic research in general has also been heavily preoccupied with investigating aspects of phonological variation (Cotter 2020:568; see e.g., Al-Wer 2007b, 2011, 2020; Cotter 2013, 2020), with the vast majority of studies focusing on consonantal variation (Al-Wer 2002:63–64, 2007b:67; see e.g., Abdel-Jawad 1981, 1986; Al-Essa 2009; Al-Khatib 1988; Al-Wer 1991, 1999; Al-Wer and Al-Qahtani 2016; Haeri 1991). By contrast, the vocalic system of Arabic varieties remains understudied (Al-Wer 2002:63–64, 2007b:67; Cotter 2020:576). Cotter (2020:576) notes that this is especially true in dialect contact situations, arguing that “more robust investigation of vocalic change in Arabic dialects represents a pressing area of concern for Arabic sociolinguistics.” Perhaps even more conspicuously, (morpho-)syntactic variation remains under-investigated in Arabic sociolinguistics (e.g., Brustad 2000; Choueiri 2019), despite the potential for broad, comparative studies in this domain (Owens 2013:10–11).

Building on earlier foundational Arabic sociolinguistic research, the present study aims to address the aforementioned gaps in the research literature by examining the effects of sustained contact between the majority variety of Lebanese Arabic (hereafter, LA) and the minority variety of Palestinian Arabic (hereafter, PA) in the city of Beirut. Broadly speaking, the present research addresses a classic question in historical and contact linguistics (Owens 2013:4): What happens to the structure of a minority dialect in contact with a socially dominant and majority dialect (Thomason and Kaufman 1988)? Following Thomason (2001:92–93), a solid case for establishing the effects of contact on the structure of a minority dialect can only be made based on a variety of variables in multiple linguistic components, rather than on the basis of a single

linguistic feature. Accordingly, drawing on comparative variationist sociolinguistics (Labov 1972a; Poplack and Meechan 1998; Poplack and Tagliamonte 2001), and research on contact linguistics (e.g., Britain and Trudgill 2005; Matras 2009; Thomason 2001; Winford 2003), this study combines synchronic and diachronic data sources to compare three variables in LA and PA. The first is a phonological variable, known in the Arabic grammatical tradition as *imāla* (lit. ‘inclining, bending to’) (Owens 2006:197), involving the raising and fronting of the long vowel /a:/ to [e:] in word-medial position, e.g., [ka:n] alternating with [ke:n] ‘he/it was’. The two remaining variables targeted are morpho-syntactic variables: verbal negation and future temporal reference. The primary objective of the present study is to examine the extent to which PA shows evidence of becoming more structurally similar to LA in different linguistic components.

1.2 Context

Contact-induced change is said to be facilitated (but not fully determined) by several social and linguistic factors (Thomason 2001, 2010:33; Thomason and Kaufman 1988). These factors help explain the rationale for focusing on the contact situation involving PA and LA.

1.2.1 Social factors

The major social factors that are propitious to convergent change include (i) the intensity and length of contact; (ii) the majority versus minority status of the varieties in contact; and (iii) the relative prestige attributed to different varieties (see e.g., Thomason 2001, 2010; Thomason and Kaufman 1988; Winford 2003).

The situation of PA and LA in Beirut satisfies the first criterion believed to be favourable to contact-induced change. Speakers of these two varieties of Levantine Arabic (Al-Wer and de

Jong 2018:527) have been in contact in Beirut since the 1948 Arab–Israeli war and the resultant exodus of thousands of Palestinians to Lebanon and to neighbouring countries. Approximately 470,000 Palestinian refugees are registered with UNRWA in Lebanon (UNRWA n.d.),¹ though as of 2020, it is estimated that around 300,000 Palestinian refugees live in Lebanon (Anera 2021).² About 180,000 of these refugees live in the country's 12 Palestinian refugee camps (Government of Lebanon and United Nations 2021:4).³ Palestinians inside and outside camps, including Lebanese-born Palestinians, generally cannot obtain Lebanese citizenship⁴ (Ghandour 2017:10).⁵ Because they do not have citizenship from a recognized state, they are *de facto* stateless. Under Lebanese law, they are classified as foreigners with no recognized state documents (El Masri 2020), resulting in numerous barriers to basic rights, such as education, healthcare and employment (Government of Lebanon and United Nations 2021:4). The exclusion of Palestinians from socioeconomic life in Lebanon aims to prevent their permanent settlement in the country (El Masri 2020:5). But while these restrictions may be shared among all Palestinian refugees in Lebanon (Horesh 2020:5), not all of them are isolated spatially from Lebanese society. In fact, the Palestinian community is highly stratified in terms of its isolation from the host society (Hennessey 2011:2), with some Palestinians being reportedly indistinguishable from their Lebanese counterparts in terms of their speech patterns, at least superficially (see Hennessey 2011:2, 30, 109, 128, 140), whereas others may be less integrated

¹ United Nations Relief and Works Agency for Palestine Refugees in the Near East

<https://www.unrwa.org/where-we-work/lebanon>

² American Near East Refugee Aid

<https://www.anera.org/where-we-work/lebanon/>

³ Lebanon Crisis Response Plan 2017 - 2021 (2021 update)

https://reliefweb.int/sites/reliefweb.int/files/resources/LCRP_2021FINAL_v1.pdf

⁴ Despite this restriction, Ghandour (2017:65) notes that Lebanese citizenship was granted to some Palestinians, especially Christian Palestinians or wealthy urbanites in historical Palestine, between 1948 and the 1960s (see also Sayigh 2005). In addition, some Palestinians who were inhabitants of the so-called Seven Villages (Palestinian villages that were previously part of Lebanon) were granted Lebanese citizenship in 1994 (see Kaufman 2006).

⁵ See also United Kingdom: Home Office, Country Policy and Information Note - Lebanon: Palestinians, June 2018, Version 1.0. <https://www.refworld.org/docid/5b32026b4.html> (United Kingdom: Home Office 2018)

into the host society because they may have limited contact with Lebanese people (Hennessey 2011:1). This is especially true for many (but not all) Palestinians who live in refugee camps (see Hennessey 2011).

The remaining social determinants of contact-induced change relate to the social status and the prestige attributed to the varieties in contact. While prestige is typically associated with the standard language, which in the Arabic-speaking world is Classical Arabic, it is important to note that we cannot equate metropolitan prestige norms with the standard in the context of Arabic (Al-Wer 1997; Ibrahim 1986; Labov 2015:12). Classical Arabic is valued as a symbol of Arab culture and Islamic religion, but it has no native speakers and is not a prestige target in everyday speech (Al-Wer 1997:254; Ibrahim 1986:118; see also Abdel-Jawad 1987; Haeri 1996). As Ibrahim (1986:118, 124) explains, in the Arabic-speaking world, locally recognized (spoken) prestigious varieties exist apart from Classical Arabic. These varieties are typically spoken in urban centres (Ibrahim 1986), and they are used by speaker groups who possess desirable social traits and whose speech patterns likewise become desirable (Hennessey 2011:14; see Fasold 1984:208). In Lebanon, Beirut Arabic acts as the prestigious local standard (Al-Wer and de Jong 2018:526; Hennessey 2011:16; Versteegh 2001:153), and it exerts the primary influence on other vernaculars, including PA. For example, Hennessey (2011:3) explains that some Palestinians in Lebanon may have felt the pressure to shift to the socially majority variety of Beirut Arabic because they are generally viewed with increasing hostility (Lindholm Schulz 2003:60), and they can be discriminated against based on their Palestinian speech features. Accordingly, the majority status and the prestige attributed to Beirut Arabic, along with the long-term and intense contact between this variety and PA may contribute to the convergence of PA on LA (see Thomason 2001, 2010; Thomason and Kaufman 1988).

Notwithstanding the (overt) prestige attributed to Beirut Arabic, it is important to note that the notion of linguistic prestige can “apply to any property of a linguistic trait that would lead people to imitate it” (Labov 2001:24). In other words, we should take into account the possibility that PA speech features may have covert prestige for some Palestinians (see Labov 2001:23–24). Covert prestige entails that what might be stigmatised as characteristic of Palestinian speech norms in the wider Lebanese society can be maintained and utilised by some Palestinians as a marker of their group (and/or personal) identity (Hennessey 2011:17; see Chambers and Trudgill 2004:85; Labov 2001:24, 215–16; Milroy 1980:19). As such, some Palestinians may be favourably disposed to PA vernacular norms, regardless of the external pressure they may face in the mainstream Lebanese society. As we will see in Chapter 4 on word-medial *imala*, this was the case of some (young) Palestinians, who asserted their orientation towards their variety and Palestinian national identity during their interviews.

In addition to overt and covert prestige, what might affect Palestinians’ speech patterns are their social networks. Dense, multiplex social networks among Palestinians, such as those in refugee camps (Hennessey 2011:19), as well as the relative spatial and social isolation of the camps (Hennessey 2011:12, 34) are expected to contribute to the maintenance of PA speech norms (Hennessey 2011:19–20; see also Milroy and Llamas 2013:410). This is something that one of the speakers who lives in a camp in Beirut has acknowledged:

- (1) *Ah, ah, šanna lakne mšayyane, bas uh ʔetʔaw-- btetʔawwa bi l-- bi l-ʒuzu:r m-- ʔabšan maš l- ʔahel, ok? W bi-tawa:ʃol ma- fi ma bayn n-na:s bi l-moxayyama:t (...)*
[56/P/M/45/02:07:58]

‘Yes, yes, we have a specific accent, but it is strengthened through family ties of course, okay? And with the communication among people in the camps (...)’

In contrast to some Palestinian camp dwellers, some Palestinians with weaker ties to their community and strong ties to the Lebanese community may accommodate to Lebanese prestige

patterns (Hennessey 2011:19; see also Milroy 1980; Milroy and Milroy 2012). This is especially true of those who have grown up in majority Lebanese neighbourhoods and who have presumably been exposed to Lebanese speech patterns since childhood. This hypothesis reflects the argument of Milroy and Llamas (2013:410) that “[n]etworks constituted chiefly of strong (dense and multiplex) ties support localized linguistic norms, resisting pressures to adopt competing external norms. By the same token, if these ties weaken, conditions favorable to language change are produced” (see also Britain 2004:611).

1.2.2 Linguistic factors

In addition to social factors, there are linguistic factors believed to be hospitable to contact-induced change, which help explain the rationale for targeting the three linguistic variables in this study.

1.2.2.1 *Imala* in word-medial position

The phonological variable of *imala* in word-medial position has two variants, [a:] and [e:],⁶ as illustrated in the examples below:⁷

- (2) [a:]
 henne **ka:no** mjawzi:n w ʕandon **wla:d**. [50/P/F/33/02:39]⁸
 they were.3p married.3p and have.3p kids
 ‘They were married and they have kids.’

⁶ The degree of raising and fronting may vary in PA and LA (see Hennessey 2011:59; Levin 2006:314; Naïm 2005:277; Versteegh 2009:482).

⁷ All utterances are reproduced verbatim from the speech data.

⁸ Codes refer to the following: Speaker number/P=Palestinian corpus; L=Lebanese corpus/M=Male; F=Female/speaker age/time-stamp of the utterance in the audio recording. For the transcription of Arabic examples, the transcription conventions of the book *Arabic and Contact-Induced Change* were followed as closely as possible (see Lucas and Manfredi 2020:19–20).

- (3) [e:]
 halla? **mse:fra** bi July **siye:ħa**. [4/L/F/24/02:45]
 now travelling.1sf in July tourism
 ‘I am travelling now in July for tourism.’
- (4) [a:] and [e:]
 fi **je:mʕa:t** **kame:n** bi UK. [3/L/M/25/20:41]
 there.is universities also in UK
 ‘There are also universities in the UK.’

As detailed in Chapter 2, the linguistic environments in which *imala* is found vary across Arabic dialects (Al-Khatib 1988:261), but in this study, the variable context is restricted to word-medial position because it reportedly distinguishes LA from PA as spoken in Beirut (Hennessey 2011). Hennessey (2011:59) notes that LA is characterised by the raised [e:] variant, while PA has a lowered and backed [æ:] (Kelly 2017; Naïm 2005:277; Shahin 2007, 2011).⁹ Within a contact framework, this variable qualifies as a conflict site, or structural difference between varieties, which can be used as a diagnostic of contact-induced change (Poplack and Meechan 1998).

Word-medial *imala* is overtly commented on in the targeted Palestinian speech community and explicitly stereotyped as a feature of LA, e.g.:

- (5) *They talk with the alif, a a a a, tala:ta, you know, hone biʔu:lo tle:te.* [P/F/43/15:32]¹⁰

‘They [Palestinians] talk with the alif [first letter in the Arabic alphabet], a a a a, *tala:ta* ‘three,’ you know, here [in Lebanon] they say *tle:te* ‘three.’’

One of the Palestinian speakers in the camp, who I was introduced to as a Palestinian by another speaker, even ‘corrected’ me (the interviewer) for using the word-medial [e:] variant in the word *mse:fra* ‘(I am) travelling,’ as illustrated below:

⁹ As mentioned in Chapter 2, this is in contrast to Al-Wer’s (2007b:70) account, where she claims that in many urban Palestinian dialects, most notably in Jerusalem Arabic, /a:/ in word-medial position may be variably raised, ranging from raised /æ:/ ~ /ɛ:/ ~ /e:/, e.g. /ʒɛ:ʒ/ or /ʒe:ʒ/ ‘chicken’; /mɛ:xdi:n/ ‘we/they have taken’ (see also Al-Wer 2020:559). However, other accounts of *imala* in PA seem to disagree, arguing that *imala* in word-medial position is not characteristic of this variety (e.g., Hennessey 2011; Naïm 2005, 2011; Shahin 2007, 2011; Versteegh 2009).

¹⁰ This speaker was later excluded from the sample because I could only extract a limited amount of tokens of the three variables from the audio recording. She used more English than Arabic during the interview.

(6) *laʔ, hallaʔ ma beddek tʔu:le mse:fra, bitʔu:le msa:fra* [56/P/M/45/02:26]

‘No, now you don’t want to say *mse:fra* [Lebanese pronunciation of ‘travelling’], you say *msa:fra* [Palestinian pronunciation of ‘travelling’].’

The fact that Palestinians consciously associate word-medial *imala* with LA suggests that it may be accessible as a target for long-term dialect accommodation. This is in line with Trudgill’s (1986:11) argument that in contact with speakers of other varieties, speakers tend to modify features of their speech which they are most aware of.

1.2.2.2 Verbal negation

Innovative in this study is the extension of the purview of the investigation to morpho-syntax, targeting two variables from this linguistic component. In verbal negation, LA and PA share a similar set of variants: the negative proclitic *ma*; a discontinuous construction, *ma-š*; a post-verbal enclitic *-š*; and a non-discontinuous element, *miš*. These are illustrated below:

(7) MA
ma baʕref šu huwwe l-problem. [17/L/F/34/41:00]
 NEG know.1s what it/he the-problem
 ‘I don’t know what the problem is.’

(8) MA-š
 ana **ma** baʕraf-u:-š. [58/P/M/45/1:24:15]
 I NEG know.1s-him-NEG
 ‘I don’t know him.’

(9) -š
 beṯebb-š l-chemistry. [08/P/F/20/20:51]
 like.3sf-NEG the-chemistry
 ‘She doesn’t like chemistry.’

(10) MIš
miš ʕam yle:ʔe šiʔil. [40/L/M/83/25:53]
 NEG PROG find.3sm work/job
 ‘He is not finding a job.’

It is commonly assumed that typological congruence (Thomason 2001:71, 2010:44–45), or interlingual equivalence sites (Backus 2005:326) facilitate the “enhancement of inherent structural similarities found between two linguistic systems” (Bullock and Toribio 2004:91). In other words, the degree of typological congruence between the varieties in contact is reportedly a key linguistic predictor of contact-induced change (Thomason 2001:97, 2010:40). However, this assumption should be treated with caution, because surface similarity may be masking underlying structural difference (Poplack, Zentz, and Dion 2012:222). In addition, the traditional view enunciated in the literature is that (morpho-)syntactic variables may be less susceptible to the influence of contact than phonological variables (see Hinskens, Auer, and Kerswill 2005; but see Thomason 2001:71).¹¹ This is because (morpho-)syntactic variables may occur less often in spontaneous speech than phonological ones (Cheshire 1999:61) and they may be less likely to act as sociolinguistic markers. On the other hand, phonological variables may distinguish social groups (Cheshire 1999:72, 76; Winford 1996:188), so they are more likely to be socially salient than grammatical variables¹² (Cheshire 1999:61). From a structural point of view, some (morpho-) syntactic variables may also be less amenable to contact-induced change because they may be embedded in more complex linguistic systems than phonological variables (see e.g., Milroy 2007:165; Sankoff and Blondeau 2007:581). Accordingly, the working hypothesis in this study is that the phonological feature of *imala* in word-medial position may be more hospitable to contact-induced change than verbal negation (and future temporal reference; see sub-section below).

¹¹ Thomason (2001:71) disagrees with this traditional view, arguing that under the right social and linguistic conditions, “anything goes, including structural borrowing that results in major typological changes in the borrowing language.”

¹² In this thesis, grammatical variation is used to refer to (morpho-)syntactic variation (phonology is part of the grammar, but for ease of reference, I use the terms ‘phonological variation/phonological variable’ to refer to variation in the sound system).

Apart from the surface similarities between LA and PA in the expression of verbal negation, another reason for targeting this variable is that previous studies focusing on verbal negation in spoken Arabic have not been based on the empirically accountable quantitative methodology of variationist sociolinguistics. As explained in Chapter 2, verbal negation has been mainly examined syntactically or descriptively in spoken Arabic, with previous studies drawing on decontextualized examples in order to infer the linguistic constraints on variant choice (e.g., Brustad 2000; Lucas 2007, 2010, 2015). Social constraints on variant selection in verbal negation have also not received much attention in the relevant literature (e.g., Khairallah 2014). As such, the present study aims to address these gaps by investigating the internal and external conditioning of verbal negation in natural speech within a comparative variationist approach.

1.2.2.3 Future temporal reference

Similar to the case of verbal negation, LA and PA also share surface similarities in the variable expression of future temporal reference. In both varieties, this variable system involves the choice between *rah* and the proclitic *ha-*, both reported to derive from the motion verb *ra:yih* ‘going’ (Jarad 2014:101; Leddy-Cecere 2018:49–50, 2020:606), as well as *bidd-* (lit. ‘want’) and the *b-*prefix, which is cliticized to the imperfect verb in Levantine Arabic (including LA and PA).¹³ These variants are illustrated below:

- (11) RAH
rah ʔa:xud ʔabl l-MCAT¹⁴ five days break. [8/P/F/20/11:29]
 FUT take.1s before the-MCAT five days break
 ‘I will take a five-day break before sitting for the MCAT.’

¹³ As mentioned in Chapter 6, the *b-*prefix is a multifunctional particle that can also be used to denote non-future temporal meanings, such as the progressive and habitual aspect (see Jarad 2013:71; Owens 2018:211, 219).

¹⁴ Medical College Admission Test

- (12) HA-
 ana l-yōm ma **ħa** šu:f-a. [12/L/F/28/08:53]
 I the-day NEG FUT see.1s-her
 ‘I will not see her today.’
- (13) BIDD-
 baʃd sene **bidd-e** axsar min ħaʔ-a. [21/P/F/57/1:17:27]
 after year FUT.1s lose.1s from price-its.f
 Intended: ‘After one year, its value will drop.’
- (14) B-
 bas hallaʔ mse:fra, so **b-teje** bi July. [11/L/F/25/19:47]
 but now travelling.3sf so come.3sf in July
 ‘But now she is abroad, so she will come back in July.’

The expression of future temporal reference in colloquial Arabic has not been previously examined from a variationist sociolinguistic perspective. As discussed in Chapter 2, the main issue in the literature on this variable in Arabic concerns the lack of consensus on the linguistic contexts which favour the selection of one exponent of futurity over the other (see e.g., Jarad 2013, 2014; Mitchell and El-Hassan 1994). As we will see below (see section 1.5), the comparative methodological framework and analytical approach adopted in the present study are ideally suited to circumvent the shortcomings observed in the previous literature on future temporal reference (and verbal negation) in spoken Arabic and to identify linguistic change in PA.

1.3 Research questions

The key research questions that this study aims to answer are:

- i) What are the linguistic and social constraints that determine the variable expression of the three targeted variables in contemporary varieties of PA and LA?
- ii) What evidence, gauged from linguistic constraints on variant selection, is there that PA may be converging on LA, and what are the social motivations for convergence?

- iii) What are the linguistic and social constraints that condition the variable expression of the three targeted variables in antecedent states of PA and LA, and how can earlier stages of the varieties be used to illuminate patterns of variation in the contemporary varieties?

The following sections explain how the synchronic and diachronic data sources and the theoretical framework which underpin the present study are used to answer these questions and to detect linguistic change in PA as spoken in Beirut.

1.4 Contribution of the data sources to assessing contact-induced change in PA

The synchronic data comprise around 45 hours of spontaneous speech I recorded through sociolinguistic interviews with 39 PA and 27 LA speakers in Beirut between 2016 and 2019, generating 7,671 tokens representing the three targeted variables. The LA and PA synchronic datasets comprise narratives of personal experience as well as other miscellaneous topics. As such, the synchronic datasets include community-based vernacular speech, which is the optimal speech style for the analysis of the orderly heterogeneity characteristic of natural speech (Labov 1972c; Weinreich, Labov, and Herzog 1968). Furthermore, the substantial amounts of everyday speech collected for this study ensured that the targeted variables occurred frequently enough to provide sufficient data for quantitative and accountable analysis (see Milroy and Gordon 2003:23). The presence of a comparative baseline from Lebanese Arabic is particularly important for situating the usage patterns emerging from Palestinian speakers with respect to those emerging from Lebanese speakers. As Poplack and Levey (2010:398, 406) explain, one of the requirements for demonstrating contact-induced change is systematic quantitative comparison of the structure of the contact variety with that of the presumed source variety of the change. In previous studies on contact between LA and PA in Lebanon, this prerequisite for

establishing convergent change in PA was not met because there was no systematic quantitative comparison of the underlying grammars of PA and LA (Fityan 1981; Hennessey 2011).¹⁵

In terms of social factors, speakers in the synchronic datasets were stratified by sex, age (or immigrant generation for Palestinians), level of education and place of residence for Palestinians only (distinguishing speakers living inside versus outside camps). The incorporation of multiple extra-linguistic factors into the analysis allows us to examine the social embedding of patterns of variation and change in PA, as well as the possible social motivations for convergence (or divergence). Coding for speaker age in particular was important for incorporating an apparent-time component into the present investigation. This allows us to compare the speech patterns of younger and older Palestinians in order to examine whether there is change in PA, and whether younger speakers are leading any change detected.

In addition to an apparent-time component, it was necessary to factor a real-time component into the analysis in order to assess linguistic change in PA. Constituting indispensable adjuncts to the contemporary datasets, this study incorporates diachronic data sources representing each variety. The first diachronic data source is a sub-set of the *Palestinian Oral History Archive* (POHA),¹⁶ comprising audio-visual interviews recorded between 1995 and 2006 with 80 first-generation (older) Palestinians in Lebanon, who were born in pre-war Palestine between 1897 and 1948. These interviews include narratives of personal experience, especially relating to the speakers' lives before the 1948 Arab–Israeli war. Accordingly, they are similar in nature to the sociolinguistic interviews recorded with speakers in the synchronic datasets, meaning that this diachronic vernacular benchmark can be used to draw systematic

¹⁵ Hennessey (2011) does make reference to variants characteristic of LA, but he does not mention that he has collected speech data from LA speakers, and his study does not include systematic quantitative comparisons across LA and PA speech patterns.

¹⁶ The Palestinian Oral History Archive (POHA) can be found at: <https://libraries.aub.edu.lb/poha/>

comparisons with the PA synchronic dataset. The availability of socio-demographic information about the speakers (including sex, education and place of residence) also renders the POHA a valuable dataset because it allows us to uncover the embedding of changes not only in the linguistic structure, but also in the social structure (i.e., the social distribution of linguistic features in the past) (see Auer et al. 2015:4, 6; Bergs 2012:96). The second diachronic data source is the *Lebanese Popular Theatre Corpus* (LPTC), which is based on 34 televised plays dating to 1960–1983. Despite potential caveats in using scripted discourse as a diachronic speech surrogate (see Chapter 3), the plays selected are a valuable resource for tapping into Lebanese Arabic spoken in past decades because they were meant for popular consumption and they were performed in LA.

By utilising diachronic data sources, the present study attempts to address what Lucas and Manfredi (2020:22) describe as a fundamental problem in inferences of linguistic change in spoken Arabic, namely that “there is simply not enough sociohistorical information available to be able to infer with confidence what precise mechanisms underlie the changes in question.” Indeed, because linguistic change originates in vernacular speech, the ideal historical data sources would be recordings of spontaneous speech from the past. In the present study, such recordings were located in the POHA to reconstruct spoken PA in the recent past. It is also possible to resort to diachronic surrogates of orality, such as popular theatre, to reconstruct antecedent stages of LA too, following previous studies on languages other than Arabic (e.g., Digesto 2019; Poplack and Malvar 2007). As the subsequent chapters will show, these diachronic datasets allow us to elucidate the evolutionary trajectory of the three targeted variable systems (in terms of variant rates and conditioning of variant selection) over several decades in order to identify linguistic change in PA and ascertain its source.

1.5 Contribution of the theoretical framework to assessing contact-induced change in PA

The key tenet of the variationist framework in which the present study is embedded is that variability is a hallmark of speech (Poplack 2021:46), and this variability is characterized by structured heterogeneity (Weinreich et al. 1968:99–100). This means that the same linguistic function can be expressed in different forms (Poplack and Tagliamonte 2001:7), and the choice between two or more alternatives in natural discourse is typically constrained by multiple independent social and linguistic variables (Bayley 2013). The structure (or grammar) underlying this heterogeneity in spontaneous speech can be inferred from the distribution and conditioning of competing variants in natural speech (Poplack and Levey 2010:398) through quantitative modelling (Tagliamonte 2006; Young and Bayley 1996:254).

The conceptual and analytical framework of variation theory allows us to transcend the shortcomings in earlier (socio)linguistic research on spoken Arabic in several ways. By making use of a systematic, empirically accountable approach, we can extract the three targeted variables from vernacular speech in adherence to the Principle of Accountability (Labov 1969, 1972b), including “every case where the variable element occurs in the relevant environments as we have defined them” (Labov 1969:737–38). This allows us to determine the factors within the linguistic and extra-linguistic environments that contribute to variant selection in the three targeted variables in natural speech. As detailed in Chapter 2, several previous works on spoken Arabic (especially those targeting grammatical variation) are mainly based on intuited or elicited materials (e.g., Jarad 2013, 2014; Lucas 2010). Even when they do rely on natural speech, they adopt non-quantitative approaches and/or approaches that do not respect the Principle of Accountability (e.g., Khairallah 2014). In some cases, scholars also pay greater heed to theory rather than to the data, and they cherry-pick data which back up their theories (e.g., Lucas

2010).¹⁷ As such, the importance of the variationist framework is that it enables us to test previous claims about (linguistic and social) constraints on variant choice against usage-based data.

Another key aspect of the variationist framework is that it draws on the comparative method of historical linguistics (e.g., Baldi 1990; Campbell 1998; Hoenigswald 1960; Meillet 1967; see Poplack and Levey 2010) by comparing usage patterns in relevant benchmarks. In other words, within this theoretical framework, the key diagnostic for assessing contact-induced change is not simply the presence of a candidate for change in a (minority) variety, but its conditioning in discourse, as inferred from multivariate analysis of the environmental constraints governing a variable system. This allows the analyst to uncover structural similarities or differences between different benchmarks (or speaker groups) in terms of the conditioning of variant selection. If structural parallels in the conditioning of variant choice are uncovered between different benchmarks, we can infer that there is a relationship between their underlying grammars. On the other hand, where there are structural differences, “we have grounds for concluding that the phenomena in question belong to different linguistic systems” (Tagliamonte 2013:131).

In the present study, examining whether any change has taken place in PA and whether

¹⁷ As detailed in Chapter 2, in Lucas’s (2010) study on verbal negation in Palestinian Arabic, he used an acceptability judgment questionnaire to collect data from Palestinian speakers from different locations in Israel, the West Bank and the Golan Heights (Lucas 2010:171). He found that the majority of informants generally rejected negating past tense verbs with *-š* (except for one informant) (see Lucas 2010:173). The majority of informants also judged the pseudo-verb *šand-* ‘have’ with *-š* as ‘odd’ or ‘impossible’ (Lucas 2010:174). However, as Wilmsen (2014:102) states, Lucas (2010) ignores or discounts “the evidence from the varieties of Levantine Arabic writ large, wherein other researchers have documented variability in the phenomenon, including negating past tense verbs and the pseudo-verb *‘and* with a post-positive *-š* alone” (see e.g., Bauer 1926; Cowell [1964] 2005; Driver 1925; Feghali 1928; Fleisch 1974; Thackston 1996). In addition, Wilmsen (2014:103) notes that some speakers in Lucas’s (2010) sample accepted constructions that he dismissed as inadmissible or ungrammatical. “Those speakers may have had exposure to—or may themselves have been speakers of—regional or older dialects retaining the variants that [Lucas (2010)] dismissed as being unacceptable” (Wilmsen 2014:103). Lucas (2010) may have ignored evidence from previous literature, as well as some of his informants’ responses in order to support theory-based claims he made earlier (Lucas 2007:423–26) regarding the ‘ungrammaticality’ of *-š* with *šand-* ‘have’ or with past tense verbs (Lucas 2010:167).

this change is internal or contact-induced requires systematic comparisons of variant rates and, more importantly, the linguistic conditioning associated with the three targeted variables in contact PA (synchronic speech data) with LA (the presumed source of the change) and with the pre-contact stage of PA (see Poplack and Levey 2010:406). As Poplack and Levey (2010:400) explain, assessing (contact-induced) change is based on comparing across relevant benchmarks the constraint hierarchy, or “the configuration of environmental factors affecting the probability that a given variant form will be selected, along with the direction of their effects.” If the hierarchy of constraints conditioning the variable occurrence of a candidate for change in the contemporary contact PA variety is the same as that of its pre-contact precursor, while differing from that of LA, no change has taken place in PA. If the candidate for change in contact PA features a constraint hierarchy different from those of both its pre-contact precursor and LA, we can infer that change has occurred in PA, but not one that is contact-induced. Only if a candidate for change in contact PA features a constraint hierarchy different from that of its pre-contact precursor, but parallel to that of LA, can we infer that there is contact-induced change in PA due to the influence of LA (see Poplack and Levey 2010:400–401).

1.6 Road-map of the thesis

The remainder of this thesis is organized as follows. In Chapter 2, I review the relevant literature on the three targeted variables and critically evaluate previous works, and I explain how this study builds on the previous literature and addresses major methodological and analytical issues. In Chapter 3, I introduce the data sources for this study, including a discussion of the sample constitution and data collection methods. I also discuss the theoretical framework and methodology used for data analysis. Chapter 4 provides a detailed quantitative analysis of raising of /a:/ to [e:] in word-medial position in the synchronic and diachronic datasets. Chapters 5 and 6

present the quantitative results for verbal negation and future temporal reference, respectively.

Chapter 7 concludes the thesis with a summary of the results; discusses their wider implications and the limitations of the study; and offers directions for future research.

2 Literature Review

This chapter is divided into four main sections: Section 2.1 reviews the literature on raising of /a:/ to [e:] (*imala*). Section 2.2 presents a literature review on verbal negation, and section 2.3 focuses on the variable expression of future temporal reference. Section 2.4 critically evaluates previous treatments of the three variables and explains how the present study builds on the previous literature through methodological and analytical enhancements.

2.1 Raising of /a:/ to [e:]

Even though advances in Arabic sociolinguistics are widely based on the study of phonological variables, these advances are restricted to a small number of variables (Miller 2007:2), especially consonants (Al-Wer 2002:63–64), such as (q), (dʒ), (θ), and (ð) (see e.g., Abdel-Jawad 1981, 1986; Al-Khatib 1988; Al-Wer 1991, 1999; Haeri 1991; Holes 1983, 1986). On the other hand, there is less sociolinguistic research on vocalic variation in spoken Arabic (Al-Wer 2007b:68; Cotter 2020:574, 576), especially on variable vowel raising, or *imala* (see e.g., Abo Mokh and Davis 2020; Al-Wer 2002, 2007b, 2020; Cotter 2013; Habib 2012; Hennessey 2011).

The term *imala*, which literally means ‘inclining, bending to’ (Owens 2006:197; see also Levin 2006:311) is used by medieval Arabic grammarians to refer to the fronting and raising of Old Arabic /a:/ toward /i:/, and the old short vowel /a/ toward /i/ (Levin 1992:74, 2006:311). Owens (2006:197) and Habib (2012:54) explain that in Classical Arabic, *imala* refers to the change from /a:/ to /e:/ or /ie/ in the context of a conditioning high vowel /i/, e.g., *kila:b* > *kilieb* ‘dogs.’ As such, in *imala*, the phonetic configuration of /a:/ is made to resemble and approach that of /i/ (Owens 2006:197–98).

There are several reasons why *imala* should attract academic interest. First, it is an example of long-standing variability, as it is a feature of both Old Arabic and some modern Arabic dialects (Owens 1993:251). Second, it is attested in certain Arabic dialects but not others, so it provides fertile territory for comparative studies of contact-induced change if speakers of one Arabic variety without *imala* come into sustained and regular contact with speakers of another Arabic variety with *imala* (see e.g., Al-Wer 2002, 2007b, 2020; Cotter 2013).

Imala can occur word-medially, e.g., *sa:xin* > *se:xin* ‘hot’ (Habib 2012:53), or word-finally (at the end of feminine nouns and adjectives that are morphologically marked by the ending [a] in Arabic) (see Barkat 2008:678), e.g., *madrasa* > *madrase* ‘school’ (Versteegh 2009:482; see also Al-Wer et al. 2020:11; Barkat 2008:678; Levin 2006:311; Owens 2006:197). Additionally, there is *imala* of the long vowel /a:/ (see Al-Hashmi 2016:300; Habib 2012:54; Hellmuth 2013:53), e.g., *ka:n* > *ke:n* ‘he/it was’, and *imala* of the short vowel /a/ (see Habib 2012:54; Owens 2006:197, 209–12),¹⁸ e.g., *ʔakal* > *ʔakel* ‘he ate’ (Habib 2012:53). In this study, I focus on *imala* of the long vowel /a:/ in word-medial position because it is reportedly a conflict site between LA and PA as spoken in Beirut (Hennessey 2011), meaning that this is a form “that differs functionally and/or structurally and/or quantitatively across the varieties in question” (Poplack and Meechan 1998:132). Word-medial (long) *imala* is reportedly not attested in PA as spoken in Beirut (Hennessey 2011; see also Naïm 2005, 2011; Shahin 2007, 2011), e.g., *ka:mil* ‘complete’ (Hennessey 2011:61), while it is reported to be a characteristic of LA, e.g., *ke:mil* ‘complete’ (Hennessey 2011:61). This conflict site can be used as a diagnostic of contact-induced change in PA by examining whether Palestinian speakers are converging on Lebanese speakers in terms of the use of word-medial *imala* (i.e., whether Palestinians in Beirut use the

¹⁸ According to Habib (2012:52), the phonetic quality of short *imala* “is assumed to be the cardinal vowel [e], because it is slightly higher and longer than the [ɛ] sound in the English word *bed*.”

raised [e:] variant at rates equivalent to those found in LA speech, and crucially, whether Palestinians are reproducing the same linguistic constraints on variant selection as those operating in LA).

Several previous accounts of *imala* are limited to describing the constraints on its use based on normative prescriptions (see e.g., Al-Nassir 1993; Elihay 2004; Naïm 2011; Owens 1993, 2006; Retsö 2011). Conspicuously infrequent are studies on *imala* that have privileged natural speech data over anecdotal reports (e.g., Al-Wer 2002, 2007b, 2020; Cotter 2013; Habib 2012; Hennessey 2011). The vast majority of such studies have additionally focused on *imala* in word-final position only (Al-Wer 2002, 2007b, 2020; Cotter 2013), with the result that less is known about the acoustic properties of word-medial *imala*, as well as the linguistic and social constraints on variant selection in different Arabic dialects where it is found.

In the ensuing sections, I present a brief overview of *imala* in Old Arabic in section 2.1.1, where the reported constraints on the occurrence of *imala* are claimed to be applicable to modern Arabic dialects (Owens 2006). In section 2.1.2, I focus on *imala* in spoken Arabic varieties. In section 2.1.3, I discuss previous accounts of *imala* in PA and LA, and in section 2.1.4, I discuss the social conditioning of *imala* in PA as spoken in Beirut (Hennessey 2011).

2.1.1 *Imala* in Old Arabic

Sibawayh, a Persian linguist who had a prominent influence on Arabic linguistics, was the first to introduce the term *imala* in his treatise *Al-Kitab* (lit. ‘the book’), written in the eighth century A.D.¹⁹ (Al-Nassir 1985:1). The primary sources he used for his examination of the Arabic language include recitations of the Qur’an and poetry, as well as speakers of different Arabic

¹⁹ There is no indication about the year when the book was published, but it seems that it was finished between 170 and 180 A.H. (786-796 A.D.) (Al-Nassir 1985:10).

dialects (Al-Nassir 1985:12). Sibawayh's description of *imala* was apparently based on the linguistic situation in his time in Al-Basra and its surroundings in southern Iraq (Levin 2006:311). He acknowledges that the use of *imala* is not categorical; rather, he notes its occurrence and non-occurrence as dialectal variation across varieties of Arabic when the consonants preceding or following /a:/ are the same (Al-Nassir 1985:161).

Sibawayh identifies several constraints on the use of *imala*, and these constraints are mainly related to the phonological segments preceding and following /a:/. He observes that *imala* occurs in the environment of the semi-vowel *ya* or /i/ in the preceding or following context (Al-Nassir 1985:163; Habib 2012; Owens 2006:197, 203), e.g., *kile:b* 'dogs'; *mase:jid* 'mosques' (Owens 2006: 197); *bayye:f* 'vendor/seller' (Al-Nassir 1985:167; Owens 1993:252). Sibawayh later observed that there is variation, where *imala* may not be realized, even in the context of /i/, or where *imala* may occur without the presence of /i/ (Owens 2006:204).

Sibawayh claims that *imala* cannot occur if one of the following consonants (called *huruf al-mustaʿliya* 'elevated or raised consonants') (see Hellmuth 2013:53) is adjacent to /a:/: emphatics [ṭ ḍ ṣ ẓ], uvulars [q ɣ x] and gutturals [ħ ʕ ʔ] (Al-Nassir 1985:169; Hellmuth 2013:53; Levin 2006:312; Owens 1993:252, 2006:200–201) (see Table 2.1 below). This is attributed to pharyngeal articulation (Al-Nassir 1985:170; Habib 2012:55; Owens 2006:200; Shahin 2011). As Hellmuth (2013:53) states, these consonants "influence the phonetic realization of neighboring segments, both vowels and consonants." Hellmuth (2013:53) explains that the most salient effect of these consonants is backing/lowering of immediately adjacent vowels, e.g., [ṭa:b]

‘he recovered’.²⁰ These consonants are said to disfavour *imala* in the following and preceding phonological context (see e.g., Davis 1995; Owens 2006).

Sibawayh also observes that /r/ disfavors *imala* in the preceding segment, but *imala* can occur if /r/ is after /a:/, e.g., *ḥime:r-i-k* ‘your donkey’ (Owens 2006:202; see also Al-Nassir 1985:174–75). But Al-Nassir (1985:169) claims that /r/ after /a:/ does not always allow *imala*, as in *da:ra* ‘he turned.’ He explains that the root of this verb is *dwr* ‘turning,’ and it does not have an /i/ in the paradigm (i.e., *dur-tu* ‘I turned,’ but not **dir-tu*), so the absence of /i/ in the first-person verb form does not allow *imala*.

Imala is said to be favoured if /a:/ is adjacent to front consonants (like palatal or labial consonants), e.g., *ba:b* > *be:b* ‘door’ (Levin 1998:77–80, 2006:312; Owens 1993:252, 2006:202). This type of *imala* was only noted by Levin (1998), but it was not noted by Sibawayh. Owens (2006:202) claims that this type is found in modern Arabic dialects, and Levin (2006:313–14) specifies that this type of *imala* is found in Lebanese Arabic, the Druze dialects of the Horan and the Golan, and Palmyra (or Tadmur in Arabic, which is an ancient Semitic city in Homs Governorate, Syria).

Sibawayh’s work remains a valuable resource because it shows that *imala* is a long-standing feature of spoken Arabic with considerable time-depth. His account is important for this study’s examination of *imala* because some constraints he mentions are relevant to modern Arabic dialects, especially with regards to the consonants that disfavour *imala* in the preceding segment (Habib 2012; Owens 2006), so these constraints are operationalized and systematically tested in this study as well as constraints associated with the following phonological context.

²⁰ Hellmuth (2013:53) explains that the influence of these consonants on neighbouring segments is referred to as ‘emphasis spread’, the domain of which “is nonlocal in character, reaching the entire word in some spoken dialects and even beyond it into adjacent words in some cases” (See Chapter 4 for more details on ‘emphasis spread’).

2.1.2 *Imala* in modern dialects of Arabic

Imala in word-medial position reportedly occurs in several Arabic varieties (see Table 2.2), such as Northern Syrian Arabic (Habib 2012) and Lebanese Arabic (Hennessey 2011; Naïm 2005), including Beiruti Arabic (Hennessey 2011), as well as in the Arabic spoken in Mesopotamia (Iraq and Turkey) (Owens 1993:253–54). Levin (1998:83) notes that *imala* also occurs in Christian Baghdadi Arabic (see also Abu-Haidar 2005:224) and Jewish Baghdadi Arabic (see also Mansour 2005:234).

Previous accounts of *imala* claim that the degree of raising and the linguistic contexts favourable to the use of *imala* vary across varieties of Arabic in which it is found (Al-Khatib 1988:261; see also Barkat 2008:678; Cotter 2020:574; Owens 1993:253–54, 2006:212). However, there are several consonants that allegedly disfavour *imala* in the preceding phonological context (Al-Nassir 1993; Cantineau 1960:96–97; Habib 2012; Levin 1978, 1992; Owens 2006:200), which are broadly the same in several Arabic varieties. As mentioned in the previous section, these consonants are gutturals and emphatics (or primary pharyngeals) (see Shahin 2011), as well as uvulars (dorsal primary pharyngeals) (see Habib 2012:55) and the secondary pharyngeal [r], as summarized in Table 2.1.

Table 2.1: Consonants reported to disfavour vowel raising in the preceding phonological context (after Habib 2012:54–55)

Emphatics		Uvulars		Gutturals		Voiced alveolar trill	
[t]	hi:ʔa:n ‘walls’	[q]	niqa:š ‘discussion’	[h]	ha:wil ‘try’	[r]	dira:se ‘a study’
[d]	ramaɖa:n (Month of Ramadan)	[ɣ]	luɣa:t ‘languages’	[ʕ]	ʕa:mmiyye ‘colloquial’		
[s]	ša:r ‘he/it became’	[x]	xa:lte ‘my aunt’	[ʔ]	riʔa:se ‘presidency’		
[z]	maʔa:hir ‘appearances’						

It is beyond the scope of this chapter to discuss the acoustic properties and the reported linguistic constraints on *imala* across Arabic dialects, but Table 2.2 below summarizes the degree of vowel raising reportedly found in word-medial and word-final positions across several Arabic varieties. The shading in the table below means that, to my knowledge, there is no information about word-medial or word-final *imala* (its presence or lack thereof, or its acoustic properties) in the available literature on the dialect in question.

Table 2.2: Default *imala* variants in word-medial and word-final positions across some Arabic varieties²¹

Variety	WORD-MEDIAL POSITION				WORD-FINAL POSITION		
	[a:] > [æ:]	[a:] > [e:]/[ē]	[a:] > [e:]/[o:]	[a:] > [i:]/[ī]	[a] (no <i>imala</i>)	[a] > [e]/[ē]	[a] > [i]
Palestinian Arabic (Al-Wer 2020:561; Hennessey 2011:59; Shahin 2007:530; Versteegh 2009:482, 483)	√ ²²					√	
Lebanese Arabic (Beirut) (Barkat 2008:678; Borg 2005:538; Hennessey 2011:59; Levin 2006:314; Naïm 2011:925)		√				√	
Lebanese (Tripoli) (Al-Wer 2007a:607)			√				
Lebanese (Shiite community in South Lebanon) (Naïm 2011:925)				√			
Jordanian Arabic (Al-Wer 2020:561)					√ ²³		
Syrian Dialects (Damascus;			√				

²¹ The realizations of *imala* in different Arabic dialects included in the table are the default choice, meaning that these are the typical realizations of *imala* in the dialects where it is found unless [a] or [a:] appears in the environment of any blocking segment (i.e., velarized, emphatic, and pharyngeal sounds).

²² As I mention later in this chapter, Al-Wer (2007b:70) claims that /a:/ in word-medial position in Jerusalem Arabic ranges from raised /æ:/ ~ /ɛ:/ ~ /e:/, e.g., /ʒɛ:ʒ/ or /ʒe:ʒ/ ‘chicken’; /mɛ:xdi:n/ ‘we/they have taken’, though she does not present evidence for her claim through quantitative reasoning based on a representative sample of speakers.

²³ [a] > low vowel [a], except after coronal sounds, where [a] is raised to [ɛ] (Al-Wer 2020:561).

Homs) (Habib 2012:51; see also Barkat 2008:678)							
Syrian (Sukhne, northeast of Palmyra in Homs) (Barkat 2008:678; Borg 2005:538; Levin 2006:314; Naïm 2011:925; Palva 2005:607)		√		√			
Syrian (Aleppo) (Al-Wer 2007a:607)		√					
Syrian (Homs) (Barkat 2008:678)							√
Syrian (Damascus; Horan) (Barkat 2008:678–79)						√	
Algerian and Moroccan Dialects (Barkat 2008:678)					√		
Egyptian Arabic (Cairo) (Barkat 2008:681; Woidich 2005a:300, 2005b:324)	√						
Coptic Arabic (Egypt) (Richter 2005:497)						√	
Western Delta, Lower Egypt (Wilmsen 2006:3)							√ ²⁴
Jewish Baghdadi Arabic (Jastrow 2006:417; Levin 2006:313; Mansour 2005:234; Palva 2005:607; Watson 2005:23)		√		√			
Christian Baghdadi Arabic (Abu-Haidar 2005:224; Jastrow 2006:417; Levin 2006:313)		√					√
Bedouin Negev Arabic (Henkin 2007:362)		√					
Spoken Arabic in Tigris, Mesopotamia (Palva 2005:607)				√			
Bahraini Arabic (some rural Baḥārna dialects) (Holes 2005:243)				√			
Antiochia Arabic (spoken in		√	√ ²⁵	√ ²⁶			

²⁴ [a:] > [i] (pausal *imala*, e.g., *kalbe* ~ *kalbi*# ‘bitch’) (Wilmsen 2006:3; see also Levin 2006:314).

Antakya, Turkey) [Jewish dialects, the Sunni dialect of Reyhanlı, and dialects of two Alawi villages] (Arnold 2005:113)							
Cilician Arabic (spoken in Cilicia, Turkey) (Procházka 2005:390)		√					

2.1.3 *Imala* in Palestinian and Lebanese Arabic

2.1.3.1 Palestinian Arabic

Palestinian Arabic as spoken in Beirut is reported not to have word-medial *imala* (Hennessey 2011:59; see also Naïm 2005:277, 2011:925; Versteegh 2009:482–83). Naïm (2005:277) claims that PA is characterized by /ā/ or [a:] in word-medial position, and Hennessey (2011:59) observes in his speech data collected from Palestinians in Beirut that PA is characterized by the lowered and backed [æ:]. Rosenhouse (2009:482) explains that in Jerusalem Arabic spoken by Palestinian-heritage speakers,²⁷ the vowel /a/ has fronted and back allophones, depending on the environment: the forward allophone [æ] is used in the context of labial and coronal consonants; the back allophone [a] is used in the context of laryngeal and pharyngeal consonants; and the higher back allophone [ɑ] or [ɔ] is used in the context of emphatic consonants (see also Hennessey 2011:64). In contrast to these accounts, Al-Wer (2007b:70) claims that /a:/ in word-medial position in Jerusalem Arabic ranges from raised /æ:/ ~ /ɛ:/ ~ /e:/, e.g., /ʒɛ:ʒ / or /ʒe:ʒ/ ‘chicken’; /mɛ:xdi:n/ ‘we/they have taken’. As such, she claims that, at least in Jerusalem Arabic,

²⁵ [a:] > [o:] in Alawi village dialects (in the vicinity of emphatic consonants), e.g., *fillāh* ‘farmer’, but *xiyyōt* ‘tailor’ (Arnold 2005:113).

²⁶ In some dialects of Christians and Alawites (Arnold 2005:113).

²⁷ In particular, Rosenhouse (2009:491) states that “Jerusalem Arabic speakers belong to at least one of three large ethnic and religious groups: Muslim, Christian, and Jewish.” In other words, Jerusalem Arabic “includes three communal subdialects of the Muslims, Christians, and Jews, although the differences among them are not large” (Versteegh 2009:481).

word-medial *imala* is attested. However, she does not present quantitative evidence for her claim. Rosenhouse (2009:482–83) acknowledges that there are rare cases of word-medial *imala* in Jerusalem Arabic, and these are found in the words *se:ʕa* ‘hour, watch’ and *mbe:riḥ* ‘yesterday,’ but she does not explain why only these two words allow word-medial *imala*, or whether she has actually observed these words in natural speech (see also Al-Wer 2007b:69).²⁸

Palestinian Arabic does, however, allow word-final *imala* at the end of feminine nouns and adjectives (see Barkat 2008:678), where the default choice is [e] (Al-Wer 2020:561; Naïm 2011:925; Shahin 2007:530; Versteegh 2009:482, 483). However, word-final *imala* is reportedly blocked by consonants like /r/, e.g., *sayyara* ‘car’ (Shahin 2007:530), as well as by other pharyngeal consonants, e.g., *ja:mʕa* ‘university’; *ʕulṭa* ‘authority’ (Al-Wer 2020:561).

2.1.3.2 Lebanese Arabic

Lebanese Arabic is reportedly characterized by word-medial *imala*, as well as word-final *imala* (Barkat 2008:678; Borg 2005:538; Hennessey 2011:59; Levin 2006:314; Naïm 2011:925). Word-medial *imala* in Beiruti Arabic in particular is characterized by the raising of [a:] to [e:] (Barkat 2008:678), while word-final *imala* is characterized by the raising of [a] to [e] (Barkat 2008:681; Naïm 2011:925).

Several scholars provide a list of linguistic constraints on *imala* in word-medial position in LA as spoken in Beirut. Levin (2006:314) states that in the dialect of Beirut, *imala* occurs when it is surrounded by two front non-emphatic consonants, e.g., *ke:n* ‘he/it was’ (see also Hennessey 2011:61). On the other hand, the consonants summarized in Table 2.1 are said to disfavour *imala* in Beiruti Arabic when they occur before or after /a:/, e.g., *ʕa:lib* ‘student;’ *bla:t* ‘floor tiles’ (Hennessey 2011:64). However, while *imala* is reportedly not favoured if /r/ is in the

²⁸ Hennessey (2011:63) actually observed that *mba:riḥ* ‘yesterday’ (with no *imala*) is common in the speech of his Palestinian interviewees in Beirut.

preceding segment (Hennessey 2011:61), it is said to be favoured when /r/ is in the following segment, but only if /r/ is followed by /i/, e.g., *še:riʃ* ‘street’ (Hennessey 2011:62). But if /r/ is not followed by /i/, *imala* is said not to occur, e.g., *ba:rak* ‘he blessed’ (Hennessey 2011:62–63).

2.1.4 Social conditioning of *imala* in PA spoken in Beirut

One of the few studies which focused on word-medial *imala* is Hennessey’s (2011) account of its use by Palestinians in Beirut. Hennessey’s aim was to investigate the linguistic assimilation of Palestinians to the Lebanese speech community in Beirut. He focused on the social conditioning of word-medial *imala* in the speech of 35 Palestinian speakers, some of whom live in refugee camps. Hennessey (2011:1, 12, 19–20) explains that dense, multiplex social networks among Palestinians in refugee camps, as well as the relative spatial and social distance of these camps from Lebanese society may lead Palestinian camp-dwellers to retain PA vernacular patterns. On the other hand, as Hennessey (2011:147–48) argues, “[w]eakened ties to Palestinian communities and strengthened ties to Lebanese communities [...] encourage adoption of Lebanese speech features.” In other words, for some Palestinians, especially those who live outside camps and who may be more integrated into the Lebanese society (Hennessey 2011:145), the lack of the linguistic input that refugee camps provide and exposure to Lebanese speech patterns could prompt accommodation to LA (Hennessey 2011:127–28).

Based on occurrences and non-occurrences of *imala* in word-medial position, Hennessey found that *imala* is favoured by some groups of speakers more than others. In terms of sex, results showed that the rate of *imala* in the speech of Palestinian females is almost double that in the speech of males. This is evidence of sex-differentiation in speech, and it is in line with the sociolinguistic gender pattern, which refers to the propensity of females to use linguistic forms that are associated with prestige (see e.g., Armstrong 1998:139; Chambers and Trudgill

2004:84–85; Milroy, Milroy, and Hartley 1994:352; Trudgill 1972:180). The prestige variant in this case is the *imala* variant (or the [e:] variant) because it is characteristic of metropolitan LA (i.e., the Lebanese dialect of Beirut; see Hennessey 2011:16, 92; Procházka 2020:84; Versteegh 2001:153).

In terms of age, Hennessey found that *imala* is favoured by those between the ages of 25 and 49. Older Palestinians (50+), some of whom were born in pre-war Palestine and some in Lebanon (Hennessey 2011:95), tend to retain their Palestinian speech features (Hennessey 2011:94; see also Fityan 1981:iii). Hennessey cites an earlier study on the linguistic patterns of Palestinians in a refugee camp outside Beirut (Fityan 1981), and explains, following Fityan (1981), that older Palestinians may have had “no social aspirations which might stimulate them to modify their speech to fit in the wider community, the Lebanese community” (Fityan 1981:48), while the younger Palestinians’ “concern about the social acceptability of speech, and their aspirations for a better social status led them to adapt their speech behavior [to Lebanese speech patterns]” (Fityan 1981:49).

Education also plays a role in the use of *imala*, which is reported to be more frequently used by Palestinians with post-secondary education than those with primary or secondary education (Hennessey 2011:111). But Hennessey (2011:111) explains that this finding does not mean that education per se is correlated with the use of *imala*; rather, education is a measure of the degree of access to and integration into the Lebanese society (see Al-Wer 2002:42). Many Palestinians attend UNRWA sponsored schools, some of which are in refugee camps, where there are primarily Palestinian social networks. However, higher education is only accessible to (younger) Palestinians in institutions outside the camps. As such, Palestinians’ engagement in post-secondary education in these Lebanese institutions presupposes some degree of regular

interaction with Lebanese speakers, ostensibly favouring the diffusion of LA features to Palestinian students (Hennessey 2011:110–12). In fact, Hennessey found that one of the major influences on the use of *imala* is the type of social network structure in which speakers are embedded. As mentioned previously, he acknowledges that the camps reinforce Palestinian-centred social networks (Hennessey 2011:131), which has profound effects on speech patterns. He found that the rate of *imala* in the speech of Palestinian speakers outside the camps is almost triple that in the speech of camp dwellers.

Hennessey (2011:130) explains that Palestinians in the camps are free to move in and out of the camps, but not necessarily in terms of permanent residency changes. This means that camp dwellers may still have some degree of integration into the Lebanese society (e.g., some camp residents may have Lebanese mothers or wives, others may attend Lebanese schools/universities outside the camps, and others used to live outside the camps before they moved to the camps) (see Hennessey 2011:138–39). Hennessey (2011:140) explains that Palestinian camp-dwellers with less access to and integration into the Lebanese society in particular retain Palestinian speech patterns, while those with higher levels of exposure or with deep-rooted connections to the Lebanese society tend to show “mixed or Lebanese-leaning speech tendencies” (Hennessey 2011:140, 145). As such, Hennessey (2011) found that among Palestinian camp-dwellers, it is not only the dense social networks characteristic of the refugee camps which may contribute to the maintenance of Palestinian speech patterns, but also the restricted access of some of these camp residents to the Lebanese society (Hennessey 2011:140).

Despite Hennessey’s (2011) valuable contribution in shedding light on the contact between LA and PA in Beirut and the agents of change, his study suffers from a number of problems, mainly in terms of how he infers (contact-induced) change in PA. This is because his

study relies on overall usage rates and focuses on social factors only, without examining the linguistic conditioning of variant selection (i.e., examining, as mentioned previously, whether Palestinians in Beirut are reproducing the same linguistic constraints on variant selection as those found in LA). I return to these issues in the critical summary at the end of this chapter.

2.2 Verbal negation

Negation in spoken Arabic has attracted substantial scholarly attention (e.g., Alqassas 2012, 2019; Aoun, Benmamoun, and Choueiri 2010; Benmamoun 2000; Benmamoun et al. 2014; Brustad 2000; Hoyt 2006, 2010; Lucas 2007, 2010, 2015; Obler 1975; Ouhalla 1993; Palva 2004; Shlonsky 1997; Soltan 2007, 2014; Wilmsen 2013, 2014; Woidich 1968). Much previous research has examined verbal negation in spoken Arabic from the perspective of generative syntax, often using fictional examples rather than ones extracted from spontaneous speech (e.g., Alqassas 2012; Aoun et al. 2010; Benmamoun 1992, 2000; Fassi Fehri 1993; Hoyt 2006, 2010). Other studies on verbal negation in Arabic have relied on acceptability judgment questionnaires to examine linguistic constraints on variant choice (Lucas 2010), while only a few studies used natural speech data to examine verbal negation (e.g., Khairallah 2014). Several studies have also focused on the historical development of negation, specifically, the emergence of the post-verbal negative enclitic *-š*, which appears in many vernacular varieties of Arabic (Wilmsen 2014:1). This variant has generated controversy in the pertinent literature, with Lucas (2007, 2015) arguing that it developed from the noun *šayʔ* ‘thing,’ while Wilmsen (2013, 2014, 2020) argues that it developed from the polar interrogative marker *ši*.

The remainder of this section is organized as follows: Section 2.2.1 briefly introduces verbal negation in Standard Arabic. Section 2.2.2 discusses verbal negation in spoken Arabic,

and section 2.2.3 reviews studies on verbal negation in Lebanese Arabic (Khairallah 2014) and Palestinian Arabic (Lucas 2010).

2.2.1 Overview of verbal negation in Standard Arabic

Standard Arabic (SA) has five negative particles (Moutaouakil 1993): *la*, *lan*, *lam*, *laysa* and *ma*. The first three particles are tense-marked (Farwanah 2021:169). Soltan (2021:387) explains that “[t]ensed negation refers to structures where the negation particle itself expresses not only the negative import of the clause, but marks its tense property as well. In such contexts, the verb appears in the non-tensed imperfective form [...]” The particle *la* is used with present tense verbs (Aoun et al. 2010:110; Farwanah 2021:169; Soltan 2021:387), e.g., *la na-qrāʔ-u* ‘we do not read’ (Soltan 2021:388), and it can function as a negative imperative, e.g., *la talʕab* ‘don’t play’ (Mousa 2019:76). The negative particle *lan* indicates negation of a future event, e.g., *lan na-qrāʔ-a* ‘we will not read’ (Soltan 2021:387), while *lam* is used for negation of a past event, e.g., *lam na-qrāʔ* ‘we did not read’ (Soltan 2021:388; see also Aoun et al. 2010:110; Farwanah 2021:169). *Laysa* inflects for person, number and gender agreement. It can occur in the context of non-verbal predicates (Aoun et al. 2010:110), e.g., *laysa-t muʕallimat-an* ‘she is not a teacher’ (Aoun et al. 2010:111; see also Soltan 2021:388–89), as well as in present tense sentences containing verbal predicates (Aoun et al. 2010:110), e.g., *laysa xaalid-un yaktubu š-šifr-a* ‘Khalid does not write poetry’ (Moutaouakil 1993:85, as cited in Aoun et al. 2010:111). The fifth negative particle in SA is *ma*, which is also used in modern spoken Arabic dialects (Aoun et al. 2010:111). *Ma* is also found in all ancient Semitic languages, including Classical Arabic in the Qur’an (Lucas 2018:45). This variant shows neither tense nor agreement inflection (Soltan 2021:389), and it is said to be used for predicate and verbal negation, where it can negate past

tense and present tense verbs (Moutaouakil 1993:80–81).²⁹ As we will see in the next section, apart from *ma* (and in some dialects, *la*), other negative particles in SA diverge markedly from those in contemporary Arabic dialects (Aoun et al. 2010:112; Procházka 2021:228).

2.2.2 Overview of verbal negation in Arabic dialects

Several Arabic dialects share the use of the preverbal negator *ma*, such as Syrian Arabic spoken in Damascus (Brustad 2000; Farwaneh 2021:170), Kuwaiti Arabic (Brustad 2000:285), Jordanian Arabic (Alsarayreh 2012:42), Lebanese Arabic (Aoun et al. 2010:96; Farwaneh 2021:170), Najdi Arabic (Benmamoun 2000), Gulf Arabic (Farwaneh 2021:170; Holes 1990), and (mainly urban) Palestinian Arabic (Awwad 1987; Mousa 2019).³⁰

Negation with *ma-š* has been attested in rural varieties spoken in some parts of northern Lebanon, such as in Baskinta, a village in Mount Lebanon (Abu-Haidar 1979:109–10), as well as in varieties spoken across coastal North Africa, such as: Morocco (Harrell 1962:152–56; Woidich 1968), Tunisia (Cohen 1975:268; Wilmsen 2021:257), Algeria (Laraba 1981:147; Souag 2005:166–67), Libya (Owens 1984:157), and Upper Egypt (Brustad 2000:279; Khalafallah 1969; Wilmsen 2021:257).³¹ *Ma-š* has also been attested in (mainly rural) Palestine, parts of Jordan (Awwad 1987; Mousa 2019; Palva 2004; Shahin 2000), parts of Yemen (Behnstedt 1985; Watson 1993), and parts of Oman and on Zanzibar (Reinhardt 1894). Soltan (2021:393) explains that when *(ma)-š* is used for verbal negation, “the host of the negation circumfix can be a full verbal complex that contains not just the verb but its cliticized arguments

²⁹ Some sources also refer to ‘perfective’ and ‘imperfective’, which are the labels given to the two non-periphrastic finite tense forms in Arabic (Lucas 2010:167). Perfective verbs refer to past tense verbs, while imperfective verbs refer to present tense verbs (Lucas 2010:167).

³⁰ Interestingly, Mousa (2019:97) reports that the younger generation of rural Palestinian Arabic speakers (in the Northwest part of the West Bank in Palestine, specifically in the vicinity of the city of Tulkarm) tend to use *ma*, switching to the urban dialect because it is the prestige target.

³¹ Broadly speaking, Wilmsen (2021:257) claims that negation with *ma-š* is characteristic of “the dialects of the North African littoral, throughout Egypt, the southern and highland Levant, some Arabic dialects of southern Mesopotamia, and some of the southern Arabian Peninsula.”

as well”, e.g.:

- (1) **ma**-šuf-t-u-hum-š.
 NEG-saw-1SG-EV-them-NEG
 ‘I did not see them.’ (Cairene Egyptian Arabic; Soltan 2021:393)

In terms of the linguistic constraints on the use of *ma*-š, it is claimed that it cannot negate future tense verbs (2) in some Levantine Arabic varieties, such as in Jordanian (Al-Momani 2011) and Palestinian Arabic (Farwanah 2021:171; Mousa 2019:100–101). Instead, the pre-verbal *ma* used alone can reportedly negate future tense verbs in Levantine Arabic (Farwanah 2021:171).

- (2) *l-walad **ma** rah yenam-š.
 the-boy NEG FUT sleep.3sm-NEG
 ‘The boy will not sleep.’ (Al-Momani 2011:484)

Another variant of verbal negation is the discontinuous morpheme *a*-š,³² e.g., *a-fīyi-š imše* ‘I can’t walk’ (Khairallah 2014:57). This variant is reportedly found in southern and mid-mountainous regions of Lebanon, such as in Keserwan District in the Mount Lebanon Governorate (Khairallah 2014:17). In some parts of the Levant, such as northern Lebanon and on the Syrian coast (Behnstedt 1997:450–51; Behnstedt and Woidich 2005:101), as well as the Jordanian city of Salt northwest of Amman (Palva 2004:227, 231), verbs can be negated with the preverbal *a*- (which is reportedly a reduced form of *ma*; Lucas 2010:186; Wilmsen 2014:62),³³ e.g., *a-katab* ‘he did not write’ (Palva 2004:227), though this is reportedly rare (Palva 2004). Bauer (1926:122) also states that negation with *a*- (with or without -š) is found in Palestine (as cited in Wilmsen 2014:61), but this was not observed in the PA synchronic dataset analysed in

³² Some researchers argue that *a*- is a reduction of the negative element *ma*, specifically in verbs beginning with a labial consonant, e.g., *a-biddi-sh* ‘I don’t want’ (see e.g., Abu-Haidar 1979:110; Blau 1960:193–98; Feghali 1928:221). But Driver (1925:197) claims that *a*- is actually a former interrogative marker, prefixed to a verb with the post-verbal negative enclitic -š, e.g., *a ba’ā-sh fīh* ‘there is not anymore’, where Driver comments that this derives from the literal meaning ‘is there not anymore?’

³³ Many researchers (Abu-Haidar 1979:110; Alqassas 2012:84; Blau 1960:193–98; Feghali 1928:221; Hoyt 2006:116–17, 2010:96; Lucas 2010) suggest that *a*- is a reduction of the negative *mā* in indicative verbs preceded by the aspectual prefix *b-* [...] (Wilmsen 2014:62). However, other researchers (Obler 1975:106; Palva 2004:227) argue that “there must be more to it than simply phonological conditioning, for it can appear with prohibitives, those, of course, beginning with the 2nd person prefix *t-*, a coronal consonant” (Wilmsen 2014:62).

the present study, while in the PA diachronic data source (the POHA), *a-š* was only used once by a male speaker (in contrast, *a-* and especially *a-š* were used slightly more frequently in the LPTC).

Some Arabic varieties also use the post-verbal enclitic *-š* alone to negate verbs, such as the sedentary dialects of historic Palestine (Lucas 2010, 2018), as well as the Arabic varieties spoken in Salt in Jordan (Palva 2004), the highlands of the Horan Plateau of Syria (Cantineau 1946:386–90), southwestern Syria (Behnstedt 1997:450–51), San’a in Yemen (Watson 1993), Upper Egypt (Khalafallah 1969), and in the mountainous regions of Lebanon (Abu-Haidar 1979:110; Fleisch 1974; Khairallah 2014:17). It should be noted that *-š* and *ma-š* are reported not to feature in Beirut LA (Wilmsen 2014:150; see also Behnstedt and Woidich 2005:101; Fleisch 1974:349, 353).

In addition to *ma*, *ma-š* and *-š*, some dialects use *la* for the negation of the imperative, such as Najdi Arabic (Benmamoun 2000), Gulf Arabic (Holes 1990), Kuwaiti Arabic (Brustad 2000:294) and Syrian Arabic, though the latter also uses *ma* for negative commands (Brustad 2000:294). In villages in the Egyptian Delta and in some parts of Jordan, *la* can be used with the post-verbal enclitic *-š* for the negative imperative (Brustad 2000:295). In some Arabic dialects, *la* is also used in disjunction structures equivalent to ‘neither ... nor’ (Soltan 2021:393), such as in Moroccan Arabic (Harrell 1962:156), Egyptian Arabic (Mitchell 1956:46), Syrian Arabic (Cowell 1964:390) and Kuwaiti Arabic (Holes 1990:72).

An additional construction used for verbal negation is the non-discontinuous element *miš* (Brustad 2000:281; see also Al-Sayyed and Wilmsen 2017), which is reportedly derived from the feminine *mā hi* ‘not her/it’ > *mā hi-š* ‘she/it isn’t’ > *miš* (Wilmsen 2014:150, 173, 178,

2021:257).³⁴ *Miš* was originally used for predicate negation; however, its use for verbal negation has been attested as early as the 19th century. In one of the earliest grammars of Egyptian Arabic, Vollers (1890) stated that though it is rare, verbs are negated with *miš* (as cited in Al-Sayyed and Wilmsen 2017:153). By the middle of the 20th century, several grammars of Egyptian Arabic attested this phenomenon. For example, Aboul-Fetouh (1969) stated that *miš* can negate past tense verbs and present tense verbs in Egyptian Arabic. The same phenomenon was attested in Egyptian Arabic a decade and a half later by Eid (1983), as well as in the early 21st century by Aoun et al. (2010:30). Al-Sayyed and Wilmsen (2017:159) also claim, based mainly on their observations of online chatrooms using Egyptian Arabic, that *miš* is used to negate past tense and present tense verbs (see also Soltan 2014:98, 2021:391) in instances of what they call ‘contrastive negation,’ which involves making a contrast between two sets of conditions, e.g.:

- (3) bi-ya-akhud fulus, miš bi-y-gib fulus.
HAB-3sm-take.IPFV money NEG HAB-3sm-get.IPFV money
‘He takes money, not gets money.’
(Egyptian Arabic; Al-Sayyid and Wilmsen 2017:160)

Miš is also said to be favoured with present progressives in Egyptian Arabic (Al-Sayyed and Wilmsen 2017:167; Wilmsen 2014:175) and in rural Palestinian Arabic (Mousa 2019:101; see also Farwanah 2021:171), as in (4).

- (4) miš ʕam yiktib.
NEG PROG write.3sm
'He is not writing.'
(Levantine Arabic; Farwaneh 2021:171)

³⁴ Wilmsen (2014:178) explains that “the negative constructions derived from parent copular negators **mā hūs* ‘he isn’t’ [> *muš*] and **mā hīs* ‘she isn’t’ have become nearly insensitive to gender in *muš* and *miš*, with most dialects applying either one indiscriminately to the masculine, feminine, or plural.”

Additionally, *miš* is claimed to negate future tense verbs in Egyptian Arabic (Doss 2008:87; Soltan 2021:391; Wilmsen 2014:175) and in Palestinian Arabic (Mousa 2019:100; see also Farwanah 2021:171), as in (5).

- (5) *miš* *rah* *ʔaukil* *bi-l-matʕam* *l-yōm*.
 NEG FUT eat.1s in-the-restaurant the-day
 ‘I am not going to eat in the restaurant today.’
 (Urban Palestinian Arabic; Mousa 2019:100)

In addition to contrastive negation, present progressives and future tense verbs, *miš* is said to be favoured with participles (Brustad 2000:289; see also Håland 2011; Wilmsen 2014:175). Brustad (2000:162) explains that participles have a verbal function in spoken Arabic, as they are used to refer to actions and events that have taken place, are taking place, or will take place in the future.³⁵ Some examples of participles are found below:

- (6) *lli* ***baʕeton*** *li* *j-jiran*.
 which having.sent.them to the-neighbours
 ‘The ones she had sent to the neighbours.’
 (Syrian Arabic; Brustad 2000:162)

- (7) *humma* ***mašyīn*** *bukra*.
 they leaving tomorrow
 ‘They are leaving tomorrow.’
 (Egyptian Arabic; Brustad 2000:227)

Eisele (1990) also explains that participles can be stative because they can represent a state, such as verbs of epistemic knowledge (like *to know*), e.g., *miš ʕarfa* ‘I don’t know’ (Egyptian Arabic; Brustad 2000:289). Farwanah (2021:171–72) states that because *miš* is favoured with participles,

³⁵ Procházka (2021:230) explains that “the main function of the active participle is to express the resultative or perfective aspect without any tense reference. Thus in the majority of [Arabic] dialects, the active participle is used to express the resultant state of an activity or a process which, at the moment of speech or another reference point, is of ongoing relevance.” The following examples illustrate participles in Damascus Arabic: *šu ʔābex ʔl-yōm* ‘What have you cooked today?’ (and is now ready for eating) (Procházka 2021:230); *bānt-a nāym-e* ‘Her daughter is sleeping’ (she has fallen asleep and is still sleeping) (Procházka 2021:231).

this can explain why it is also favoured with future tense verbs and with present progressive verbs in Levantine Arabic: “The motivation behind the use of *miš* with future and progressive predicates is that the future and progressive markers *rah/ħa* and *ħam* are derived from the corresponding participial sources *raayih* and *ħammāl* respectively, and participles are normally negated by *miš*” (see also Benmamoun 2000; Wilmsen 2020:19).

Table 2.3 summarizes the major variants reported to be used for verbal negation in several Arabic dialects (see also Appendix A for the distribution of negation strategies in Levantine Arabic).

Table 2.3: Major exponents of verbal negation in some Arabic varieties

Variety	ma	ma-š	-š
Egyptian Arabic in Upper Egypt		√	√
Syrian Arabic	√		
Syrian Arabic in parts of Syria			√
Jordanian Arabic	√		
Jordanian Arabic in parts of Jordan		√	√
Lebanese Arabic (Beirut)	√		
Lebanese Arabic in southern and mid-mountainous regions in Lebanon		√	√
Palestinian Arabic	√	√	√

2.2.3 Verbal negation in Lebanese and Palestinian Arabic

This section discusses two studies on verbal negation, one focusing on Lebanese Arabic, and one on Palestinian Arabic. In Lebanese Arabic, Khairallah (2014) examined verbal negation in speech data collected from speakers in Zeitoun village of Keserwan in Mount Lebanon. She found that there are several variants for negating verbs: *ma*, *ma-š*, *a-š*, *-š* and *miš*. Khairallah focused on verb tense and the initial consonant of negated verbs as linguistic predictors of variant choice, based on the previous literature. Several scholars claim that both *ma-š* and *-š* can negate present tense verbs, but only the former can negate past tense verbs, while *-š* alone is reported not to do so in Levantine Arabic (Abu-Haidar 1979; Farwanah 2021:171; Hoyt 2006,

2010; Lucas 2007, 2010; Obler 1975). For example, Obler (1975:32–33) explains that “the option of dropping negative *mā* when *-š* is suffixed is otherwise exclusive to Palestine and southern Lebanon, but there it does not apply to past tense verbs. Instead, [past tense] verbs must be preceded by *mā*.” Past tense verbs in spoken Levantine Arabic typically do not begin with labial consonants (e.g., *akal-t* ‘I ate’ vs. *b-akul* ‘I eat’), so some scholars add that *-š* cannot occur with non-labial-initial verbs, e.g., **akalt-š* ‘I didn’t eat’ (Blau 1960; Hoyt 2006). Hoyt (2006:16) explains that omission of *ma* is only possible with verbs beginning with labial obstruents such as [b] or [f]. Before Hoyt, Blau (1960:198) hypothesized that *ma* is deleted because of its phonological similarity to the *b*-prefix that is characteristic of the imperfect (present tense verbs) in Arabic. As such, *-š* is reportedly only allowed with labial-initial verbs, but with non-labial initial verbs, *ma* is said to be retained, either alone or with *-š* (i.e., *ma-š*). However, other scholars claim that *-š* can be used with non-labial-initial verbs (Bauer 1926; Driver 1925:197; Obler 1975:107; Palva 2004; Wilmsen 2014). In LA, Thackston (1996) also claims that *-š* negates non-labial-initial verbs in the past tense, e.g., *kuntⁱ-š* ‘I wasn’t.’

In examining her data, Khairallah (2014:15) transcribed negative sentences and counted occurrences of the variants with past tense and present tense verbs, and with labial-initial and non-labial-initial verbs, including pseudo-verbs.³⁶ Based on variant frequencies with the two verb tenses, Khairallah found that *ma* was the most frequently used variant to negate present tense verbs, though other scholars also stated that *ma*, along with *ma-š*, are frequently used to negate past tense verbs (Abu-Haidar 1979; Hoyt 2006, 2010; Lucas 2010; Obler 1975; Palva 2004). Khairallah (2014:40) found that past tense verbs occurred most frequently with *ma-š* (out of 219 tokens with past tense verbs, 134 included *ma-š*), but only 15 out of 219 tokens with past tense

³⁶ Brustad (2000:153) explains that “a pseudo-verb can be a nominal or prepositional phrase that is used semantically to convey a verbal meaning, often but not necessarily possessive or existential in nature”, e.g., *fī* ‘there is’; *bidd-* ‘want’; *šand-* ‘have’; and *maš-* ‘have (on me)’.

verbs were negated with *-š*. Accordingly, consistent with claims in the literature, *-š* does not seem to be favoured with past tense verbs. This variant also seems to be disfavoured with non-labial-initial verbs, as Khairallah (2014:43) found that *-š* occurred only once with the non-labial-initial pseudo-verb *šand-* ‘have,’ and it occurred a few times with non-labial-initial verbs that are in the present tense, e.g., *takil-š* ‘she does not eat’ (Khairallah 2014:45).

Focusing on PA as spoken in Israel, the West Bank and the Golan Heights, Lucas (2010), in the same vein as Khairallah (2014), examined the effect of verb tense and initial consonant of (negated) verbs on variant choice. He collected data using an acceptability judgment questionnaire, asking informants to rate negative sentences as ‘normal’, ‘odd’ or ‘impossible’. Consistent with the majority of the claims in the literature, Lucas (2010:173) found that *-š* is widely accepted with labial-initial verbs or pseudo-verbs. For example, similar to Khairallah (2014:43), Lucas (2010:174) found that the majority of informants judged the non-labial-initial pseudo-verb *šand-* ‘have’ with *-š* as ‘odd’ or ‘impossible’ (with the exception of one informant; see Lucas 2010:174). On the other hand, *ma-š* was reported to be acceptable with non-labial-initial verbs by Lucas’s (2010) informants. *Ma* was found to be acceptable with all pseudo-verbs.

In terms of verb tense, Lucas (2010:172) found that the overwhelming majority of informants preferred *-š* with present tense verbs and *ma-š* with past tense verbs (with the exception of one informant who accepted negating past tense verbs with *-š* alone; see Lucas 2010:173). The following example was judged as acceptable by the majority of informants only with *ma-š*: *ma masaḥna-š t-ṭa:wle* ‘we didn’t wipe the table’ (Lucas 2007:424, 2010:176). This leads Lucas to claim that *-š* alone cannot negate past tense verbs, even when such verbs begin with a labial consonant, such as /m/ in the aforementioned example (see Lucas 2007:424). This is

in line with Awwad's (1987) claim that *ma*, whether alone or with the post-verbal *-š*, is obligatory with past tense verbs in PA.

Both Khairallah's (2014) and Lucas' (2010) studies seem to support claims in the literature that *-š* is not favoured with past tense verbs and with non-labial-initial verbs. Both studies, however, suffer from a number of methodological and analytical issues, which I address in the critical summary at the end of this chapter.

Table 2.4 summarizes the constraints on variant choice in verbal negation that have been reported in the literature on spoken Arabic.

Table 2.4: Reported linguistic constraints on variant choice in verbal negation³⁷

	ma	ma-š	-š	miš
Verb				
Present tense	√	√	√	
Past tense	√	√		
Future tense	√			√
Pseudo-verbs	√	√	√ ³⁸	
Participle				√
Initial consonant of the verb				
Labial-initial verbs	√		√ ³⁹	
Non-labial-initial verbs	√	√		

2.3 Future temporal reference

The variable expression of future temporal reference in spoken Arabic has attracted few systematic treatments (but see e.g., AbuAmsha 2016; Jarad 2013, 2014; Leddy-Cecere 2018, 2020). The restricted literature addressing this variable in spoken Arabic includes contradictory accounts of the functions of future variants, and, in terms of linguistic focus, considers only a few linguistic constraints on variant choice, most notably temporal distance between speech time

³⁷ The shadings in Table 2.4 (under *miš*) are used to show that, to my knowledge, there are no reports regarding the effect of the initial consonant of the negated verb on the use of *miš*.

³⁸ Lucas (2010) claimed that *-š* can only negate labial-initial pseudo-verbs, such as *fī* 'there is', but not *šand* 'have.'

³⁹ Lucas (2010) claimed that *-š* cannot negate labial-initial past tense verbs.

and the future event or state. The nature of the data in previous treatments of future temporal reference in Arabic also includes written material, such as newspapers written in Modern Standard Arabic (Parkinson 2003), where future markers differ from those found in spoken Arabic varieties, as I discuss below.

This section is organized as follows: Section 2.3.1 introduces future markers in Standard Arabic, and section 2.3.2 introduces those in spoken Arabic varieties. Section 2.3.3 discusses reports on the linguistic constraints on variant selection in spoken Arabic. Section 2.3.4 briefly discusses grammaticalization theory, which has been used to explain the development of future markers cross-linguistically (Bybee, Perkins, and Pagliuca 1994). Section 2.3.5 reports on the linguistic constraints on variant choice in natural speech which have been tested in previous variationist studies on European languages, as well as how these constraints can be used as predictors of (the degree of) grammaticalization of future markers (see e.g., Poplack and Dion 2009; Poplack and Malvar 2007; Poplack and Tagliamonte 1999; Poplack and Turpin 1999; Torres Cacoullos and Walker 2009). As detailed in that section, the linguistic constraints on the expression of the future in European languages, and their predictions about the advancement of future markers along the cline of grammaticalization, are used as a heuristic to achieve a better understanding of future temporal reference in LA and PA in the present study.

2.3.1 Overview of the expression of futurity in Standard Arabic

The future markers in Classical Arabic (CA) and Modern Standard Arabic (MSA) are *sawfa* and *sa-* (Parkinson 2003). The present tense can also be used to express futurity, depending on the context (i.e., in the presence of a temporal adverb; see Dat 2005:568; Eisele 2005:196; Jarad 2013:74), e.g., *yuya:diru l-fa:ṣimata yadan* ‘he leaves the capital tomorrow’ (Ryding 2005:608).

Some scholars argue that there is no difference between *sa-* and *sawfa*, as they are

claimed to be used interchangeably in MSA and in CA (Parkinson 2003; Stewart 1998). However, other scholars, like Jarad (2013:74–75), argue that *sa-* indicates a proximate future event, while *sawfa* indicates a distal future event, e.g., *sa-yarkabu aṭ-ṭāira ḥālan* ‘he will board the plane soon’ (Jarad 2013:74); *sawfa yadrusu aḥmad al-ʔadaba il-inglīziyya al-ṣām al-qādim* ‘Ahmad will study English literature next year’ (Jarad 2013:75). In these two examples, it is easy to see the difference in terms of temporal distance, but that is because it is the time adverbials (i.e., *ḥālan* ‘soon’ and *al-ṣām al-qādim* ‘next year’) indicating temporal distance. Poplack and Turpin (1999:139) are critical of the kind of circularity in this interpretive approach, arguing that the meaning of contextual elements such as temporal adverbials is often attributed to the variant form itself (see also Poplack and Dion 2009:582).⁴⁰

There are different claims about the origins of *sa-* and *sawfa*. Stewart (1998:121) argues that it is likely that these two variants are derived from an auxiliary verb expressing motion. In Medieval Arabic, the root *s-w-f* was used in the forms *sāwafa/yusāwifu* (which would account for *sawfa*), meaning ‘to smell’, but this is only in the context of travel, where a guide would

⁴⁰ Poplack and Turpin (1999:135) explain that most scholars claim that the exponents of futurity (in Canadian French) “are selected according to distinctions in the way the speaker envisions the future eventuality and/or the semantic and/or pragmatic import s/he wishes to convey.” However, there is little consensus as to what that import is, or which of the variants is capable of expressing it (Poplack and Turpin 1999:135). Poplack and Turpin (1999:135) suggest that this lack of agreement over the meanings or functions of competing variants “is the product of an epistemological problem stemming from difficulty in reconciling the form-function polyvalence characteristic of inherent variation with the (distributional) linguistic enterprise of ascribing a unique function to every form.” In this connection, Sankoff (1988:153) does acknowledge that while different connotations may be postulated among alternating forms, and although competing variants may have different degrees of acceptability in different contexts, these distinctions need not be pertinent every time one of the variant forms is selected (as cited in Poplack and Turpin 1999:140). Underlying the study of (morpho-)syntactic variation within the variationist framework is the hypothesis that “for certain identifiable sets of alternations, these distinctions come into play neither in the intentions of the speaker nor in the interpretation of the interlocutor” (Sankoff 1988:153). As Sankoff (1988) explains, distinctions in referential value or grammatical function among different variant forms can be neutralized in discourse. This is the fundamental mechanism of language variation and change (Poplack and Turpin 1999:140).

In the literature on spoken Arabic, the subjective, psychological nature of the semantic and/or pragmatic distinctions among exponents of futurity makes them difficult, if not impossible, to operationalise and test in an objective, replicable fashion. Within the variationist approach, we can address the question of the function and meaning of the exponents of futurity through systematic examination of the contexts of future temporal reference in spontaneous speech. In other words, to avoid circularity, our coding procedures are based not on the meaning said to be expressed by a certain variant, but on independent contextual indicators of this meaning, where present (see Chapter 6).

smell dirt or sand for traces of camel excrement to check whether the path has been travelled before (Lane 1984:1469–70), so he could tell whether he was on the right path (Stewart 1998:123). Once the root *s-w-f* was generalized, it came to mean ‘to go, to set out on a journey’, and subsequently, it came to express the future aspect (Stewart 1998:123–24). *Sawfa* was reduced to *sa-* through intermediate stages, as in *saw-* and *saf-* (Stewart 1998:125).

Interestingly, Stewart’s claim about the development of the future markers in CA/MSA from a construction involving motion is in line with Bybee et al.’s (1994:244) argument that cross-linguistically, future markers evolve from a restricted range of lexical sources, including motion verbs. As we will see below, even though future markers in spoken Arabic differ from those in CA/MSA (Zaki, Wilmsen, and Abdulrahim 2021:479), some exponents of futurity in vernacular varieties of Arabic have also developed from motion verbs.

2.3.2 Future temporal reference in some Arabic dialects

The two main lexical sources of future markers across vernacular Arabic varieties are motion and volition verbs (Leddy-Cecere 2018, 2020). Leddy-Cecere (2018, 2020) provides an account of the distribution of future markers across different varieties in the Arabic-speaking world. Using descriptive grammars (e.g., reference grammar books), dialect atlases and annotated dialectological texts (Leddy-Cecere 2018:37, 43), he argues that the broadly contiguous distribution of future markers (members of the multiply-attested grammaticalization pathways FUT < GO and FUT < WANT) across a vast geographical area in the Arabic-speaking world is “consistent with a history of diffusion via contact between speakers of neighboring dialects” (Leddy-Cecere 2020:616).⁴¹

⁴¹ Leddy-Cecere (2018:86) states that the “collected products of FUT < GO and FUT < WANT grammaticalization pathways may thus be identified as representing the results of repeated processes of [contact-induced grammaticalization]. Under this scenario, one “founding” member of each of these two pathways may be

In terms of *go*-futures, future forms based on *rāyih* ‘going’ (the active participle of the verb *rāḥ*) (see Leddy-Cecere 2018:50), such as *raḥ* and the proclitic future marker *ḥa-*, are widespread across the Arabic-speaking world (Leddy-Cecere 2018:82). *Rāyih*-derived future markers are reportedly observed in Lower Egypt and the Levant (Farwaneh 2021:160–61; Jarad 2014), extending eastward to southern and central Iraq, and southward along the Nile Valley as far as central Sudan and the western littoral of the Arabian Peninsula (Leddy-Cecere 2018:82). Moving to the west, we also find future markers derived from *rāyih* “in a band reaching across modern Libya and Tunisia to Algerian territory” (Leddy-Cecere 2018:82).

Besides *rāyih*, there are different realizations of the FUT < GO path in other areas. For example, forms derived from *māšī* ‘going’ (originally meaning ‘walking’) (see Leddy-Cecere 2018:53) are allegedly attested in Tunisia, and we find future markers from *yādī* ‘going’ in the greater Moroccan area (Leddy-Cecere 2018:83). Additionally, there are future forms derived from *māšī* used in northern Morocco (Leddy-Cecere 2018:83). As Leddy-Cecere (2018:83) argues, these distinct realizations representing the FUT < GO grammaticalization path (i.e., *rāyih*, *māšī*, *yādī*) “display complete geographic continuity.”

With respect to volition-based futures, or the FUT < WANT path, we find future forms deriving from *yabyā* ~ *yabyī* ‘want’. These are attested in the Arabian Peninsula and span across the Red Sea throughout the greater Sudanic area (Leddy-Cecere 2018:83). Moving northward from this region, we also find such forms in modern Libya. Leddy-Cecere (2018:56) explains that forms derived from *yabyā* ~ *yabyī* display phonological change of the sequence /by/ > /bb/

hypothesized to have originated independently, a development not surprising in light of the preponderance of these grammaticalization paths cross-linguistically. Subsequently, once the innovation had become established within a given area and the grammaticalization process had begun, speakers of a neighboring dialect or dialects encountered the change, comprehended its etymological source and trajectory, and replicated it in their own Arabic variety utilizing distinct but synonymous etymological resources.”

(> /b/) in some Arabic dialects (see also Al-Najjar 1991; Holes 1990; Johnstone 1967; Persson 2008), as in *b-* used in Abu Dhabi (Leddy-Cecere 2018:56).⁴²

Bordering the Arabian branch of *yabyā ~ yabyī* to the north and west, we find other realizations of the FUT < WANT pathway, those of *biddu ~ widdu* (Leddy-Cecere 2018:83), lit. ‘it is (in/by) his will’ (Leddy-Cecere 2018:57). Futures derived from *biddu*⁴³ ~ *widdu* are mainly found in the Levant, such as in Syrian Arabic (Jarad 2013) and Palestinian Arabic (AbuAmsha 2016). Leddy-Cecere (2018) states that phonetic erosion has occurred to the future marker *bidd-* in some Levantine varieties (see section 2.3.4.1), commonly beginning “with the collapse of the vowel and coda of the initial syllable” (Leddy-Cecere 2018:58; see also Jarad 2013), e.g., Aleppo *baddo ~ bdo ~ b-*⁴⁴ (Leddy-Cecere 2018:58; see Owens 2018).⁴⁵ In Levantine Arabic, the *b-* prefix is cliticized to the imperfect (present tense) verb, and it is multifunctional because it is used to express future as well as non-future temporal meanings, such as the progressive and habitual aspect (Farwanah 2021:160; Jarad 2013:71; Owens 2018:211, 219).

In addition to *yabyā ~ yabyī* and *biddu ~ widdu* (as well as the *b-* prefix), other realizations of the FUT < WANT path are reflexes of *yīšā* ‘want’ (Leddy-Cecere 2018:59), which occur in close proximity to *yabyā ~ yabyī* forms in the southwest corner of the Arabian Peninsula (Leddy-Cecere 2018:83), e.g., in some Yemeni dialects such as the Arabic spoken in Sana’a

⁴² Leddy-Cecere (2018:56) explains that “[a]s a future tense marker, the initial syllable is most often lost; many Arabic varieties attest a nearly maximally reduced realization, such as Abu Dhabi *b-*, while others retain evidence of original final vowel quality, as in Khartoum *bi-* [...]”

⁴³ As Leddy-Cecere (2018:58) explains, there are local variations in the quality of the internal vowel common to both lexical and grammatical uses of *bidd-* in the Arabic varieties in which it is found (e.g., *bidd-*, *badd-*, *bədd-*).

⁴⁴ It should be noted that while *b-* in some Levantine Arabic is superficially indistinguishable from the *b-* that is reportedly product of *yabyā ~ yabyī* (Leddy-Cecere 2018:58), “the presence of several fully attested chains of development for both families of items [...] prompts the identification of separate origins” (Leddy-Cecere 2018:58).

⁴⁵ It should be noted that according to Owens (2018), the origins and evolutionary trajectory of the *b-* prefix in several varieties of Arabic in which it is found are less clear-cut and more controversial than what has been claimed in several sources (Al-Najjar 1991; Holes 1990; Jarad 2013; Johnstone 1967; Persson 2008). Owens (2018) argues that the *b-* prefix, which is common across several Arabic varieties, has only one origin and it developed along the following pathway: *yabya* > *yiba* > *b-* (‘he wants’ > reduced to *yiba* > prefix). Owens (2018) examines how it developed from one origin and acquired new meanings/functions in different Arabic varieties.

(Leddy-Cecere 2018:59). In interior North Africa, other volition-based future tense markers encountered include reflexes of *ydawr* ‘want’ in Mali and Mauritania, as well as forms from *byā* (the perfective form of *yabyā* ~ *yabyī*) (see Leddy-Cecere 2018:61) in southern Morocco and southwestern Algeria (Leddy-Cecere 2018:83). Leddy-Cecere (2018:84) concludes that similar to the members of the FUT < GO grammaticalization pathway, the realizations of the FUT < WANT path display a fully contiguous spatial distribution.

While Leddy-Cecere (2018, 2020) is an important account of the distribution of future markers across the Arabic-speaking world, his argument about the role of contact-induced grammaticalization in the evolution of the future system in spoken Arabic warrants cautious assessment. I return to this issue below.

Table 2.5 below summarizes some of the major exponents of futurity, especially those attested in the Levant and neighbouring areas (i.e., *rah*, *ha-*, *bidd-* and *b-*) due to their relevance to the present study focusing on the Levantine varieties of LA and PA.

Table 2.5: The distribution of *rah*, *ha-*, *bidd-* and *b-* across some vernacular varieties of Arabic

Variety	<i>rah</i>	<i>ha-</i>	<i>bidd-</i>	<i>b-</i>
Egyptian (Cairene) Arabic (Abdel-Hafiz 2005; Brustad 2000; Leddy-Cecere 2018:48)		√		√
Sinai Arabic (Leddy-Cecere 2018:48, 57)	√	√	√	
Aswan Arabic (Leddy-Cecere 2018:48)		√		
Syrian Arabic (Jarad 2013, 2014; Leddy-Cecere 2018:48, 57)	√	√	√	√
Jordanian Arabic (Amman) (Alshboul, Al Shaboul, and Asassfeh 2010; Leddy-Cecere 2018:48, 57)	√	√	√	√
Palestinian Arabic/Jerusalem Arabic (AbuAmsha 2016; Leddy-Cecere 2018:48, 57)	√	√	√	√
Lebanese Arabic ⁴⁶ (Mitchell and El-Hassan 1994; Thackston 1996)	√			√
Iraqi Arabic (Abu-Haidar 1991)	√	√		
Baghdad Arabic (Leddy-Cecere 2018:48)	√			
Gulf Arabic (Holes 1990; Johnstone 1967; Persson 2008)	√			√
Kuwaiti Arabic (Brustad 2000)	√			√

⁴⁶ LA is claimed to use only *rah* and the *b-* prefix, but *ha-* and *bidd-* have also been observed in the LA data in the present study.

Meccan Arabic (Leddy-Cecere 2018:48)	√	√		
Libyan Arabic (Tripoli, Benghazi) (Leddy-Cecere 2018:48)		√		
Tunisian Arabic (Tozeur) (Leddy-Cecere 2018:48)		√		
Sudanese Arabic (Khartoum) (Leddy-Cecere 2018:48)		√		

2.3.3 Linguistic constraints on the use of future variants in spoken Arabic

This section discusses the reported linguistic constraints on the use of the four major exponents of futurity in Levantine Arabic (*rah*, *ha-*, *bidd-* and *b-*), which especially concern temporal distance and the expression of intention.

Rah ‘going to’

Several previous works discuss the functions of *rah*. In a reference grammar of Syrian Arabic, Cowell ([1964] 2005:322–23) argues that *rah* often carries a sense of immediacy, so it can be translated as ‘about to’, and he calls *rah* the particle of anticipation, indicating that an action is impending in the future, based on speaker intention or a course of events already under way. In Jerusalem Arabic, *rah*, as well as *ha-* are also reportedly used for the immediate future or future intent (Leddy-Cecere 2020:612). Jarad (2013:71) also argues that *rah* in Syrian Arabic reflects the speaker’s certainty that an event is impending in the future, and it asserts the proximity of the future event (see also Procházka 2021:230). However, in a separate publication, Jarad (2014:107) claims that *rah* in Syrian Arabic is used to describe an event which will take place in the immediate or distant future.

Other internal constraints discussed are animacy and verb type. Based on natural speech data, AbuAmsha (2016) gives examples of *rah* in PA spoken in Gaza city, showing that this variant includes the semantic elements of locomotion and agency, as in (8), but it can also be used with non-human subjects, as in (9).

- (8) *rah tirjaʕi ʕa yazeh?*
 FUT return.2sf to Gaza
 ‘Will you return back to Gaza?’

- (9) *youm l-jumʕah ʕina rah ykun ʕawaʕif.*
 day the-Friday at-1p FUT be.3sg storms
 ‘Next Friday, there will be storms.’

(AbuAmsa 2016:7)

Only Persson (2008:33) mentions the effect of clause type on the selection of *rah* in Gulf Arabic, finding that it is favoured in main clauses in speech data collected from speakers from Bahrain, Kuwait, Oman, Qatar and the UAE.

ha- ‘going to’

The proclitic future marker *ha-* is claimed to refer to the near future in several descriptive works on Arabic varieties, such as in Jerusalem Arabic (Rosenhouse 2011; as cited in Leddy-Cecere 2020:612). In Cairene Egyptian Arabic, Woidich (2005b:331) reports that it refers to ‘being on the point of’, and Holes (2004:228) states that *ha-* is particularly used for proximate intention in this variety, as it involves a strong commitment on the part of the speaker. Nevertheless, it is not clear how this is assessed, especially because psychological notions like speaker intention and commitment cannot be objectively assessed unless there is some *overt* clue as to the speaker’s attitude towards the future predication (Poplack and Malvar 2007:137; Poplack and Tagliamonte 2001:217; Poplack and Turpin 1999:145). Holes (2004:228) puts forward some examples to show that *ha-* is used to express the speaker’s intention about a future event. He argues that *ha-yiʔra l-kita:b bukra* ‘He’s going to read the book tomorrow’ indicates more certainty that this is the subject’s intention than *b-yiʔra l-kita:b bukra* ‘He’ll read the book tomorrow’, which he claims is an uncommitted prediction (Holes 2004:228). However, from his examples, it is not clear how Holes can make such a distinction between *ha-* and *b-* in terms of the expression of the

speaker's intention. Similar to Holes (2004), El-Hassan (2007:265) claims that *ha-* expresses the intention of the speaker, and it is used with the first person. In Levantine Arabic, *ha-* is claimed to be used for proximate futures, and it allegedly asserts the intention of the speaker regarding a future event (Holes 2004:228; Mitchell and El-Hassan 1994:25), but once again, these claims are based mainly on random examples that the scholars themselves propose, rather than on exhaustive and accountably sampled usage data.

Bidd- 'want' and the b-prefix

Several linguistic constraints are also reported to condition the selection of *bidd-* 'want' and the *b-*prefix. It would be expected that scholars would comment on the use of temporal adverbs that disambiguate the meaning of these variants, as their future reading depends on the ambient context. But previous accounts focus mainly on agentivity, volition and intention. Even though *bidd-* is originally a volition verb, it is said to be used without an element of volition with non-human subjects in Palestinian (AbuAmsa 2016) and Jordanian Arabic (Al-Wer 2006:514). However, Mitchell and El-Hassan (1994:39) claim that in the Levant, *bidd-* is often used with reference to speaker intent, and it mainly occurs with human subjects. Their claims are based on cherry-picked examples from a corpus of Educated Spoken Arabic⁴⁷ and on native intuitions of speakers from Egypt and the Levant (Mitchell and El-Hassan 1994:xiii). Unlike Mitchell and El-Hassan (1994), Leddy-Cecere (2020:613) claims that in Damascus Arabic, *bidd-* (and *b-*) refer to a possible future, as opposed to *rah* and *ha-*, which reportedly indicate a higher degree of certainty that the future event will take place, though Leddy-Cecere (2020) does not explain how he infers these subtle semantic meanings.

⁴⁷ Educated Spoken Arabic "is a form of conversational Arabic used by educated speakers from one or more Arab countries" (Mitchell and El-Hassan 1994:1).

The *b*-prefix is said to refer to the impending or proximate future in several Arabic varieties, such as in Yemeni Arabic (Simeone-Senelle 2006:272). But Mitchell and El-Hassan (1994:25) report that in the Levant, *b*- refers to the remote future. Their claim that *b*- refers to the distal future is based on one example: *ḥatta šufha bisʔallak yaha* ‘I’ll ask her for you when I see her’ (Mitchell and El-Hassan 1994:25). But this example is arguably interpreted as referring to an *indeterminate* future, as there isn’t any clue as to when the future event will take place.

Mitchell and El-Hassan (1994:19) also claim that *b*- involves a weaker commitment to the future action than *bidd*- and *ḥa*-; the *b*-prefix reportedly indicates suggestions as to the possibility of the future event, whereas the other co-variants indicate a definite commitment to the future event (see also Holes 2004:228). However, in Syrian Arabic, Jarad (2013:71) claims that the *b*-prefix denotes volition, willingness and intention. Again, these scholars (Holes 2004; Jarad 2013; Mitchell and El-Hassan 1994) don’t explain how they infer such nuanced semantic conditioning of variants. Mitchell and El-Hassan (1994:19) base their claim on one example: *bšufak j-jumʕa j-jay* ‘I’ll see you next week’. This example is problematic because the authors offer no objective criteria for operationalizing and testing for ‘weaker commitment’. Absent any such systematic criteria, the reader is left to accept the authors’ interpretation on faith alone.

Jarad (2013:76–77) comments on the use of the *b*-prefix for contingent events in Syrian Arabic, claiming that it is favoured in the context of conditionals, where it occurs in the main clause (see also Bybee and Dahl 1989:93–94; Bybee et al. 1994:274; Persson 2008), e.g.:

- (10) ʔiza btedros kwayes, **b**-iṣṭrī-l-ak laptop.
 if study.2sm well FUT-buy.1s-for-you laptop
 ‘If you study well, I will buy you a laptop.’

The *b*-prefix is said to express intention or volition, implying that it may be favoured with human and first-person subjects (Brustad 2000:244; Ingham 1994; Johnstone 1967). However,

Brustad (2000:243) observes that in the Kuwaiti Arabic speech data that she collected, *b-* is used with non-human subjects, so it can be used without any indication of will or intent (see also Holes 2004:247; Jarad 2013:79). In a descriptive account of future temporal reference in Libyan (Tripoli) Arabic, Pereira (2009:555) claims that *b-* expresses the future of intention, but only in affirmative or interrogative statements.

Table 2.6 summarizes the main constraints on variant choice in future temporal reference that have been reported in the literature on spoken Arabic. The shaded areas in the table mean that to the best of my knowledge, there have been no reports on how a particular variant is constrained by the linguistic factor in question (e.g., there are no reports of how *rah* is constrained by sentence type, hence the shading).

This section showed that there are several, often inconsistent, claims about the linguistic constraints on variant selection in future temporal reference in spoken Arabic. The contradictions in terms of the meanings attributed to the future variants in Arabic are due to the use of decontextualized utterances of unknown social provenance which may not reflect everyday speech. I return to these issues in the chapter summary.

Table 2.6: Reported linguistic constraints on variant choice in future temporal reference in spoken Arabic

	rah	ha-	bidd-	b-
Sentence Type				
Negative				
Affirmative				√
Interrogative				√
Clause Type				
Main	√			√
Subordinate				
Temporal reference				
Proximate	√	√		√
Distal	√			√
Animacy				
Human	√	√	√	√

Non-human	√		√	√
Grammatical Person				
First-person		√		√
Non-first-person				
Verb Type				
Stative	√			
Motion	√			

2.3.4 Grammaticalization of future tense markers

This section briefly discusses grammaticalization theory and its relevance to the investigation of future temporal reference in LA and PA (see Jarad 2013, 2014). As mentioned previously, cross-linguistically, future markers have developed from different lexical sources, such as motion and volition verbs, or from similar sources at different periods (Bybee et al. 1994:243), which produces a layering of new markers over older ones (see Bybee et al. 1994:243; Hopper 1991:23).

Grammaticalization is widely understood (see e.g., Bybee and Dahl 1989; Bybee, Pagliuca, and Perkins 1991; Bybee et al. 1994; Heine, Claudi, and Hünemeyer 1991) as the process “whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions” (Hopper and Traugott 2003:xv). The transition of an item along a cline from more lexical to more grammatical occurs in the following manner, according to Hopper and Traugott (2003:7): content word > grammatical word > clitic > inflectional affix.

Heine (2003:579) identifies four mechanisms that are specific to grammaticalization (see also Traugott 2011:29): (i) desemanticization or bleaching; (ii) extension or context generalization; (iii) decategorialization (i.e., change in category membership; see Hopper 1991); and (iv) erosion or phonetic reduction (see also Lamiroy and De Mulder 2011:303). Semantic bleaching occurs when the semantic content of the lexical item is gradually lost (Givón 1973; see Bybee

and Dahl 1989:61). In other words, this item is bleached of its lexical meaning and what is left is a grammatically enriched meaning (Traugott 2010:275; Traugott and König 1991). Extension, or the generalization of the contexts where a grammatical morpheme can be used (Bybee and Dahl 1989:63; Bybee et al. 1994:6) occurs because the meaning of the grammatical item becomes more abstract, which makes it appropriate in a wide range of contexts (Bybee 2002:1, 2011:72; Hopper and Traugott 2003:1). A grammaticalized item is also reported to undergo decategorialization, which refers to the loss of the morphosyntactic properties associated with its source's original word class (Lamiroy and De Mulder 2011:303; Leddy-Cecere 2020:605), e.g., loss of agreement inflection or word order freedom. The item undergoing grammaticalization can also undergo phonetic reduction with high frequency of use (Bybee 2001, 2003, 2011:70–71; Bybee and Pagliuca 1985:76; Bybee et al. 1994:6; Traugott 2011:28), e.g., *be going to* > *be gonna* (Bybee 2011:70; Traugott 2011:28). However, this is reportedly not required for grammaticalization to occur (Heine and Kuteva 2007:42; Lamiroy and De Mulder 2011:303).

In explaining the variable expression of futurity, Bybee et al. (1994) rely on reference grammars, teaching grammars and journal articles (Bybee et al. 1994:32), which are at some remove from natural speech where patterns of variability and change are located. However, they remain a valuable resource in explaining the common sources from which future markers develop cross-linguistically. The most prominent lexical source for the development of future markers in many languages is motion verbs (Bybee et al. 1994:267). Such verbs imply movement in time as well as in space, so the “temporal meaning that comes to dominate the semantics of [motion verbs] is already present as an inference from the spatial meaning” (Bybee et al. 1994:269). The role of intention is prominent in the evolution of motion verbs into future markers, as when the agent states that she is going somewhere to do something, she is expressing

the intention to go to a certain place to do a certain thing (Bybee et al. 1994:269). The change that takes place in the meaning of motion constructions concerns their generalization “to contexts in which an intention is expressed, but the subject is not moving spatially to fulfill that intention” (Bybee et al. 1994:269; Eckardt 2011:393). As such, the restriction with the original meaning of *going to* regarding the movement of an animate (or a first-person) subject is not necessary with the future marker *going to* (Bybee and Dahl 1989:91). Thus, *go*-futures express movement towards a goal, then intention, then prediction (Poplack and Tagliamonte 1999:316).

Another common source for the development of future markers is volition verbs (Heine 1993:31; see also Esseesy 2021:287–88). The desire future pathway starts with desire and then moves to willingness. Throughout their development, desire futures are said to go through a stage in which they are used to convey intention, first of the speaker, and then of other subjects, where prediction arises (Bybee et al. 1994:256). The following pathway is proposed by Bybee et al. (1994:256) for desire futures: desire > willingness > intention > prediction.

Present tense forms can also be used to express future temporal reference (these are referred to as aspectual futures; see Bybee et al. 1994:275), but this is contextually determined (Bybee et al. 1994:275). This means that the use of the present tense to express future temporal reference is mainly restricted to temporal clauses where there is typically an accompanying (future) time adverb (Tagliamonte, Durham, and Smith 2014:82; see also Poplack and Turpin 1999:151; Torres Cacoullos and Walker 2009:335), such as *tonight*, *tomorrow*, etc. Poplack and Turpin (1999:152) explain that the effect of time adverbials on the (futate) present “appears to be a functional or disambiguating one, since the absence of an adverbial could confer an habitual reading to the unmarked statement.”

2.3.4.1 Grammaticalization of future tense markers in Levantine Arabic

Accounts of the grammaticalization of future tense markers in spoken (Levantine) Arabic are scarce (Jarad 2013, 2014; Leddy-Cecere 2020; see Esseesy 2021). Jarad (2014) claims that this is because there is a lack of historical records of some spoken Arabic varieties, such as Syrian Arabic (Jarad 2013:71), which are necessary for the reconstruction of the processes of grammaticalization of future markers. Accordingly, Jarad (2013, 2014) relies on accounts of grammaticalization processes affecting future markers in English in his attempt to provide a reconstruction of the evolution of future markers in Syrian Arabic.

Raḥ and ḥa- ‘going to’

Jarad (2014:113) explains that the Classical Arabic motion verb *raḥ* ‘go’ “is the source concept from which the prospective future particle [*rāyih/raḥ*] developed” (see also Esseesy 2021:295; c.f. Farwaneh 2021:162–63).⁴⁸ He explains that this motion verb developed from andative use, to purposive use, to future intention (Jarad 2014:102, 106), e.g.:

Lexical verb (andative use)

- (11) ana **ra:yih** ʕa j-ja:mʕa.
 I going to the-university
 ‘I am going to the university.’

Purposive use

- (12) ana **ra:yih** ʃu:f rifʔa:t-i.
 I FUT see friends-my
 ‘I am going to see my friends.’

Future particle

- (13) ana **rah** ru:h.
 I FUT go.1s
 ‘I am going to go.’

Prefix

- (14) ʔemta **ha**-tsa:fer Maryam?
 when FUT.travel Maryam
 ‘When is Maryam going to travel?’

The development of *raḥ* ‘go’ from andative to purposive to future intention is a pathway attested cross-linguistically (Bybee et al. 1994; Croft 2000; Heine and Kuteva 2002a, 2002b). In

⁴⁸ In disagreement with Jarad (2014), Farwaneh (2021:162–63) argues that the *b*-prefix “evolved from the regional form of the [Classical Arabic] continuous modal *baqā* ‘stay/remain’” (see also Eksell 2006).

Syrian Arabic, Jarad (2014:102) argues that the change of the lexical verb *ra:h* to a future marker occurs in the context of purposive directional complements (12) but not in contexts of a directional in the presence of locative adverbs (11). This is because there is an inference of futurity from purposives (Hopper and Traugott 2003:3). Similar to the case of *going to*, the transition from the motion verb *ra:h* to the future particle *rah* is a change from the cognitive domain of space to the conceptual domain of time (i.e., future temporal reference; see Jarad 2014:102). The first stage in the transition of *rah* from a lexical to a grammatical item is mainly seen in contexts where a movement interpretation of *rah* is not possible, as in (13), where *rah* is used as a future marker, not as a motion verb. The lexical verb *ru:h* ‘(I) go’ is the motion verb in (13). In this case, there is sufficient bleaching of the lexical content of *rah* “to permit co-occurrence with its erstwhile lexical source without ensuing ungrammaticality” (Esseesy 2021:296). The second stage in the grammaticalization of *rah* involves the expansion of its meaning from movement to intention and then to prediction, and then it starts to appear in passive structures and in constructions with inanimate subjects (Jarad 2014:109), i.e., there is semantic change where “the restriction on the subject/agent to a willful sentient being is lifted” (Esseesy 2021:296). Jarad (2014) explains that the generalization of *rah* for use with non-human subjects is an important step in the transition from a motion meaning to a future one, where prediction, and not necessarily intention, is implied. Jarad (2014:106) states that *rah* also shows decategorialization to an auxiliary (see Esseesy 2021:296), as it no longer takes agreement inflection, i.e., *ra:h* (SG.M)/*ra:hat* (SG.F)/*ra:hīn* (PL) versus *rah* (invariant) (see Leddy-Cecere 2020:606). Phonetic reduction of *rah* is also reported to have occurred, resulting in the prefix *ha-* (Jarad 2014:110, 113). The development of *rah* and *ha-* can be schematized as follows: *rayih/ra:h* (lexical/motion verb) > *rah* (grammatical) > *ha-* (Jarad 2014:110–11; see also Leddy-

Cecere 2020:606).

Bidd- ‘want’ and b-prefix

The other two future markers in Levantine Arabic are *bidd-* ‘want’ and the *b-*prefix. Jarad (2013:78–81) argues that the latter developed from the verb *bi-wudd-i/ bi-widd-i* ‘it is (in/by) his will’ in Standard Arabic in the following way: *bi-wudd-i/ bi-widd-i* > *badd-/bidd-* (lexical) > *badd-/bidd-* (grammatical) > *b(i)-* (see also Esseesy 2021:296; Jarad 2017:752).

The first stage in the development of the *b-*prefix as a future marker is when the volitional meaning of the verb *bi-wudd-i* is bleached out, leaving the temporal relation of futurity. This signifies a change from agent-oriented to speaker-oriented modality and the metaphorical extension of meaning from the domain of ‘want’ to the abstract domain of time (i.e., future temporal reference; see Jarad 2013:79).

The second stage is the combination of the preposition *bi* with *wudd*, resulting in the lexicalized verb *bidd-* ‘want’ in the following manner: because the preposition *bi* ends with a vowel, the glide ‘w’ of *wudd* is dropped, and vowel truncation of ‘u’ occurs in line “with the general tendency in [...] Arabic phonotactics to rule out the occurrence of more than two vowels in a row” (Jarad 2013:80). In Syrian Arabic, *bidd-* as a volition verb co-exists with *bidd-* as a future tense marker. This is referred to as layering, where new functions and meanings have emerged, but the old ones co-exist (Hopper 1991:22). The third stage in the development of the *b-*prefix, according to Jarad (2013:80), is the phonetic reduction of *bidd-*: syllabic erosion reduces the form by shortening *bidd-* to *bi-*. The form *bi-* undergoes further erosion, where the

vowel gets dropped, and the resultant form is the *b*-prefix (Jarad 2013:80; but see Owens 2018).⁴⁹

2.3.5 Linguistic constraints and measurements of grammaticalization

In this section, I summarize the linguistic factors which have been tested in previous variationist research on future temporal reference in European languages (see e.g., Poplack and Dion 2009; Poplack and Malvar 2007; Poplack and Tagliamonte 1999; Poplack and Turpin 1999; Torres Cacoullos and Walker 2009). In addition, I discuss how these factors can be used to gauge ongoing grammaticalization. It is important to reiterate that I draw on the variable expression of the future in European languages as a heuristic in order to better understand the future in vernacular varieties of Arabic in the absence of any detailed quantitative treatments of this variable in Levantine Arabic. I am not operating with any a priori assumption that the expression of futurity in Arabic is necessarily identical to what has been found in European languages. Nevertheless, the fact that spoken Arabic derives its future markers from sources that have also provided input to the expression of futurity in European languages would lead us to expect that future markers in Arabic may behave in similar ways to their European analogues in terms of the advancement of these markers along the grammaticalization pathway.

The factors operationalized in previous studies relate to a number of lexico-semantic and morpho-syntactic properties of the contexts in which the exponents of futurity appear (see Poplack and Malvar 2007:137; Poplack and Turpin 1999:146). At the syntactic level, the *type of sentence* (affirmative, negative or interrogative) and *type of clause* (main or subordinate) have been taken into account. In terms of sentence type, affirmative or declarative sentences are considered to be less specified, more frequent and ‘unmarked’ contexts (Poplack and Malvar

⁴⁹ As stated in footnote 27, Owens (2018) points out that the historical reconstruction of the *b*-prefix is complicated by its functional polyvalence in modern varieties of Arabic, so Jarad’s (2013) proposal that the *b*-prefix is a morpho-phonologically reduced form of *bidd-* should be treated cautiously.

2007:147–49; Schwenter and Torres Cacoullos 2008:29), so if an exponent of futurity appears in such contexts, this may highlight its role as a default future marker and may indicate that it is proceeding along the grammaticalization pathway.

In terms of clause type, Bybee et al. (1994:231) argue that when new tense and aspect grammatical morphemes arise, they do so in main clauses, which are said to be innovative, while subordinate clauses are considered as conservative grammatically, retaining older syntax and morphology (Bybee 2002:1; Bybee et al. 1994:231). As such, “a consequence of the gradual spread of new [grammatical morphemes] to more environments is that some subordinate clause environments will not be affected immediately, and here the older [grammatical morphemes] will continue to be used—trapped, so to speak, in these conservative environments” (Bybee et al. 1994:231). That is why, Bybee (2002:5) argues that “the greater richness and explicit semantics of the newly grammaticized form is more appropriate in the assertive context of main clauses.” In other words, if futures constitute a prediction by the speaker (Bybee and Pagliuca 1987; Bybee et al. 1994:235, 244), they are more likely to occur in assertive main clauses than in subordinate clauses, like *when* clauses, which are contingent rather than asserted (Bybee et al. 1994:274; see Torres Cacoullos and Walker 2009:327). But Bybee et al. (1994) suggest that if future markers occur in subordinate clauses, this indicates their advancing grammaticalization, because in the earlier stages, future markers are generally associated with assertiveness, which, as mentioned above, tends to be associated with main clauses. Accordingly, if future markers occur in subordinate clauses, this means that they have lost their assertive meaning and they are in an advanced developmental stage (see Bybee et al. 1994:244, 271, 274).

At the semantic/discourse level, studies examined *temporal distance* between speech time and the future event (coding for distal, proximate or indeterminate futures), as well as *presence*

of temporal adverbs and *adverbial specification*. In coding for temporal distance, studies determined proximity only when it was reliably inferred from the discourse context (Torres Cacoullos and Walker 2009:329). Poplack and Malvar (2007:141) state that a persistent reading of *go*-futures and the futurate present is that of proximity in the future (Poplack and Tagliamonte 2001; Poplack and Turpin 1999). However, it is hypothesized that a future variant that is favoured with indeterminate futures may be proceeding along the cline of grammaticalization because such futures constitute a less specified context (i.e., they lack a specific temporal reference; see Schwenter and Torres Cacoullos 2008:30), as opposed to proximate or distal futures (see e.g., Bybee 2003:605; Bybee et al. 1991:31; Leddy-Cecere 2018:19–20).⁵⁰

In terms of temporal adverbs, these are crucial in disambiguating future and non-future readings, especially in the use of the present tense to refer to the future (Tagliamonte et al. 2014:82). Several studies found that the futurate present is favoured with (specific) temporal adverbs (Poplack and Malvar 2007:153, 157; Poplack and Turpin 1999:149, 152; Torres Cacoullos and Walker 2009:334–35). But a sign of advanced grammaticalization of a future tense marker is its use without temporal adverbs because this means that it unambiguously refers to the future, so it does not need supporting material like temporal adverbs for disambiguating purposes. For example, Torres Cacoullos and Walker (2009:341) found that *be going to* in Canadian English is favoured with no adverbs, and they explained that the “greater independence of *going to* from adverbial modification lends support to the view that temporal meaning was present as an inference from the spatial meaning of *going to* from the beginning” (Torres Cacoullos and Walker 2009:341).

⁵⁰ Bybee et al. (1991:31) explain that as the future marker further proceeds along the cline of grammaticalization, it may be associated with less specialized contexts like those signifying simple futurity, i.e., without a specific temporal distance.

At the level of the subject, *grammatical person* (and number) and *animacy* have been tested in previous variationist studies. Because future variants (such as *go*-futures and future markers derived from volition verbs) can be associated with readings of certainty, intention or volition, which relate to agentivity (Poplack and Malvar 2007:137; Poplack and Tagliamonte 1999:334), these future variants can be expected to favour (first-person) human subjects who exercise the highest degree of agentivity (Poplack and Malvar 2007:137). Poplack and Tagliamonte (1999:335) explain that the co-occurrence of a future marker with a particular type of subject can measure grammaticalization. They explain that Bybee et al. (1991) observed that “agent-oriented uses, which predicate certain conditions on the agent with respect to the completion of the action, are close to the lexical meaning of the original source material and thus tend to occur early in the evolution of the form” (as cited in Poplack and Tagliamonte 1999:335). For example, *going to* in English, which is originally a motion verb, started occurring with human subjects; as it further grammaticalized, its meaning was generalized from movement to intention to prediction (Bybee et al. 1994:5), and it was used with non-human and non-first person subjects as well. This indicates its semantic bleaching and advanced grammaticalization (see Poplack and Tagliamonte 1999:335).

At the level of the verb, studies coded for *verb type* (stative or motion verb). Poplack and Malvar (2007:138) explain that the best known lexical effect in the future temporal reference sector concerns the avoidance of *go*-futures with motion verbs until the former are bleached of their original semantic content of movement towards a goal and grammaticalized as future markers. As such, if a *go*-future retains its original meaning of movement, it should be less likely to occur with a motion verb (Poplack and Malvar 2007:138). On the other hand, the use of a *go*-

future with a motion verb implies its semantic bleaching and is consistent with its advanced grammaticalization (Poplack and Tagliamonte 1999:335).

Table 2.7 below summarizes how ongoing grammaticalization can be inferred from shifting usage constraints. The table mainly focuses on *go*-futures because of the substantial attention they received in the literature (see e.g., Tagliamonte et al. 2014).

Table 2.7: Summary of predictions for the grammaticalization of *go*-futures
(after Tagliamonte et al. 2014:89)

Measure	Initial stage of grammaticalization	Later stage of grammaticalization
Sentence type	Negatives/interrogatives	Affirmative sentences
Clause type	Main clauses	Subordinate clauses
Temporal distance	Proximate/distal futures (specialized contexts)	Indeterminate futures (less specialized contexts)
Temporal adverbs	Adverb present	No adverb
Animacy	Human subjects (movement/intention)	Non-human subjects (prediction)
Grammatical person	First-person subjects (movement/intention)	Non-first person subjects (prediction)
Verb Type	Stative verbs	Motion verbs

The aforementioned predictions based on co-occurrence possibilities discussed in the literature allow us to assess ongoing grammaticalization in the LA and PA future temporal reference systems. Crucially, to address this study's overarching objective of examining whether PA is converging on LA, the measures of grammaticalization allow us to investigate whether future markers in each variety are situated at comparable points on the cline of grammaticalization. If the hierarchy of constraints conditioning the variable occurrence of future temporal reference and their direction of effects in PA are parallel to those of LA, and if the

same future marker is situated at similar points on the grammaticalization pathway in each variety, this strengthens the case for contact-induced change in PA in the future temporal reference sector due to the influence of LA. On the other hand, if the same future marker in LA and PA is at different stages of grammaticalization, we can interpret this as evidence that the future temporal reference system is not behaving in the same way in both varieties.

2.4 Critical summary

2.4.1 Methodological and analytical gaps in the literature

This chapter reviewed the pertinent literature on *imala*, verbal negation and future temporal reference. It also highlighted that there are several methodological and analytical deficiencies that are particularly prominent in previous works on the morpho-syntactic variables in spoken Arabic. These shortcomings are summarized below.

(i) *The use of disparate methodologies, hampering cross-study comparability*

In contrast to the bulk of research on morpho-syntactic variation in spoken Arabic, in the literature on the targeted phonological variable, there is some research that deploys quantitative reasoning in the examination of word-medial *imala* in spontaneous speech data (Hennessey 2011). However, the problem is how researchers deploy this quantitative reasoning in their work. In Hennessey's (2011) study on this variable in the speech of Palestinians in Beirut, he only focused on variant frequencies and the social constraints on variant choice. Consideration of linguistic factors is also crucial as it allows the analyst to transcend reliance on overall usage rates, which could fluctuate for a variety of extralinguistic reasons, in order to consider the extent to which the fine-grained linguistic conditioning of the variable in question in one variety (here

PA) resembles that of another (here LA), affording crucial insights into the linguistic process of convergence.

- (ii) *The neglect of community-based usage data in previous treatments of morpho-syntactic variation in spoken Arabic*

In previous studies on the two targeted morpho-syntactic variables, there is a general lack of data-driven reports of variant frequencies and conditioning of variant selection (but see Persson 2008). Starting with verbal negation, we saw that Lucas' (2010) study on this variable in PA relied on native speaker intuitions through acceptability judgment questionnaires, which is problematic because what individuals might judge as acceptable or unacceptable may not reliably reflect what they actually do in their natural speech. Labov (1966) discusses this point by highlighting the problem of categorical perception. He explains that individuals may perform with subtle quantitative differences which are not accessible to conscious intuition (see also Poplack and Torres Cacoullos 2015:272). Many of these differences are not commented on at all, whereas others are exaggerated to the degree of categorical occurrences (Poplack 1982:2). Additionally, responses in questionnaires are vulnerable to Labov's (1972c) principle of subordinate shift. This means that when speakers are asked questions about their dialect/language, their responses may be heavily influenced by prestige norms. As such, they might deny using forms like the post-verbal negative enclitic *-š*, which is said to be stigmatized in major Levantine urban centres (see e.g., Wilmsen 2014 and Chapter 5). For these reasons, it is important to complement Lucas' (2010) questionnaire with a benchmark of actual speech data in order to examine the linguistic constraints on variant selection and to quantify the strategies speakers employ unreflectingly (i.e., actual usage constraints).

Perhaps even more problematic are previous treatments of future temporal reference in vernacular varieties of Arabic. This literature is generally characterized by detailed exegesis of decontextualized examples (but see AbuAmsha 2016; Persson 2008). Additionally, most authors rely on their own intuitions in attributing a meaning to a future variant (see e.g., Jarad 2013, 2014; Mitchell and El-Hassan 1994), which is circular (Poplack and Turpin 1999:146), rather than referring to independent contextual indicators in the ambient discourse that support the inferred meaning (see Poplack and Tagliamonte 2001:217).

(iii) *The flouting of the Principle of Accountability*

In some previous studies on morpho-syntactic variables in Arabic, especially on verbal negation, even though speech data were collected, some researchers have not adhered strictly to the Principle of Accountability. As explained in Chapter 1, this principle requires the analyst to report every case in which a variable form occurs out of the total number of environments where the variable element could have occurred but did not (Labov 1972b:94). In Khairallah's (2014) study on verbal negation in LA, she circumvents the Principle of Accountability by reporting the variants of interest only, counting, for example, the occurrence of -š with labial or non-labial-initial verbs (Khairallah 2014:44–45) to directly address previous claims in the literature, and omitting any consideration of how the use of this variant is contoured by the co-variants within the verbal negation system.

(iv) *A tendency to impose form-function symmetries on the analysis of the morpho-syntactic variables*

The imposition of form-function symmetries on variable systems is especially salient in previous treatments of future temporal reference in spoken Arabic, and is also found, albeit to a lesser extent, in the literature on verbal negation. In Lucas' (2010) study, he uses his

acceptability judgment questionnaire to support theory-based claims he made previously (Lucas 2007). For example, one of his informants judged acceptable the use of *-š* to negate past tense verbs, but he dismisses this as ungrammatical because it contradicts his theoretical expectations regarding the inadmissibility of *-š* with such verbs in PA (see Lucas 2007). Lucas (2010), insisting on form-function symmetry of forms used for verbal negation, ignores evidence from other varieties of Levantine Arabic, where scholars have documented variability in verbal negation, including using *-š* alone to negate past tense verbs and non-labial-initial (pseudo-)verbs (see e.g., Wilmsen 2014:102).

In previous studies on future temporal reference in Arabic, researchers also tend to rule out inherent variability in speech, claiming that each future variant has a different meaning or function, and they explain the alternation between variants in the same context as a result from the differences in the way the future eventuality is envisaged (see Poplack and Malvar 2007:134), which also explains why there is a lack of consensus regarding the linguistic predictors of variant choice. Their arguments predicated on form-function symmetries stand in contrast to form-function asymmetries characterizing the heterogeneity of speech, where multiple forms can be used for the same function or referential meaning (Labov 1969; Tagliamonte 2006:9).

(v) *Problems in assessing change, convergent or otherwise*

In the literature on word-medial *imala* and future temporal reference in Arabic, we observed that claims about linguistic change, whether contact-induced or otherwise, were made in the absence of relevant benchmarks necessary to assess change. In Hennessey's (2011) study on contact-induced change in PA as spoken in Beirut, not only is there a lack of an LA benchmark (i.e.,

presumed source variety of the change) against which to compare the speech of Palestinians in Beirut, but there is also a lack of appropriate real-time data (pre-contact variety of PA) necessary to examine the convergence of PA on LA. Similarly, Jarad (2013, 2014) does not appeal to diachronic (or even synchronic) data sources to examine the advancement of the exponents of futurity on the grammaticalization pathway in Syrian Arabic. Leddy-Cecere (2018, 2020) also makes strong claims about contact-induced grammaticalization in the future temporal reference sector in Arabic based on surface similarities between Arabic dialects (e.g., several Arabic dialects sharing the use of *rah* and *ha-*). However, it is important to note that surface similarity of forms should not be equated with structural similarity (Poplack et al. 2012:223) because there may be differences in the underlying grammatical conditioning of the exponents of futurity common across Arabic dialects.

As I explain in more detail in Chapter 3, to assess change, it is important to appeal to the underlying structure of variability, emerging from constraints conditioning variant choice, and conduct systematic quantitative comparisons across relevant synchronic and diachronic baselines in order to demonstrate whether there is linguistic change in a particular sector of the grammar, and whether this change is contact-induced or the product of internal drift, i.e., the tendency of languages to evolve in the same direction over long periods of time (or parallel internally-motivated changes; Labov 2001:416).

2.4.2 Building on the literature

The present study addresses the shortcomings in the literature on spoken Arabic highlighted above and builds on this literature by drawing on the variationist framework with its focus on spontaneous speech data and quantitative modelling. This framework enables us to examine the structured nature of the variable systems of *imāla*, verbal negation and future temporal reference

by examining variant frequencies and linguistic (and social) constraints on variant choice in natural speech data collected from Palestinian and Lebanese Arabic speakers in Beirut. Quantitative analysis allows me to validate the contextual criteria which have been said to affect variant selection in (Levantine) Arabic in the previous literature, and it may show that there is no correspondence between the impressionistic claims made in the literature and the motivations constraining variant choice in spontaneous speech.

As mentioned in Chapter 1, a key methodological advantage of the variationist framework is comparison, which is essential to the present study. I make use of the comparative variationist framework to examine whether there is change in PA by comparing the linguistic conditioning of the three targeted variables in the contact PA variety (synchronic speech data) with its pre-contact precursor (diachronic data), and I also examine whether this change is contact-induced by comparing the linguistic conditioning in the contact stage of PA with that in LA. In particular, I look at the constraint hierarchies (pattern of favouring and disfavouring effects on variant choice) and the direction of effect (the order from more to less of factors within a factor group) (see Tagliamonte 2012:122). As stated in Chapter 1, only if the conditioning patterns in the contact PA variety differ from those in its pre-contact precursor, while they are simultaneously parallel to those in LA, can we conclude in favour of contact-induced change in contemporary PA as spoken in Beirut (see Poplack and Levey 2010).

In line with the comparative framework, I also compare constraint rankings across different age groups to track patterns of variation and change in apparent-time. Linguistic change in PA, if any, could be detected in apparent-time if young speakers are diverging from older speakers and are participating in ongoing change (see Poplack and Levey 2010:409). Other social factors, such as sex, education and (Palestinians') residence in refugee camps are also

taken into consideration in investigating the social motivations for variant choice in the three variables. In the next chapter, I explain the main tenets of the comparative variationist framework, and I describe the synchronic and diachronic datasets used in this study.

3 Theoretical Framework and Data

This chapter is divided into three main sections. Section 3.1 explains the theoretical framework of comparative variationist sociolinguistics. Section 3.2 presents the synchronic and diachronic data sources used in this study, and section 3.3 finishes with a chapter summary.

3.1 Theoretical framework

This study draws on the framework of comparative variationist sociolinguistics (Labov 1966, 1972b, 1972c, 1982; Poplack and Tagliamonte 2001; Weinreich et al. 1968). This framework is based on several key tenets and working principles. Its main tenet is that *inherent variability* is a hallmark of speech. This means that speakers engage in a number of choices among discrete alternatives in discourse which have the same referential meaning or grammatical function within their respective grammatical sectors (Poplack and Tagliamonte 2001:88). This alternation among choices in discourse relates to a principal construct in the variationist framework, which is the *linguistic variable* (Labov 1963, 1966), defined as “two or more ways of saying the same thing” (Labov 1972b:271). Because multiple forms within the same variable system can encode the same meaning or grammatical function, they display what is known as *form-function asymmetry* (Labov 1969; Tagliamonte 2006:9). In other words, there is no privative relationship between form and function in a variable system, contrary to much normative or prescriptive treatments of variability which seek to establish a one-to-one relationship between form and meaning or grammatical function.

The alternation between two or more variants within a variable system is not random; rather variation in speech is characterized by *orderly or structured heterogeneity* (Weinreich et al. 1968:99–100). This means that the choices speakers make between variable linguistic forms

are systematically constrained by multiple linguistic and social factors that reflect the underlying grammar associated with a given variable (Bayley 2013:85–86; Poplack et al. 2012:207; Sankoff 1988:151).

3.1.1 Data collection in the variationist framework

In terms of data collection, the starting point for variationist analysis is the *speech community* (Labov 1966, 2006), which is a group of speakers who share a set of evaluative norms regarding language use. We refer to the speech community to situate the speech patterns of an individual speaker in the context of her community and to distinguish idiosyncratic or otherwise exceptional occurrences from patterns that are regular among members of a speech community.

We also refer to the makeup of the speech community to examine extra-linguistic factors that may facilitate contact-induced change, such as the length and intensity of contact with other speech communities, as well as the status of the speakers belonging to that speech community (i.e., majority group or minority group) and the status of their language variety (i.e., majority or minority variety). For example, long-term and sustained contact, as well as the minority status of a speech community (and the minority status of their spoken variety) are widely believed to increase the chance of structural convergence of the minority variety on the majority variety (Thomason 2001). As such, taking the speech community as a reference point is important because contact-induced change, if any, can only be established based on diffusion, and for diffusion to proceed, a community is necessary (Poplack 2021:50; see also Trudgill 1986).

In the variationist framework, the optimal speech style to examine structured heterogeneity in speech is the *vernacular*, which is “real language in use” (Milroy 1992:66) and in “which the minimum attention is given to the monitoring of speech” (Labov 1972b:208). Labov (1984:29) emphasizes that the vernacular is the “most systematic data for linguistic

analysis.” That is because the vernacular is the speech style that is most free from hypercorrection or style-shifting (Tagliamonte 2006:8). In other words, the linguistic and social factors that constrain variation are said to “operate most robustly, and most systematically, in vernacular speech, where corrective norms exert less influence” (D’Arcy and Tagliamonte 2015:256; see also Tagliamonte 2006:8). The vernacular is also the repository of non-standard forms, which may be stigmatized, so tapping into the vernacular means that speakers may use linguistic forms that are of interest to the analyst. Labov (1972b:92) points out the problem of the *Observer’s Paradox* in this regard, and he explains that the aim of linguistic research is to investigate how people use language when they are not being systematically observed, but we can only obtain such speech data by systematic observation. Variationists address this problem by using the *sociolinguistic interview* (Labov 1984) to collect vernacular data. Even though it is called an ‘interview’, it is actually designed to resemble a casual conversation (Schilling-Estes 2007:171). The interview consists of a series of hierarchically structured sets of questions, which Labov (1973) refers to as conversational modules (see also Labov 1984:33). One of the aims of these modules is to elicit narratives of personal experience, where it is anticipated that the speakers become so involved in recounting their experience that they pay minimum attention to speech, optimizing the production of the vernacular.

In addition to personal narratives, what also helps in eliciting vernacular speech patterns and in attenuating the Observer’s Paradox is if the interviewer is an in-group community member (see Cukor-Avila and Bailey 2001; Wolfson 1976). As I explain later in this chapter, the vernacular tends to be used more frequently by speakers when they are interviewed by either a familiar person, or an in-group community member who shares similar characteristics with them, such as nationality, dialect, experience, social background, etc. (Childs 2019:32; Cukor-Avila

and Bailey 2001:267). In other words, the speech data that an in-group member collects may better represent the typical linguistic behaviour of speakers than if they were collected by an out-group interviewer (see Cukor-Avila and Bailey 2001:268; Fasold 1972; Rickford and McNair-Knox 1994; Wolfson 1976:203). For example, in the study of Fasold (1972), African-American interviewees used vernacular features with African-American interviewers more than they did with European-American interviewers. As such, the identity of the interviewer may influence the type of data collected (Fought 2004:455).

3.1.2 Data analysis in the variationist framework

In terms of data analysis, a key procedure involves *circumscribing the variable context*, which requires elimination of “all the contexts in which the [...] alternant forms contrast, i.e. do not say the same thing” (Labov 1978:6). This necessitates abiding by a central principle in the variationist paradigm, namely the *Principle of Accountability* (Labov 1972b:72), which entails that every variant that is part of the variable context must be reported with the proportion of cases in which it occurred in the relevant environment, compared to the number of cases in which it might have occurred (Labov 1972b:94). In other words, the analyst should include not only all the cases where the variant occurred, but also where it could have occurred but did not.

Because competing variants of a linguistic variable are constrained by features of the linguistic and extra-linguistic contexts in which they occur, these contexts can be operationalized as factors (independent variables) that represent the hypotheses regarding what motivates variant choice (see Labov 2004). We empirically test the hypotheses regarding the linguistic and social constraints on structured variation through *quantitative modelling* (Tagliamonte 2006:12). This allows the analyst to identify the nature of the underlying grammar by examining the distribution and conditioning of competing variants. In other words, quantitative modelling allows us to

express in quantitative terms the strength of association between a variant form and the surrounding linguistic environment, as well as the social phenomena that are associated with this variant form (Bayley 2013:86).

Within quantitative analysis, we conduct a distributional analysis to examine how often variants occur in the linguistic contexts we operationalized, and the social factors associated with variant selection. But it is important to also subject the data to multivariate analysis (MVA) because it allows us to examine which independent variables are significant in variant choice when all of the factors are considered simultaneously. MVA allows the analyst to uncover patterns of variability; a pattern is defined as “a series of parallel occurrences (established according to structural and/ or functional criteria) occurring at a non-negligible rate in a corpus of language use” (Poplack and Meechan 1998:129). In this thesis, MVA yields two key lines of evidence crucial to the interpretation of the results. The first is statistical significance at the .05 level of social and linguistic factors. The second and more important line of evidence that MVA yields is the direction of effects, or the order from more to less of factors within a factor group (Tagliamonte 2006:235, 2012:122), which shows the pattern of favouring and disfavouring effects on variant choice, or the constraint hierarchy. The constraint hierarchy provides a window on a portion of the underlying grammar (Poplack and Tagliamonte 2001:92).

3.1.3 The comparative variationist method

The constraint hierarchy and the direction of effects can be used for comparative purposes as key diagnostics in assessing change (contact-induced or otherwise). Poplack and Meechan (1998:130) state that variant frequencies and especially the conditioning of linguistic variability are to some extent language-specific, and once they are determined via quantitative analysis, they serve as a diagnostic of language membership. The comparative method is essential to the

present study, which aims to examine whether there is evidence that PA is converging on LA in different linguistic components. In this thesis, convergence is defined as “the achievement of greater structural similarity in a given aspect of the grammar of two or more languages, assumed to be different at the onset of contact” (Silva-Corvalán 1994:4–5).

Poplack and Levey (2010:398) detail several prerequisites that need to be met in order to infer that a linguistic feature is the result of contact-induced change. A solid case for contact-induced change from PA to LA should rest on the demonstration that: (i) the feature is present in the presumed source variety of the change (LA); (ii) the feature is absent in the pre-contact variety (prior stage of PA) or a contemporary non-contact variety; or if present, the feature is not conditioned in the same way in the pre-contact or non-contact varieties as in the source variety; and (iii) the feature in the contact variety (contemporary PA) parallels in a non-trivial way the behaviour of its counterpart in the source variety (LA). The comparison of the linguistic conditioning of variant selection across relevant benchmarks (see below) transcends the comparison of rates, which can vary for extralinguistic reasons, but variant conditioning or the underlying grammar is expected to be more constant and it reflects more stable constraints on variant selection (Poplack and Levey 2010:400; Poplack and Tagliamonte 2001:93–94).

In this study, I compare the linguistic conditioning of the three targeted variables in contemporary PA with: (i) a diachronic benchmark to examine the possibility of change, and with (ii) LA to examine the possibility of contact-induced change in PA. As explained in Chapter 1, if the hierarchy of constraints, particularly their direction of effect, associated with the candidate for change (word-medial *imala*, verbal negation and future temporal reference) in the contact PA variety is the same as that of the pre-contact PA variety, while differing from that of LA, no change in PA has taken place. If the candidate for change in contact PA features a

constraint hierarchy that differs from those of both the pre-contact PA variety and LA, then change has occurred in PA, albeit not one that is contact-induced. Only if a candidate for change in contact PA features a constraint hierarchy which differs from that of the pre-contact PA variety, but which parallels that of LA, can we conclude that there is contact-induced change in PA due to the influence of LA (see Poplack and Levey 2010:400–401).

In addition to real-time comparisons between contact and pre-contact PA varieties, I also compare constraint rankings across different speaker age groups to track patterns of variation and change in apparent-time.⁵¹ The apparent-time construct is based on the assumption that the vernacular remains relatively stable throughout the course of an adult lifetime (Cukor-Avila and Bailey 2013:248; Sankoff and Blondeau 2007). Comparing constraint hierarchies across younger and older generations offers a window on change in progress, or an inference of change over time in the speech communities under investigation.

3.2 Data

3.2.1 Synchronic speech data

3.2.1.1 Recruitment of speakers

The synchronic data in this study come from 39 Palestinians and 27 Lebanese speakers recorded in Beirut between 2016 and 2019. Palestinians are stratified by age/immigrant generation, speaker sex, education (distinguishing speakers with and without post-secondary education) and place of residence (distinguishing those speakers living inside refugee camps in Beirut as well as those living outside the camps). Data from camp-dwellers were collected mainly in Burj El

⁵¹ It should be noted that comparisons of constraint rankings across datasets/speaker groups are done only where the data permit. In some cases (e.g., see Chapter 5 on verbal negation), variability is so limited (in the younger speaker datasets) that I cannot apply statistical methods (MVA) to assess the data, but I leverage what variation there is to elucidate any noteworthy patterns.

Barajneh camp in the southern suburb of Beirut, but data from some speakers were also collected in other two camps in Beirut: Shatila and Mar Elias. These camps are shown on the map in Figure 3.1.

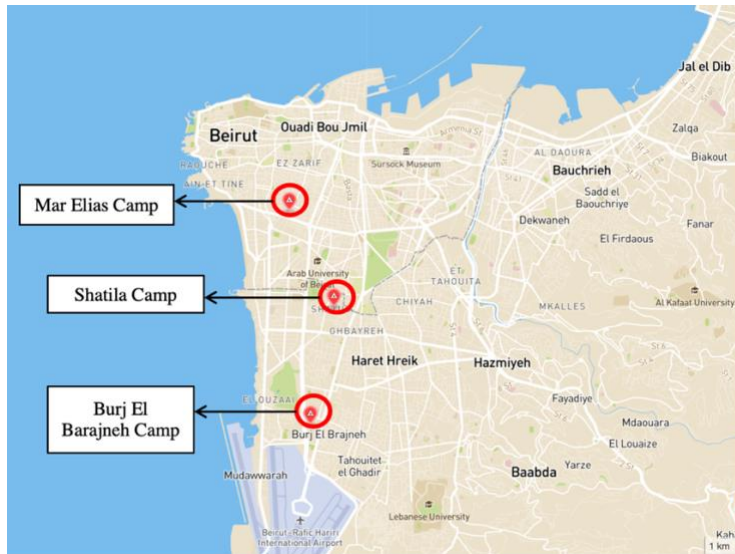


Figure 3.1: Refugee camps targeted in this study⁵²

For the Lebanese control group, the same stratifying measures were imposed on the speaker sample, excluding refugee camp residence. Details of the Palestinian and Lebanese sample constitution are summarized in Table 3.1. These details were obtained via a socio-demographic questionnaire administered to speakers on completion of the sociolinguistic interview.

⁵² Adapted from *Economic Corporation Foundation*. <https://ecf.org.il/issues/issue/1432> Copyright 2010 by Economic Corporation Foundation

Table 3.1: Characteristics of the Palestinian and Lebanese sample populations

Palestinians		Lebanese	
Age/Generation	N	Age	N
Old (73-92)/1 st Generation	9	Old (72-89)	5
Middle-aged (40-66)/2 nd Generation	19	Middle-aged (40-58)	5
Young (20-39)/3 rd Generation	11	Young (19-34)	17
Sex		Sex	
Male	19	Male	11
Female	20	Female	16
Education		Education	
Post-secondary	21	Post-secondary	22
Primary/Secondary	18	Primary/Secondary	5
Place of residence		(Camp residence is not applicable for Lebanese speakers)	
Inside camp	15		
Outside camp	24		
Total	39	Total	27

Informants were recruited through snowball sampling (Milroy and Gordon 2003:32), particularly through the ‘friend-of-a-friend’ method (Milroy 1980). My status as a Palestinian born and raised in Beirut (outside refugee camps) facilitated my access to both the Lebanese and Palestinian speech communities in Beirut.

Interviews with many Lebanese speakers were successful in eliciting narratives of personal experience. My familiarity with several of the Lebanese speakers made the interviews with them successful possibly because in interactions with a familiar interlocutor, the interview is more likely to mirror an everyday linguistic interaction (Cukor-Avila and Bailey 2001:258; Kyratzis and Ervin-Tripp 1999:1325; Schilling-Estes 2013:124), where non-standard or

vernacular features are expected to be used (Childs 2019; Coupland 1980; Rickford and McNair-Knox 1994; Russell 1982). This emphasizes the importance of rapport between interviewers and interviewees in minimizing the impact of the Observer's Paradox (Labov 1972b). Additionally, interviews with Lebanese speakers who were not familiar to me were also successful. These speakers were not aware of my Palestinian heritage before the interview. When debriefed about the project after the interview, most of them asked why I was interested in Palestinian speech patterns, and they were surprised to learn about my Palestinian heritage possibly because of my Lebanese-influenced speech patterns. As such, my perceived insider status as Lebanese (during the interview) may have helped in eliciting more casual speech.

For Palestinians outside camps, I relied on my friends and their social networks, as well as my family members and their networks to recruit speakers. As such, my in-group status may have facilitated recruiting a more diverse group of Palestinian speakers (in terms of age and educational level) than Lebanese speakers. Many of the Palestinian speakers outside the camps who were not familiar to me learned that I was Palestinian during the interview when they asked me whether I was Palestinian as well. Interestingly, the fact that I use word-medial *imala* while many Palestinian interviewees did not, or did so at reduced rates (see Chapter 4), and explicitly commented on my use of *imala*, afford some tentative evidence that my own speech patterns did not unduly influence theirs, despite their recognition of my own Palestinian heritage. However, their knowledge of my Palestinian heritage may have elicited subject matter and narratives that they might have been less willing to share with a non-Palestinian. For example, a Palestinian man living outside the camps who knew that I was Palestinian had the following to say about Palestinians living inside versus outside camps in Lebanon:

- (1) *ʔa:ʔed bi mujtamaʕ muylaq falaʃti:ne w msakkar ʕa ʔa:lo, w ma beʔessiʃ ʔa:lo bi yurbe, beʔess ʔa:lo ɖumun ʔahlo, ma beʔess ʔa:lo ɣari:b. Baynama l-falaʃti:ne lli barra:t l-*

moxayyam byetʃa:mal bi jafa:we ktir, bi ʔasa:we, bi ʃanjahiyye, bi ʃonʃoriyye, beṯessi: inno feʃlan ʔa:ʃed yarib, [...] beṯess ḥa:lo inno ʔa:ʃed yarib bi ha l-mujtamaʃ, manno uh mutaja:nes maʃahom, w miʃ ʔa:der uh yextelef ʃannon laʔanno ḥaya:to bedda testamerr maʃon.

[30/P/M/56/1:40]

‘He (the Palestinian inside the camp) is in a closed Palestinian society and he has locked himself (in that society), and he doesn’t feel that he is alienated, he feels that he is within his family, he doesn’t feel that he is an outsider. On the other hand, the Palestinian outside the camp is treated very harshly, ruthlessly, condescendingly, he is racially discriminated against, you feel that he is actually alienated, he (the Palestinian outside the camp) feels that he is alienated in this (Lebanese) society, he does not live in harmony with them (the Lebanese), and he cannot be different from them because his life has to continue with them.’

My in-group status as a Palestinian also facilitated my access to refugee camps in Beirut. My first and main point of access to the camps was a Palestinian family friend. He introduced me to three other people (two of whom are camp-dwellers) who also facilitated my access to the camps. The three speakers, in addition to the family friend, introduced me to the majority of the 15 camp residents I interviewed. All of these speakers knew that I was Palestinian because I was introduced to them as such by our mutual acquaintance.

3.2.1.2 Nature of the data

Speech data were collected through sociolinguistic interviews. Most interviews lasted between 30 minutes and around three hours. While most interviews were conducted with one speaker at a time, some interviews (especially in the camps) were conducted with three or four speakers who knew each other. The interviews focused on speakers’ experiences in an effort to elicit narratives of personal experience and attenuate the effects of audio-monitoring that might be induced because informants were aware that they were being recorded. Many interviews are a rich repository of vernacular forms. These are illustrated in the excerpt below. A Palestinian man,

who lives in Burj El Barajneh camp, talked about his friend's late cancer diagnosis and his subsequent death. The interview was with a small group of people known to the speaker.

(2) *mitil ha:la:t [Akram]⁵³, [Akram] telef minteşer intişa:r b jismo [...] bas l-fikra [Akram] ma laha?u:ş yaşmlu:lo işi. Halla?- halla? bteşrfe şu- şu ?a:l d-dactor? [...] Roḥet zurto, ?alle “Şam noḡlob şa:ḥbak hada d-dactor, miş Şam biji:b.” [...] ?oltello “ma:şi halla? minşu:f.” Da?eit la d-doctor, “kifak, bidna d-doctor [Bassem], fi Şenna rfi:qna [Akram], hek hek l-?oşşa.” ?a:l “[Yasser] wala yhemmak, ana baḥke maşo, w bokra Şa s-sabŞa-sabŞa w noşş baku:n ana wiyya maş başed w minşu:fo, xalaş?” [...] ?oltello la [Akram] “bokra biku:n Şendak.” ?a:l “mumta:z.” ?ejo ta:ni yōm la Şendo, w şa:fo d-dactor, ?axadu:lo Şayyene, w Şemlu:lo fḥoşa:t [...]. Şa l-moyreb babŞat la ?ibno, ba?ollo “şu Şamme, ṭammenne şu?” ?alle “[...] huww telef maşo mn d-daraje r-ra:bŞa.” ?eday hek, Şam bakteb ana wiyya Şa l-whatsapp, ?eday- da?eit la d-dactor, ?alle “[Yasser] şa:ḥbak xa:leş”, ?oltello “ṭab Şamaliyye, işi, kaza”, ?alle “ḥara:m hada ymedd ?i:do Şaleih, iza bethebbu: mahadaş y?arreb şobo.” ?alle “wala Şila:j wala işi, ma elo yeir d-duŞa:? w ha l-musakkina:t bas la ḥatta Rabna yeftekro [...] w iza hay, xalli: yeṭlaş Şa l-beit, w xalle motto tku:n bi l-beit, xalaş, fiş lo fa:yde bi l-mustaşfa.” Feşlan telef Şala l-beit. Ma fiş leiltein w t-ta:lte, ka:n [...]*

[58/P/M/45/1:41:07]

‘Cases like Akram’s, for Akram it turned out that it (the cancer) was spread throughout his body. But the thing is that with Akram, they couldn’t do anything for him. Now- now do you know what- what the doctor said? I went and visited him (Akram), he said, “We are calling this doctor friend of yours, he is not responding.” I told him, “Okay, we will see now.” I called the doctor, “How are you? We want Doctor Bassem, we have our friend Akram, and this is the story.” He said “Yasser, don’t worry, I will talk to him, and tomorrow around seven (or) half past seven I will be here with him (Bassem) and we will see him (Akram), okay?” I told Akram, “He (the doctor) will be here tomorrow.” He said, “Excellent.” They came over the next day, and the doctor saw him, they took a sample, and they did some tests. At sunset, I send (a text message) to his (Akram’s) son, I tell him “What? Tell me what (happened)?” He said “It turned out he has Stage IV (cancer).” My hands were like that (trembling), I am texting him on WhatsApp, my hands- I called the doctor, he told me “Yasser, your friend is done for”, I told him “What about a procedure or something?” He told me “(It wouldn’t help if anyone did anything for him), if you love him, (don’t do anything to him).” He told me “Without treatment, without anything, he only needs prayers and some pain relievers until (he dies), and if anything, let him go

⁵³ Names have been changed to maintain the privacy of the speaker and the people in the narrative.

home, and let his death be at home, it's useless for him to be in the hospital.” He (Akram) actually went home. In two or three nights, he was ... (dead).’

In the excerpt above, I have highlighted some vernacular features, including variables targeted in this thesis (indicated in boldface). These are listed below:

a. *Absence of word-medial imala*

In extract (2), all word-medial environments that are theoretically hospitable to *imala* show the absence of vowel-raising characteristic of the Lebanese benchmark variety (see Hennessey 2011:59) e.g., *mumta:z* ‘excellent’.

b. *Verbal negation*

The excerpt features the variable expression of verbal negation, where the speaker alternated between the variants *ma-š*, *miš*, and *-š*, e.g.:

- i. *ma laḥaʔu:-š yašmlu:lo iši* ‘they couldn’t do anything for him’
- ii. *miš šam biji:b* ‘he is not responding’
- iii. *fī-š/ ma fī-š* ‘there isn’t...’

c. *Future temporal reference*

The speaker also used the *b*-prefix in the excerpt above to express future temporal reference, e.g.:

- i. *bokra biku:n šendak* ‘he will be (here) tomorrow’.

d. *Examples of other vernacular features*

- i. The English-origin established loanword *d-dactor* ‘the doctor’,⁵⁴ where the Arabic definite determiner is assimilated to the first coronal consonant

⁵⁴ As gauged from dictionary listedness.

- ii. Variation between /q/ and the glottal stop, e.g., *rʃi:qna* (vs. *rʃi:ʔna*) ‘our friend’ and *l-ʔoʃʃa* (vs. *l-qoʃʃa*) ‘the story’.

3.2.2 Diachronic data

In addition to the synchronic datasets, this study uses diachronic data sources to elucidate the evolutionary trajectory of the three targeted variables (see Milroy and Gordon 2003:176). Several variationist studies of European languages have combined synchronic and diachronic data sources to understand the process of language change (e.g., Digesto 2019; Poplack and Dion 2009; Poplack and Malvar 2007; Poplack and St-Amand 2007), but this approach has yet to acquire traction in Arabic sociolinguistics because the study of historical orality in Arabic remains incipient (e.g., Holes 2018; Lentin 2018; Owens 2006).

The ideal source for studying language change is a prior stage of LA and PA vernacular because linguistic change originates in vernacular speech. In other languages where speech recordings from the past are not available, historical investigations have relied on written documents as a representation of speech (e.g., Culpeper and Kytö 2010; Hickey 2010; Moore 2011), though not all historical documents may be a fair representation of speech. In Arabic, there are written historical documents, but they involve varieties like pre-Islamic Arabic (Al-Jallad 2020; see Herin 2019) and Andalusí Arabic (see Vicente 2020), so they cannot be used as surrogates for spoken Arabic varieties which emerged in a context of first language acquisition.⁵⁵ There are also written dialectological works on spoken varieties of Arabic, such as LA (e.g., D’Alverny 1970; Feghali 1928; Thackston 1996) and PA (e.g., Bauer 1926; Blanc 1953; Piamenta 1966; Schmidt and Kahle 1918; Seeger 1996), but they mainly focus on describing the

⁵⁵ van Putten (2020:57) acknowledges that representation of spoken Arabic dialects is almost absent in the written record for much of the Middle Ages. Vicente (2020:226) similarly states that there is hardly any written material reflecting the everyday dialectal level in Arabic, though with some ingenuity, surrogates of non-contemporary varieties of spoken Arabic might be located (such as the diachronic sources used in the present study).

dialect in question and its features. They are also typically based on only a few speakers (Vicente 2019:107), whose speech may feature stereotypical phrases and may not reflect everyday unmonitored speech (Behnstedt and Woidich 2013:313). As such, these works are a poor approximation to naturalistic speech that is targeted in variationist sociolinguistic research. Accordingly, I have resorted to diachronic sources, such as oral history interviews (POHA; see section 3.2.2.1), as well as televised plays intended for popular consumption (LPTC; see section 3.2.2.2).

3.2.2.1 Palestinian Oral History Archive (POHA)

The first diachronic data source comes from a sub-set of the *Palestinian Oral History Archive* (POHA), comprising audio-visual interviews recorded between 1995 and 2006 (not transcribed) with first-generation (older) Palestinians now residing in Lebanon. The POHA was chosen because it is the only known corpus of interviews with older Palestinians in Lebanon exclusively, where the interviews are in colloquial PA.

The POHA offers several advantages for this study. As mentioned in Chapter 1, the interviews in this corpus resemble sociolinguistic interviews, in the sense that they are not structured and there is little interference from the interviewer. As we will see later, these interviews include non-standard features of vernacular PA and they display the inherent variability characteristic of unreflecting speech. This renders them comparable to the contemporary data, which allows me to compare the structure of prior variability with counterparts in the Palestinian synchronic corpus.

The vernacular used in the POHA has another advantage for this study. Because the vernacular is the first speech style acquired in preadolescence, it is expected to remain the most stable, “forming the basis of historical continuity among successive language states” (Poplack

and St-Amand 2007:710). As such, the vernacular is the truest indication of how far a change has progressed (Coveney 1996:278). In this sense, the vernacular in the POHA can be used as an approximation to pre-contact PA. Speakers in the POHA were born between 1897 and 1948 (at the time of their interviews, their ages ranged between 47 and 101 years old). Out of the 80 speakers in the POHA sampled for this study (see Table 2), 60 of them fled Palestine when they were 17 years old or older (the oldest speaker, born in 1897, left Palestine when he was 51 years old).⁵⁶ The remaining speakers were preadolescents or teenagers (whose ages ranged between 11 and 16 years old) when they left Palestine, with the exception of two speakers, one who was 7 years old, and one who mentioned that she was a few months old when she left (she was born in 1948). As such, the majority of the 80 speakers retained in the present investigation were around, or beyond, the stage of vernacular stabilization, which occurs roughly at the age of 17 (Denis et al. 2019:45; Tagliamonte and D’Arcy 2009:66). Vernacular stabilization means that basic grammatical patterns are relatively stable or fixed, and there are limits to acquiring variables in the phonological, morphological and syntactic sub-components (Tagliamonte and D’Arcy 2009:65; see also Kerswill 1996). Labov (2001:447) explains that vernacular stabilization is the period when “the declining ability to acquire native syntactic intuitions hits bottom.” Evidence suggests that by the age of 6, children acquire the phonological features of their local variety (Kerswill 1996:190), and for speakers over the age of 14, it becomes increasingly difficult to acquire a new dialect (Chambers 2003:179). Accordingly, it is expected that speakers in the POHA would retain conservative Palestinian vernacular norms. In other

⁵⁶ As I mention later in this chapter, speaker metadata are available on the POHA website (including every speaker’s year of birth), so I could easily subtract a speaker’s year of birth from the year 1948 (when the Arab–Israeli War occurred).

words, to the extent that their vernaculars have remained stable, their speech will most likely reflect the state of PA when they acquired it as children (see Labov 1963, 1966).

Extra-linguistic characteristics of speakers in the POHA also afford comparability between their speech and the PA synchronic dataset. Similar to Palestinian speakers in the latter, all the speakers in the POHA were living in Lebanon after they fled Palestine. An effort was also made to select interviews with speakers who live in Beirut (whether inside or outside refugee camps) because Palestinian speakers in the synchronic dataset also reside in Beirut, so the area of residence in Lebanon can be controlled for across the two comparison groups.

Interviews in the POHA included several themes, such as pre-war life and childhood in Palestine, traditions and folklore, relationships between Palestinians and the Jewish community in pre-war Palestine, personal accounts of the Palestinian exodus in 1948, and other personal narratives. An example of a personal narrative is illustrated below. The variables targeted in this study are indicated in boldface, and other vernacular features are underlined.

- (3) *rohna marra, ʕazamu:na ʕala bala- ʕala manʔaʔa hona:k esma l-baʔʕa bi l-ʔuds. Hay ʕazamu:na bedu:n ma yeʕerfo l-iza:ʕa, ʔaʕle uh ʕa:ʔiliyye yaʕne, rohna bi s-sirʔa eʕna meʕa:n ma ʔada yeʕref yaʕne. Laʔki:lak ʕu ʕa:r maʕna. Naʕabo ʔ-ʔa:wla:t, ʕi ʕaʕar ʔa:wla:t, w ʕa:ʔela:t, ʕa:yef, ʔaʕadna. Ana bi hada:k l-waʔet, abaskariʕ, w ma baʕrab yaʕni xamer. ʔaʕadna, fi bana:t betʕazzem, ʕaffo l-akel w ja:bo l-ka:sa:t, [...] L-ka:sa:t l-fa:ɖe la l-ʕaraʔ, baʔa mi:n yalli beʕobb? L-bana:t. [...] Weʕlet la ʕande ʕam betʕobb, ʔoltella “ya sitt ana btesmaʔi:le, baʕrabiʕ ʕaraʔ”, ʔa:lat “beddak teʕrab.” ʕabbat noʕʕ l-ka:se w kasrata b-mayy. Ballaʕna bi l-ʕasha, kollon yeʕrabo illa ana, ka:ste ʔodda:me [...] Barmat hiyye w ejet laʔat l-ka:se baʕda mala:ne, ʔa:l “leh ma bteʕrab ka:stak?” ʔoltella “ana ʔoltellek xamer ma baʕrab, ana ʕaraʔ baʕrabiʕ”, ʔa:l “beddak teʕrab.” Ma hiyye ʕa:rat sakra:ne kama:n. Fi batzakkar xza:ne hiyy ʔba:lna, saʕbat ha j-ja:ru:r tabaʕ l-xza:ne, ʔa:lat fard hek ha l-ʔad bi:je, yaʕne sitte mi:le taʔri:ban. Ejet- ʔaʔtato b zonna:ra w ejet la ʕande. Mesket l-ka:se, [...] ʔa:l “bteʕrab walla ʔaʔaɖde l-fard bi ra:sak?” Yaʕne ʕufta sakra:ne, ana ʔolt yemken ʔarrek ʔi:da hada:k l-waʔet, hasabna hsa:b, ʕa:yef, ʔoltella “la ya sitt, baʕrab, baʕrab xalaʕ.” [...] ʔallat wa:ʔfe ʔodda:me la ʕribt l-ka:se ʕa s-ʕa:fi. Hay ʕa:rat maʕe bi l-ʔuds. W l-- skirna, ʕaʕa w ʕaraʔ w katha,*

skirna, ma ḥasseina ḥa:lna illa kol wa:ḥad mšallah laḥa:l, ḍalleina la ṣ-ṣubeḥ. Hay ṣa:rat maʿe bi l-ḥuds.

[POHA/05/16.04.2004/M/77/16:04]⁵⁷

‘We went one time, we were invited to- to an area there called Al-Ba’a’ (Al-Baqa’a) in Jerusalem. They invited us without the knowledge of the radio station (where the speaker worked),⁵⁸ it was a family party, we sneaked out so nobody would know. Let me tell you what happened with us. They set up the tables, around ten tables, and families, you see, we sat down. At the time, I don’t get drunk, and I don’t drink alcohol. We sat down, there are women who are taking (guests to their tables), they put the food and they got the glasses. The empty glasses are for arrack (liquor), so who is pouring (arrack)? The women. A woman reached to my table and she is pouring (arrack), I told her “Excuse me miss, I don’t drink arrack”, she said “You will drink.” She poured up to half of the glass and she mixed it (the arrack) with water. We started with dinner, everyone is drinking except for me, my glass is in front of me. The woman turned around and came and found the glass is still full, she said “Why aren’t you drinking from your glass?” I told her “I told you, I don’t drink alcohol, I don’t drink arrack”, she said “You will drink.” She was also drunk by now. I remember there was a cabinet facing us, she opened this drawer of the cabinet, she grabbed a gun of that size (a small gun), around 6 mm [sic]. She came- she put it around her belt and came to me. She held the glass, she said “Will you drink or do I unload the gun in your head?” I saw that she is drunk, I thought she might make a (wrong) move with her hand at the time, I was apprehensive, you see, I told her “No miss, I will drink, I will drink.” She kept standing next to me until I drank the whole glass. This happened with me in Jerusalem. And the- we got drunk, dinner and arrack and so, we got drunk, each of us was lying down, we stayed until dawn. This happened with me in Jerusalem.’

The vernacular features in excerpt (3), including variables targeted in this study, are listed below:

a. *Absence of word-medial imala*

- i. Non-raised variant, e.g., *l-iza:ʕa* ‘the (radio) station’; *t-ṭa:wla:t* ‘the tables’

⁵⁷ Utterances are transcribed verbatim from the POHA recordings. Codes are to be interpreted as follows: POHA=Palestinian Oral History Archive/speaker number/date of recording (Day. Month. Year)/M=Male; F=Female/speaker age/time stamp of the excerpt in the recording.

⁵⁸ The speaker mentioned before that he was part of a band and they worked at a radio station, but they were not allowed to attend private parties.

b. *Verbal negation*

The speaker alternated between the variants *a-š*,⁵⁹ *ma*, and *-š*, e.g.:

- i. *abaskariš* ‘I don’t get drunk’
- ii. *ma bašrab* ‘I don’t drink’
- iii. *bašrabiš* ‘I don’t drink’

c. *Examples of other vernacular features*

- i. The discourse markers *yašne* (lit. ‘meaning/means’) and *ša:yef* ‘you see’
- ii. The use of the voiced alveolar sibilant /z/ instead of the voiced dental fricative /ð/, e.g., *l-iza:ʕa* ‘the (radio) station’ (vs. *l-iða:ʕa* in MSA)

The three variables targeted in the present study (*imala*, verbal negation and future temporal reference) were extracted from interviews with 80 older Palestinians living inside and outside refugee camps in Beirut. Details of the sample constitution are summarized in Table 3.2.

Table 3.2: Characteristics of the speakers in the POHA sampled for this study

Sex	N
Male	40
Female	40
Education	
Illiterate	6
No school	36
Primary	24
Secondary	8
Post-Secondary	6
Place of residence	
Inside camp	40
Outside camp	40
Total	80

⁵⁹ As mentioned in Chapter 2, *a-š* was only used once in the POHA, and it was used by this male speaker.

Details about the speakers in the POHA were recoverable from the speaker profiles and the interviews, available on the POHA website.⁶⁰ In this website, there is a digital link for each interview, which provides additional information pertaining to speaker metadata, such as his/her name, date and place of birth (in Palestine), where s/he resides in Lebanon, as well as when the interview was recorded and the link to the interview.⁶¹ Only a few profiles did not include some key information, such as speaker age or place of residence. However, missing speaker information was recoverable from the interviews (e.g., the interviewer would ask the speaker when and where s/he was born). Details about the educational level of the speakers were extracted from the interviews, where they were asked whether they attended school and if so, which educational level they had reached.

3.2.2.2 Lebanese Popular Theatre Corpus (LPTC)

The *Lebanese Popular Theatre Corpus* (LPTC) includes 34 televised plays (not transcribed) intended for popular consumption, dating from 1960 to 1983. Their genres range from comedy, satire to social realism. Details about the plays are summarized in Table 3.3.

⁶⁰ The POHA website can be found here: <https://libraries.aub.edu.lb/poha/>

⁶¹ Other information found in the speaker profile includes the interviewer's name, duration of the interview, the permanent link of the interview, the subjects/themes discussed during the interview, landmarks or private institutions discussed by the speaker (e.g., schools or churches) in pre-war Palestine, etc.

Table 3.3: List of plays in the Lebanese Popular Theatre Corpus (LPTC)

Year	Title		Playwright
1960	<i>Tawaheen Al Layl</i>	‘Mills of the Night’	Wadih El Safi
1962	<i>Jisr Al Qamar</i>	‘Bridge of the Moon’	Rahbani brothers ⁶²
1964	<i>Bayya’ Al Khawatem</i>	‘The Ring Vendor’	Rahbani brothers
	<i>Wadi Al Ghazzar</i>	Name of a fictional village	Maurice Awwad
1965	<i>Dawaleeb Al Hawa</i>	‘The Pinwheels’	Rahbani brothers
1966	<i>Ayyam Fakhreddeen</i>	‘Days of Fakhreddeen’	Rahbani brothers
1968	<i>Al Shakhs</i>	‘The Person’	Rahbani brothers
1969	<i>Jiwar El Ghaym</i>	Name of a fictional village	Rimon Haddad
	<i>Jibal El Sawan</i>	Name of a fictional village	Rahbani brothers
1970	<i>Faraman</i>	‘Firman’ (equivalent to ‘decree’)	Romeo Lahoud
	<i>Ya’eesh Ya’eesh</i>	‘Long Live, Long Live’	Rahbani brothers
1971	<i>Min Layaleena</i>	‘From Our Nights’	Walid Gholmieh
	<i>Madeenat El Farah</i>	‘The City of Happiness’	Elie Choueiri
	<i>Al Awasef</i>	‘Storms’	Romeo Lahoud
	<i>Sah-n-Nom</i>	Equivalent to ‘Rise and Shine’	Rahbani brothers
1972	<i>Ayyam Saif</i>	‘Summer Days’	Elias Rahbani
	<i>Natouret El Mafateeh</i>	‘The Guardian of the Keys’	Rahbani brothers
1973	<i>Akh ya Baladna</i>	‘Oh Our Country’	Fares Youwakim
	<i>El Fnoun Jnoun</i>	‘Art Is Madness’	Fares Youwakim
	<i>Sahriye</i>	‘Soirée’	Ziad Rahbani
	<i>Al Mahatta</i>	‘The Station’	Rahbani brothers
1974	<i>Loulou</i>	Name	Rahbani brothers
	<i>Mawsem El Tarabeesh</i>	‘The Season of the Fez’	Azar Habib
	<i>Sitt El Kel</i>	‘The Woman Above All’ (‘Respectful/Admired Woman’)	Wassim Tabbara
	<i>Nazl El Sourour</i>	Literally ‘Happiness Hotel’	Ziad Rahbani

⁶² Assi and Mansour Rahbani

1975	<i>Mais El Reem</i>	Name of a fictional village	Rahbani brothers
	<i>Helwe Ktir</i>	‘Very Pretty/Beautiful’	Wassim Tabbara
1976	<i>El Amira Zmurrud</i>	‘Princess Zmurrud’	Romeo Lahoud
1977	<i>Esmak bi Albi</i>	‘Your Name Is in My Heart’	Romeo Lahoud
	<i>Yasmine</i>	Name	Romeo Lahoud
	<i>Bint El Jabal</i>	‘Mountain Girl’	Romeo Lahoud ⁶³
1978	<i>Bi-n-Nesbeh La Bokra...Chou?</i>	‘What About Tomorrow?’	Ziad Rahbani
1980	<i>Film Ameriki Tawil</i>	‘A Long American Film’	Ziad Rahbani
1983	<i>Shi Feshil</i>	‘Failure’	Ziad Rahbani

All plays selected are performed in colloquial LA and include representations of different colloquial situations, as well as vernacular linguistic features. However, it is important to note that there are numerous caveats attached to the use of the LPTC for reconstructing historical orality. The first problem concerns the limited availability of the plays. Lebanese theatre dates back to the 1800s (Chalala 2018),⁶⁴ but plays before 1960, to my knowledge, are not available online. Plays from the 1960s onwards were accessible on YouTube, and an effort was made to include as many plays performed in colloquial LA as possible in the present study.⁶⁵ However, to my knowledge, there is little to no explicit information about any play, such as its genre, the playwrights’ style, the linguistic features of their work, and the social attributes of the characters in the play (e.g., social class, age, etc.). This information is crucial because ideally, the plays selected should mirror the informal spoken LA of different social classes so that they can accurately represent an earlier stage of colloquial LA (see Culpeper and Kytö 2010:7). Nevertheless, despite the paucity of publicly accessible information about the plays, much useful

⁶³ The play *Bint El Jabal* ‘Mountain Girl’ is based on George Bernard Shaw’s 1913 stage play *Pygmalion*.

⁶⁴ Jalal Khoury (1933-2017): Brechtian Realist Forged by 1967 War, and the Birth of Modern Lebanese Theater <https://www.aljadid.com/node/2103>

⁶⁵ Some plays on YouTube were available in audio format only, and these were excluded because I could not see the characters (i.e., I could not determine who is speaking).

information about the social status of the characters (inferred largely from their occupations; see Chapter 4), their sex and approximate age could be gleaned from careful content analysis of the plays while watching them.

Another problem in the LPTC is that in some plays, some morpho-syntactic features tend to occur infrequently. For example, the post-verbal negative enclitic *-š* used alone was not common in the plays. This brings us to the problem of negative evidence (see Ayres-Bennett 2004:14–15; Poplack and St-Amand 2007:708). This means that a form may be absent or rare either because this was the case in the earlier stage of the variety in question, or it just so happens that the playwright did not use it, or it was not appropriate to the genre of the play (see Romaine 1988:1462), or perhaps prestige norms may have blocked the use of some non-standard forms.⁶⁶ This is especially because in the plays, negation with *-š* (whether *-š* used alone, or *ma-š*, or *a-š*) was typically used to characterize lower class and uneducated speech (see Chapter 5), such as the speech of a peasant, homeless person, beggar, thief, garbage collector, shepherd, etc. These characters also used *a-* or *a-š* for verbal negation, so there is always the possibility that these non-standard features were exaggerated for comedic or parodic purposes (see Ayres-Bennett 2004:31, 74, 81), which may distort the spoken language of the time.

An example of negation with *-š* used to characterize lower status or uneducated speech comes from the 1977 play *Bint El Jabal* ‘Mountain Girl’, which is based on George Bernard Shaw’s 1913 stage play *Pygmalion*. In the Lebanese play, the protagonist, Lisa, is a flower seller who speaks a “vulgar” variety of Lebanese Arabic characteristic of her village. She meets a linguist (phonetics professor), Henry Adib, who helps her become presentable in the high society by teaching her to speak “proper” Lebanese Arabic, which does not include the stigmatized

⁶⁶ As mentioned in Chapter 2, *ma-š* and *-š* are reported to be stigmatized as rural speech in Levantine capitals (Wilmsen 2014:152).

forms she uses. The excerpt below shows how Henry tells Lisa to avoid using stigmatized variants for the expression of verbal negation:

- (4) [Lisa] *Walaw, a-betxalli:ne ʔu:la-š marra te:ne?* [Henry] *ma betxalli:ne, ma betxalli:ne.* [Lisa] *Mi:niyy a-bexhallik-š?* [Henry] *Lek ya Lisa, ana šam bšaheḥlik, šam bʔellik beʔu:lo “ma betxalline ʔu:la marra te:nye?”* [Lisa] *ʔaš fley, ma hayde yeir ši, haydi:k felem, w hayde ḥdi:s. ʔena bʔu:l “a-betxalli:ne ʔu:la-š,” w inta betʔu:l “ma betxalli:ne ʔu:la” [...]* Ma kollo ʔawl b-ʔawl, *ma hak-š?* [Henry] *W lek ya Lisa, inte hallaʔ je:ye tetʔallame l-lafʔ l-mazbu:t, ma hek? Ntabahte inno ʔeltillik “ma hek?”* Ma ʔeltillik *“ma hak-š?”* Yašne inte le:zim taʔrfe tmayze bein l-lafʔ š-šaḥi:h, w l-lafʔ l-mubtazal. Hallaʔ iza baddik, ʔuli:ha marra te:nye, w iza betri:de nsi:le kelmit *“ma hak-š.”*
[LPTC/Bint El Jabal/1977/185, 189/F, M/1:04:44]⁶⁷

‘[Lisa] Won’t you let me say it [*a-betxalli:ne ʔoula-š*] one more time? [Henry] Won’t you let me [*ma betxalli:ne*], won’t you let me [*ma betxalli:ne*]. [Lisa] Who’s the one who won’t let you [*a-betxallek-š*]?’ [Henry] Lisa, I am correcting you, I am telling you they say ‘won’t you let me [*ma betxalli:ne*] say it one more time?’ [Lisa] It’s okay, this is different, this one (the way you say it) is for teaching, and the other one (the way I say it) is speech. I say *a-betxalli:ne ʔu:la-š* ‘won’t you let me say it’ and you say *ma betxalli:ne ʔu:la* ‘won’t you let me say it’ [...] It’s all speech, isn’t it [*ma hak-š*]?’ [Henry] Lisa, now you are coming to learn proper pronunciation, aren’t you [*ma hek*]?’ Did you notice that I said *ma hek* ‘aren’t you?’ I didn’t say *ma hak-š* ‘aren’t you?’ So you have to know how to distinguish between proper pronunciation and vulgar pronunciation. Now if you want, say it one more time, and please forget the word *ma hak-š* ‘aren’t you’.’

The excerpt above suggests that negation with *-š* is stigmatized, while negation with the pre-verbal *ma* only is prestigious. What is notable in this particular instance is that the characterization of non-standard speech is achieved not through phonological or lexical features, but in terms of morpho-syntactic variation.

Another challenge of the LPTC concerns its stylistic range. This compendium of plays contains speech from higher status personages (e.g., mayor, princess, king, etc.) and lower status personages (e.g., beggar, thief, prostitute, cobbler, farmer, etc.), so we would expect that the speech of some characters, especially that of higher status ones, would be at some remove from

⁶⁷ Utterances are transcribed verbatim from the LPTC. Codes are to be interpreted as follows: LPTC/name of the play/year/speaker number /M=Male; F=Female/time stamp of the excerpt in the play.

vernacular speech targeted in the present study. Some of the plays, especially the earlier ones, are also set in a village milieu, in what Khoury (2009)⁶⁸ calls the “idealised” Lebanese mountain village, where the characters include some villagers and peasants, the mayor or the governor, the guard, etc. Even though the language in these plays is LA (Salem 1999),⁶⁹ in some scenes, such as romantic rendezvous, characters tell stories which are somewhat lyrical, expressing their thoughts in highly stylized language and using formulaic expressions, which may not reflect natural conversational situations (Khoury 2009; see Ayres-Bennett 2000:327, 2004:29).

Despite the acknowledged drawbacks of the plays, the LPTC remains a valuable diachronic data source for several reasons. First, in contrast to some of the earlier plays, by the 1970s and 1980s, many of the plays were set in urban contexts (Salem 1999), capturing scenes that better resemble daily life in an urban community. This is especially true for the plays in the late 1970s and early 1980s by Ziad Rahbani, who questioned the myth of a simple, village-like Lebanon and attempted to capture everyday situations which most Lebanese people could identify with (Salem 1999). Some of his plays included political themes that were the talk of the town at the time (related to the 1975 civil war in Lebanon). In these plays, Salem (1999) explains that “[t]o the Lebanese, Ziad spoke their language” and he “typically [adopted] the colloquialisms of his generation.” Bazzi (2014)⁷⁰ similarly states that this playwright echoed “the language of the people on the streets.” According to Salmon ([1965] 1987), this is characteristic of the more skilful playwright who depicts contemporaneous daily life and authenticates the action in the plays by approximating everyday speech (as cited in Culpeper and Kytö 2010:17).

⁶⁸ Mansour Rahbani: Dramatist and writer of the classic songs of a Lebanese golden age

<https://www.theguardian.com/music/2009/apr/17/obituary-mansour-rahbani>

⁶⁹ Imagining Lebanon Through Rahbani Musicals

<https://www.aljadid.com/content/imagining-lebanon-through-rahbani-musicals>

⁷⁰ Ziad Rahbani: His life as a performance

<https://english.alaraby.co.uk/english/artsandculture/2014/12/29/ziad-rahbani-his-life-as-a-performance>

Second, the plays include many scenes with face-to-face interactions and turn-taking, as well as pauses, hesitations and interruptions, which reflect everyday spoken interactions (see Ayres-Bennett 2000:327, 2004:29; Culpeper and Kytö 2010:11, 16). Third, systematic examination of the LPTC confirms that the latter displays the vernacular features and inherent variability characteristic of contemporary LA. In other words, there are systematic correspondences between modern LA and the LPTC in terms of variable usage (see e.g., Chapter 4), which validates the latter as a rendition of LA speech.⁷¹ This is in line with Schneider's (2013:72) argument that we can assess the validity of historical sources based on the presence of dialectal forms and variation, as well as internal consistency across different sources. If there are similarities in constraint hierarchies and frequencies across different sources, we can assume that there is some approximation to the oral usage of the time. Fourth, we find non-standard features used in the LPTC, such as certain variants of verbal negation (e.g., *ma-š*, *-š* and *a-š*), possibly because the speech of the characters in some plays, especially those that take place in a village, may reflect varieties outside of Beirut (i.e., varieties characteristic of rural areas in Lebanon). As mentioned previously, there is evidence of social conditioning of variant selection in the LPTC in terms of verbal negation (see Chapter 5), where stigmatized variants, especially *ma-š* and *a-š*, were used more frequently by lower-status characters.

As is the case with any diachronic surrogate of orality, the plays must be used with caution, but whatever deficiencies they may entail are offset by the diachronic perspective they afford on an early variety of LA. The use of discursal features typical of everyday interactions in many plays, as well as features that characterize colloquial LA show that the LPTC can be

⁷¹ Systematic correspondences mean that the variable features in question pattern in the same way in the plays and in the LA synchronic dataset (and this was examined by comparing variant rates and the fine conditioning of variant selection across the LPTC and the LA synchronic dataset).

considered as an approximation to that variety, and it can be used to compare the structure of prior variability in LA with patterns in the Lebanese synchronic dataset.

Some vernacular features in the LPTC are illustrated in the excerpt below from the play *Bi-n-Nesbeh La Bokra...Chou?* ‘What About Tomorrow?’ (1978). In this excerpt, Zakaria, who works as a bartender in a pub, is talking about an incident that happened with his son to his friends: Najib, a chef at the pub, and Ramez, a vegetable vendor. The scene takes place in a bar.

- (5) [Zakaria] xayye he:da z-za:hir- fhemna l-Ɂussa l-mazbu:ta rjeɁna. He:da z-za:hir ke:yni:n bi l-maɁab, w absar šu gheltiɁ ha n-na:zra w Ɂateto mla:haza la šaki:b inno Ɂam byetwahšan bi l-leɁbe. Eh sa:Ɂabit l-football maɁo, šabaɁa fiya Ɂa noss di:n wejja. BalaɁetlak ha s-suffayra ya mɁallim, w lehɁeto. ɁelɁit tsawfir w terkud wara. Ša:r kelme:lo yerkuɁ w ši, Ɂallak “badda tezneɁne bši ze:wye.” Šu Ɂemil? Šu Ɂemil? Allah lahamo, “ma tku:n kesle:n ya šaki:b!” Fatal min wara:ha, sahaba min šaɁra w labaja bi l-ɁareɁ. Ayya mɁawwera? MaɁab zefet baddak tɁu:l, ayya mɁawwera? He:da ku:Ɂa min hon, ku:Ɂ Ɂi:da, ma Ɂaš fi ku:Ɂ, deyre Ɂi:da ša:rit. Ana bas šefta la l-mɁalme, xabaɁ Ɂa ra:se. [Najib] ya latif! Šefta int la l-mɁalme? [Zakaria] W la betɁu:l ke:nit mɁalme. [Najib] Hki:t maɁa yaɁne? [Zakaria] Šefta, hki:t, ana hki:t, hiyye bteheke?! Hiyy betɁawfir hiyy, ma Ɂaš teheke he. [Najib] Ya hara:m. [Zakaria] He:ketne l-mudi:ra [...] Ɂe:letle “he:da Ɂibnak baɁd raḥ naɁti: mehle la j-jemɁa, ya huww baddo yetɁayyar, ya faɁsu:lo Ɂa ši madrase te:nye.” NaɁaletle Ɂale:me:to tabaɁ ha j-jumɁa Ɂa warɁa. Yehru? Ɂardo, huww law Ɂando ši beɁawwid Ɂan shi, ke:n beɁu:l l-we:hid. Lek ha l-Ɂale:me:t. [Najib] Farji:ne ta šu:f. [Zakaria] Lek malla Ɂale:me:t. [Ramez] Miš mne:h Ɂale:me:to, ha? [Zakaria] Afiš Ɂale:me:t, refɁe:to Ɂam ye:xdo l-Ɂale:me:t kella, miš Ɂam byebɁe:lo Ɂale:me:t. [...] Raḥ yemredne, yelɁan bayy bayyo, raḥ yemredne. Šu baddo, šu fi b-ra:so ma baɁrif [...]

[LPTC/Bi-n-Nesbeh La Bokra...Chou?/1978/14/M/43:29]

‘[Zakaria] Brother, apparently this- we later understood the true story. Apparently, they were in the playground, and I don’t know how this superintendent made the mistake of giving a warning to Shakib (Zakaria’s son) that he is being aggressive in the game. It just so happened that the football was with him, he hit her with it in her damn face. She swallowed this whistle man, and she ran after him. She kept whistling and running after him. He kept running and everything, he thought “She will trap me in some corner.” What did he do? What did he do? God inspired him “Don’t be lazy Shakib!” He turned around her back, he pulled her from her hair and dropped her to the ground. Has he hurt her? You can say it’s an asphalt playground, has he hurt her? Her elbow from here, her elbow on her arm, there isn’t an elbow anymore, her arm became straight. When I saw the teacher, I went crazy in the head. [Najib] Oh my goodness! Did you see the teacher? [Zakaria] You would never think she was a teacher. [Najib] Did you speak with her?

[Zakaria] I saw her, I spoke, I spoke, could she speak?! She whistles, she cannot speak anymore. [Najib] What a pity! [Zakaria] The principal spoke with me, she told me “We will give this son of yours time until Friday, either he will change, or you search for another school for him.” She copied his marks for this week on a piece of paper. Fuck him, if he had something to cover up (for his behaviour), (we wouldn’t have said anything). Look at these marks. [Najib] Let me see. [Zakaria] Look at these marks. [Ramez] His marks are not good, (are they)? [Zakaria] There are no marks, his classmates are taking all the marks, there aren’t any marks being left for him. He will make me sick, son of a bitch, he will make me sick. I don’t know what he wants, what’s in his head [...]

Examples of the targeted variables and other vernacular features are listed below:

a. *Word-medial imala*

- i. Raised variant, e.g., *ke:nit* ‘it/she was’
- ii. Non-raised variant, e.g., *ša:rit* ‘it/she became’, where /a:/ is preceded by an emphatic, which is said to disfavour *imala* (Habib 2012:55)

b. *Verbal negation*

The speaker alternated between the variants *a-š*, *miš* and *ma*, e.g.:

- i. *afiš* *šale:me:t* ‘there are no marks’
- ii. *miš* *šam byebʔe:lo šale:me:t* ‘there aren’t any marks being left for him’
- iii. *ma* *bašrif* ‘I don’t know’

c. *Future temporal reference*

The speaker used *raḥ* and *badd-* for the expression of future temporal reference, e.g.:

- i. *raḥ* *našʔi: mehle la j-jemša, ya huww baddo yetyayyar [...]* ‘we will give him time until Friday, either he will change [...]
- ii. *raḥ* *yemredne* ‘he will make me sick’

d. *Examples of other vernacular features*

- i. *w ši*, equivalent to the general extender ‘and everything’
- ii. Profanity, e.g., *yeħru? řarđo*, equivalent to ‘fuck him’

3.3 Summary

The combination of synchronic and diachronic data sources in this study addresses the paucity of socio-historical research on Arabic dialects, and crucially, it allows me to establish continuity between earlier stages of LA and PA and contemporary varieties by extending the time-depth of the analysis to some 100 years (as mentioned previously, the oldest speaker in the POHA was born in 1897, and the oldest speaker in the PA synchronic dataset was born in 1926, while the oldest speaker in the LA synchronic dataset was born in 1929). This provides a richer understanding of the process of language change in Arabic, particularly in PA. In the ensuing chapters, I explain the coding procedures for the three targeted variables and discuss the results for each variable.

4 *Imala* in Word-Medial Position

The first linguistic variable which is targeted in the present study to examine the extent to which PA is converging on LA in Beirut is *imala* in word-medial position, or the raising and fronting of the long vowel /a:/ to [e:]. This chapter presents the results for this variable in the Lebanese and Palestinian diachronic and synchronic datasets.

This chapter is structured as follows. Section 4.1 defines the variable context, and section 4.2 explains how the social and linguistic factors were coded, as well as the hypotheses they test. Section 4.3 presents the quantitative results regarding the variable raising of word-medial /a:/ to [e:] in the diachronic and synchronic datasets. Section 4.4 is an excursus on word-final *imala* (and its relation to word-medial *imala*). Section 4.5 discusses the results and presents some concluding remarks.

4.1 Circumscribing the variable context

As mentioned in Chapter 3, a key methodological procedure in the variationist framework is circumscribing the variable context, which entails defining the contexts in which the variants of a linguistic variable may alternate without change in referential meaning or function (Labov 1969). In the present study, all instances of raised word-medial /a:/ (where the raised variant equates with [e:] characteristic of the Lebanese control variety; see below)⁷² and non-raised word-medial /a:/ were extracted from the synchronic and diachronic datasets. The variants are illustrated in the examples below (in boldface), which are reproduced verbatim from the data:

⁷² As I explain below, because the raised variant in LA is typically realized as [e:] (Barkat 2008:678; Borg 2005:538; Hennessey 2011:59; Levin 2006:314; Naïm 2011:925), a vowel was counted as raised in PA only if it matched the corresponding [e:] in the Lebanese control variety (and I examined a random set of tokens in Praat to corroborate my auditory coding procedure).

- (1) [e:]
ke:nit ḥelwe **l-ḥaye:t.** [43/L/M/85/12:14]
 was.3sf nice the-life
 Intended: 'Life was good.'
- (2) [a:]
 ḥasse-t-o **insa:n** laṭi:f.
 felt.1s-him person nice [30/P/M/56/0:16]
 'I felt that he is a nice person.'
- (3) [a:]
 ejo **r-ra:hba:t** akhathu:-ni. [POHA/25/25.01.2003/F/68/06:50]
 came.3p the-nuns took.3p-me
 'The nuns came and took me.'
- (4) [e:]
 yalla **šabe:b.** [LPTC/Akh ya Baladna/1973/09/M/41:13]
 come.on guys
 'Come on, guys.'

Tokens were coded based on an auditory basis, but spectrographic analysis was conducted on a random set of tokens using Praat (Boersma and Weenink 2016) to check whether the variants had been accurately coded. From an acoustic point of view, Barkat (2008:678) explains that *imala* in modern Arabic vernaculars is characterized by the lowering of F1 (vowel formant) and the raising of F2.⁷³ Spectrogram analysis was also conducted to check whether raising and fronting of the long vowel /a:/ to [e:] in word-medial position in PA speech (when it occurs) is phonetically similar to the raised [e:] variant in LA speech. If the raised variant had a low F1 value (in Hertz or Hz) and a high F2 value, and the values were similar in LA and PA

⁷³ Rogers (2000:304) states that formants "indicate clusters of harmonics with high intensities, relative to their neighbours. Formants are particularly useful in distinguishing vowels. These are numbered from the one with the lowest frequency: F1, F2, F3, etc." Zsiga (2013:136) explains that F1 is inversely proportional to vowel height. F1 is high for low vowels, and low for high vowels. Zsiga (2013:136) continues that "*the distance between F1 and F2 is inversely proportional to vowel backness*. The more back the vowel, the closer F1 and F2 are; the more front the vowel, the further apart F1 and F2 are. Moving the tongue body back tends to equalize the space in front of the tongue (F2 resonance) and the space behind the tongue (F1 resonance)." In short, the first formant (F1) is (inversely) related to vowel height, while the second formant (F2) is related to the degree of backness: the more front the vowel, the higher the F2 (see also Barkat 2008:669) (Praat Manual: https://corpus.eduhk.hk/english_pronunciation/index.php/2-2-formants-of-vowels/).

speech, then word-medial *imala* was considered to be phonetically similar across the comparison varieties. Table 4.1 shows the number of tokens extracted from each dataset.

Table 4.1: Number of tokens of raised and non-raised variants of word-medial *imala* extracted from the diachronic and synchronic datasets

Dataset	N
POHA	4440
LPTC	3741
LA synchronic data	1404
PA synchronic data	2285
Total	11870

4.2 Factor groups

Several factors were coded for to test hypotheses derived from the relevant literature about the social and linguistic predictors of variant selection in word-medial *imala*. These factors are explained in the following sub-sections.

4.2.1 Social factor groups

Synchronic data

Each speaker in the synchronic speech data was coded for age (and immigrant generation for Palestinians), speaker sex, level of education (primary/secondary versus post-secondary), and place of residence for Palestinian speakers only (distinguishing speakers living inside and outside refugee camps in Beirut).

In terms of age or immigrant generation for Palestinians, speakers were divided into three age groups: first-generation (older) Palestinians between the ages of 73 and 92; second-generation (middle-aged) Palestinians between the ages of 40 and 66; and third-generation (younger) Palestinians between the ages of 20 and 39. Age or immigrant generation is expected to play a key role in determining variant choice. For first-generation Palestinians, it is

hypothesized that they will be more likely to retain conservative Palestinian speech patterns (i.e., they are not expected to adopt the [e:] variant in word-medial position characteristic of Beirut Arabic) (see Hennessey 2011).

Older Palestinians in this study were born in pre-war Palestine between 1926 and 1945. As such, their ages ranged between 3 years old and 22 years old when they left Palestine in 1948 due to the Arab-Israeli war. Out of the nine older Palestinian speakers sampled for this study, three were 17 years old or older when they arrived in Lebanon, one was a 14-year-old teenager, and the remaining were 8 or 9 years old, with the exception of the youngest first-generation speaker who was born in 1945. Accordingly, at least a few of the older Palestinians were around, or beyond, the age of vernacular stabilization. As explained in Chapter 3, vernacular stabilization occurs roughly at the age of 17 (Denis et al. 2019:45; Tagliamonte and D’Arcy 2009:66). Around that age, basic grammatical patterns are said to be relatively stable or fixed, and there are limits to acquiring variables in the phonological, morphological and syntactic sub-components (Tagliamonte and D’Arcy 2009:65; see also Kerswill 1996).⁷⁴ Evidence suggests that by the age of 6, children acquire the phonological features of their local variety (Kerswill 1996:190), and for speakers over the age of 14, it becomes increasingly difficult to acquire a new dialect (Chambers 2003:179). Accordingly, it is expected that the older generation of Palestinians will use the most conservative Palestinian speech patterns. This is consistent with findings of previous studies focusing on dialect contact situations in the Arabic-speaking world, where older speakers were more likely to retain conservative speech patterns and resist (contact-induced) change, while younger speakers tended to lead change in progress towards the socially dominant speech norms (see e.g., Cotter 2013, 2020; Cotter and Horesh 2015; Fityan 1981; Hennessey

⁷⁴ Though there is a possibility of ‘longitudinal change’ (Sankoff 2005:1011) or ‘lifespan change’ (Sankoff and Blondeau 2007:562), in which “individual speakers change over their lifespans in the direction of a change in progress in the rest of the community” (Sankoff 2005:1011).

2011). This is because in dialect contact situations, if change has occurred, and if that change is contact-induced, its leaders are usually younger speakers, who diverge from their elders, and those in most intense contact with the majority dialect (see Poplack and Levey 2010:396, 402, 409).

Speaker sex in the Palestinian synchronic data is hypothesized to be a predictor of variant choice. In particular, it is expected that the *imala* variant (or the word-medial raised [e:] variant) characteristic of metropolitan LA (Hennessey 2011:35, 92; see Versteegh 2001:153) will be favoured by Palestinian females, while Palestinian males may show a preference for the non-raised variant, consistent with Hennessey's (2011:93) finding that the rate of *imala* in the speech of Palestinian females in Lebanon was almost double that in the speech of males. This effect appears to be related to sex-differentiation in speech, where females are more likely than males to use the supralocal feature (see Armstrong 1998:139; Chambers and Trudgill 2004:84–85; Labov 1972b:243; Milroy et al. 1994; Tagliamonte 2012:36; Trudgill 1972:180), which, in this case, is the word-medial *imala* variant associated with Beirut Arabic (Hennessey 2011:16). Hennessey (2011:16) explains that Beirut Arabic has become the prestige form of speech in Lebanon. Barkat (2008:681) also claims that in the Levant, socially dominant dialects are characterized by vowel raising. It is because of the prestige associated with Beirut Arabic, characterized by vowel raising in word-medial position (Hennessey 2011), that Palestinian females are expected to use the [e:] variant more than males.

In terms of educational level, it is expected that Palestinian speakers with post-secondary education are more likely to use word-medial *imala* than those with primary or secondary education. However, it is important to note that it is not education per se that contributes to variant selection because education functions as a proxy measure of degree of access to and

integration into the Lebanese society (see Al-Wer 2002:42; Hennessey 2011:111–12). As highlighted in Chapter 2, post-secondary education is only accessible to (younger) Palestinians in institutions outside the camps,⁷⁵ so Palestinians in these institutions are expected to regularly interact with LA speakers. In other words, because of the presumed regular exposure of Palestinians with post-secondary education to LA speakers in higher education institutions, it is expected that the former would be more likely to use the word-medial [e:] variant characteristic of LA.

It is important to note that in this study, education interacts with speaker age, such that all young Palestinian (and Lebanese) speakers have post-secondary education, while all older Palestinians, with the exception of one older Palestinian female, have primary or secondary education. It was extremely difficult to find less educated young Palestinians, because even in the camps, there are many educated young Palestinians, as one camp resident explained:

(5) *Fi na:s sam betfallamo, eh w sam byowşalo la daraja:t kti:r şa:lye [...] ana başref ktir na:s sam byaşmlo Ph.D., henne kollon min l-mokhayyam [...]*

[51/P/M/26/04:33]

‘There are people who are getting educated, yes and they are reaching very high (educational) levels [...] I know many people who are pursuing a Ph.D., they are all from the camp [...]

It is worth mentioning here that post-secondary education may be afforded by some Palestinians because UNRWA provides university scholarships for students who excel academically.⁷⁶ This does not mean, however, that there are no younger, less educated Palestinians, but to my knowledge, there are no official statistical figures regarding the rate of educational attainment among young Palestinians in Lebanon. A 2012 report by the International Labour Organization

⁷⁵ There are no universities inside refugee camps in Lebanon; there are only some UNRWA (United Nations Relief and Works Agency for Palestine Refugees in the Near East) schools (Hennessey 2011:32–33, 110).

⁷⁶ What We Do: Education

<https://www.unrwa.org/what-we-do/university-scholarships?program=33>

(ILO)⁷⁷ shows that only 13% of all Palestinians in Beirut only, regardless of age, are educated to post-secondary level (International Labour Organization (ILO) 2012:123), though it is not clear if this figure only concerns Palestinians inside the camps or whether it includes non-camp-dwellers as well. However, there were no less educated young Palestinians (or Lebanese speakers) in my own social network or any of the social networks of the speakers recorded for this study. In the absence of baseline data from younger, less educated (Palestinian and Lebanese) speakers, it is only possible to make provisional claims about the effect of education on variant behaviour among younger people.

Whether Palestinians live inside or outside camps in Beirut is also expected to be a key predictor of variant selection. Place of residence is correlated with the social networks of Palestinian speakers. Palestinians in refugee camps are expected to retain Palestinian vernacular patterns because the camp environment reinforces dense, multiplex social networks among Palestinian camp-dwellers (Hennessey 2011:1, 12, 19–20), as well as because of the relative social separation of refugee camps from mainstream Lebanese society. Hennessey (2011:1) explains that many camp residents have limited daily contact with those outside the camps, and in Beirut, “while both camp and urban sprawl over the past half century have served to physically envelope the camps within the greater metropolis, the camps’ residents have not necessarily felt a comparable sense of integration into the larger Lebanese society” (Hennessey 2011:1; see also Hanafi 2009). The social marginalisation of the Palestinian refugee camps (though the refugee camps in Beirut are not spatially isolated from the larger urban area),⁷⁸ and

⁷⁷ Palestinian Employment in Lebanon - Facts and Challenges: Labour force survey among Palestinian Refugees Living in camps and Gatherings in Lebanon
https://www.ilo.org/wcmsp5/groups/public/---arabstates/---ro-beirut/documents/publication/wcms_236502.pdf

⁷⁸ As Hennessey (2011:49) explains, refugee camps in Beirut “can be better thought of as ‘slum areas and under-developed urban sprawl’, constituting ‘urban enclaves’ rather than fenced-off and spatially isolated camps” (see Hanafi 2009).

the general hostility of the Lebanese society towards Palestinians (Hennessey 2011:20, 35, 42–43, 49) are expected to influence Palestinian camp-dwellers' linguistic behaviour and contribute to maintaining PA vernacular norms.

For Palestinians who live outside the camps, they may be more likely to adopt Lebanese speech patterns because they may have weak ties to other Palestinians and strong ties to the Lebanese community (Hennessey 2011:33–34, 103, 131, 140). Hennessey (2011:127–28) explains that Palestinians outside the camps lack the linguistic input that social networks inside camps provide, and their exposure to Lebanese speech patterns could prompt the adoption of these speech patterns. Additionally, Hennessey (2011:109) states that many Palestinians in Lebanon are aware early on in their lives that they speak differently from their Lebanese counterparts and that the identity their particular speech patterns project is sometimes met with a judgmental reaction. Accordingly, many Palestinians might be compelled to limit the degree of linguistic distinctiveness between themselves and their Lebanese counterparts (Hennessey 2011:109). This might be especially true for Palestinians outside camps who engage in regular, face-to-face interactions with Lebanese people. One of the Lebanese speakers actually brought up the rather judgmental reactions in Beirut when dialects other than Beirut Arabic are spoken. This Lebanese young male explained during the interview that he speaks Lebanese outside home and at work, but at home, he speaks a Syrian variety because he went to school in a city in southwestern Syria and acquired the dialect. The below excerpt shows what he said when I asked him why he doesn't speak his Syrian dialect outside home:

- (6) *Most of the people, especially bi Beirut yaʕne, miš they don't accept different accent, but ʕandon hek- ma byetʔabbalo aw- aw- byetmaskharo aw they look at you heke differently, fa again bterjaʕe l-mujtamaʕ w l-culture byejbrik sometimes to be like how the people are.*

[03/L/M/25/33:44]

‘Most of the people, especially in Beirut, it’s not that they don’t accept different accents (referring to dialects), but they have this- they don’t accept or- or- or they mock or they look at you differently, so again, you go back (to) the society and the culture which force you sometimes to be like how the people are.’

Despite the hypothesis regarding the effect of place of residence on Palestinians’ speech patterns discussed above, it is worth noting that Hennessey (2011:12–13) points out that some Palestinians who live in closer proximity to Lebanese people than to Palestinians still approximate to PA vernacular norms. This is because these vernacular norms may have covert prestige (see Trudgill 1972) that makes them resistant to change (Hennessey 2011:17). This means that what might be stigmatized as Palestinian speech features in the wider Lebanese society is used by Palestinians as markers of their group identity (Hennessey 2011:17; see Chambers and Trudgill 2004:85; Labov 1994:24). Some Palestinians, then, regardless of their geographical proximity to and contact with Lebanese people, may be favourably disposed to PA vernacular norms. As we will see later in this chapter, this was the case of some (young) Palestinians, who expressed their positive affective orientation towards their variety and Palestinian identity during their interviews.

Conversely, some Palestinians who live in refugee camps may adopt Lebanese speech patterns. As mentioned previously, the camps in Beirut are not completely isolated, and as Hennessey (2011:34) explains, Palestinians may be drawn to the prestige forms of either LA or PA, depending on social, economic, political and ideological factors that differentiate the members of this community in Beirut. As such, if Palestinians (inside the camps) have strong ties to the Lebanese community (Hennessey 2011:34), perhaps because of their employment conditions (Hennessey 2011:113–14) or because they go to university outside the camps, they may be more likely to adopt Lebanese speech norms. For example, young Palestinian camp-dwellers in the present study, all of whom are educated to post-secondary level, may adopt

Lebanese speech patterns, at least to some extent, due to their presumed regular interaction with Lebanese speakers in higher education institutions. However, the problem here, as previously highlighted, is that there isn't a benchmark of less educated younger speakers within the camps (as well as outside the camps) in this study. In other words, it is not possible to compare the speech behaviour of less educated young Palestinian camp-dwellers, who may have fewer opportunities to interact with Lebanese speakers due to their lower level of education (i.e., they do not go to post-secondary institutions), with the speech behaviour of educated young Palestinian camp-dwellers who have ample opportunities to regularly interact with Lebanese speakers in post-secondary institutions. This remains an important line of inquiry to pursue in future research.

For the LA synchronic dataset, as mentioned previously, the same social factors discussed above were coded for, excluding place of residence. To my knowledge, there are no quantitative sociolinguistic investigations of word-medial *imala* in LA. The relevant literature only includes claims that *imala* is a feature of indigenous Lebanese varieties (Barkat 2008:678; Borg 2005:538; Hennessey 2011:59; Levin 2006:314; Naïm 2011:925). As such, I explore the social (and linguistic) conditioning of word-medial *imala* in LA as an exploratory manoeuvre.

Diachronic data

For the POHA, the same stratifying measures as the PA synchronic dataset were imposed on the speaker sample, excluding age/generation because all speakers in the POHA are older Palestinians. As such, I only coded for speaker sex, educational level and refugee camp residence. As detailed in Chapter 3, the expectation is that older Palestinians in the POHA will retain their Palestinian vernacular features because when they left Palestine, the majority of them

(60 out of the 80 speakers) were around, or beyond, the stage of vernacular stabilization, which means that they would have been less likely to adopt Lebanese speech features.

For the LPTC, I also coded characters in the plays for sex and approximate age (distinguishing older, middle-aged and younger characters).⁷⁹ As mentioned in Chapter 3, sex and age group could be relatively easily identified from watching the plays. I also capitalized on the differential social status of characters in the plays based on the occupations they portrayed.⁸⁰

To code for social status, it was necessary to refer to a specific taxonomy to define this independent variable. Habib (2010:176) explains that social class/status is a hazy and controversial concept because it depends on composite indices, such as education, occupation, house upkeep, income, schooling, etc. In the LPTC, only occupation was relatively easily recoverable from the plays themselves (e.g., characters portraying café owners, waiters, mayors, farmers, etc.). As such, it was used as a broad-brush measure of the social status of the characters in the plays.

In Lebanon, there is no official social class taxonomy. However, Khuri (1960, 1969), a Lebanese anthropologist, devised a system of social ranking (specific to Lebanon) which includes four social classes that differ in terms of occupation: Class I, including leaders who may exercise regional influence or who may be engaged in national politics; Class II, including wealthy landowners and merchants, university graduates, medical doctors, engineers, lawyers, etc.; Class III, including small landowners and shopkeepers, minor clerks, secretaries and soldiers; and Class IV, including sharecroppers, peasants, day labourers, dustmen, vegetable

⁷⁹ As mentioned in Chapter 3, the plays included in the LPTC were televised, so I had visual clues to draw on to inform my coding protocols.

⁸⁰ This factor group was not included in the other datasets because of several problems. As I discuss in this chapter, in addition to the absence of an official social class taxonomy in Lebanon, social class/status is difficult to operationalize and depends on composite indices, such as occupation and income (among others), which I do not have access to for all speakers in the synchronic datasets and in the POHA.

peddlers, etc. Based on Khuri's (1969) taxonomy, I coded for three broad social groups: higher social status (including Class I and Class II); middle social status (including Class III); and lower social status (including Class IV).

Higher status characters in the LPTC include a mayor, owner of a restaurant/café/bar/pub, famous singer, etc. Middle status characters include a police officer or sergeant, real estate brokers, school teacher, etc. Lower status characters include a beggar, prostitute, waiter, flower seller, garbage collector, cobbler, etc. Any (tertiary) characters in the LPTC whose occupations were not readily available (e.g., random villagers) were excluded from the analysis. Additionally, there were some female characters who did not have any work, but they portrayed the wives or daughters of males who had work (e.g., wife of a judge). As such, female characters who did not have any work were classified according to the social status of their husbands or fathers.

In terms of the hypotheses for the LPTC, as I mentioned for the LA synchronic dataset, it remains to be seen how social factors will contribute to variant selection, as there are no claims in the relevant literature about the social predictors of variant choice in LA due to the dearth of quantitatively-driven studies of word-medial *imala* in that variety.

4.2.2 Linguistic factor groups

For all datasets, each token was coded for two linguistic factor groups: preceding phonological segment and following phonological segment. The effect of the preceding phonological segment on variable vowel raising has attracted much attention in the literature (see e.g., Al-Nassir 1985, 1993; Barkat 2008; Habib 2012; Levin 2006; Owens 1993, 2006). As mentioned in Chapter 2, there are several consonants that reportedly disfavour vowel raising in the preceding phonological segment, such as preceding emphatics [t̤ d̤ s̤ z̤], uvulars [q ɣ x], gutturals [ħ ʕ ʔ], and the secondary pharyngeal [r] (Hennessey 2011). Habib (2012:55) explains that the

disfavouring effect of these consonants on vowel raising is due to pharyngeal articulation. In addition to the secondary pharyngeal [r], gutturals and emphatics are primary pharyngeals, and uvulars are dorsal primary pharyngeals (Shahin 2011).⁸¹

An assimilatory pattern that is relevant to the effect of the pharyngealized consonants on *imala* is emphasis spread in Arabic. This pattern occurs when emphasis (i.e., pharyngealization), which is characterized by the [+ RTR] feature (Retracted Tongue Root) (Davis 1995:466), spreads to neighbouring sounds (Davis 1995:465; Zemánek 2005:205; see also Al-Bataineh 2019; Card 1983; Hellmuth 2013; Herzallah 1990; McCarthy 1994; Shahin 2002; Watson 1999; Younes 1982; Zawaydeh 1998). Emphasis also spreads to adjacent vowels, causing them to become retracted (i.e., emphatic sounds lower and back neighbouring vowels) (see Al-Bataineh 2019).⁸² But as Davis (1995:465–66) explains, the pattern of emphasis spread varies across Arabic dialects. Herzallah (1990) states that in her rural northern Palestinian dialect, emphasis spreads leftward from the emphatic consonant to the beginning of the word, while rightward emphasis spread is frequently restricted to a following low vowel (as cited in Davis 1995:466). Similarly, Davis (1995:473, 487) shows that in a rural southern PA variety and a rural northern PA variety, leftward spread of emphasis starts from the emphatic consonant and extends to the beginning of the word (see also Kulikov, Mohsenzadeh, and Syam 2020:17; McCarthy

⁸¹ Shahin (2011:1) explains that “pharyngeals are segments with primary or secondary articulation in the pharynx.” The difference between primary and secondary articulation is as follows: “Primary articulation refers to either (1) the place and manner in which the stricture is made for a consonant or (2) the tongue contour, lip shape, and height of the larynx used to produce a vowel. The primary articulation may still permit some range of movement for other articulators not involved in its formation. For example, an ‘apico alveolar’ articulation involves the tip of the tongue but leaves the lips and back of the tongue free to produce some degree of further stricture in the vocal tract. This latter is called a secondary articulation.”

<https://www.britannica.com/topic/articulation-speech>

⁸² Pharyngealized vowels, or vowels affected by pharyngealization from a neighbouring emphatic consonant, have been described as being backed and lowered. The lowering effect is caused by the retraction of the tongue root (RTR) as a result of the co-articulatory effect of the constriction of the pharynx witnessed when emphatic consonants are produced in Arabic (Al-Hashmi 2016:43). Al-Hashmi (2016:43) explains that acoustically, “pharyngealized vowels have been found to result in an increase on the F1 (open/close jaw; high/low) axis of the vowel diagram and a lowering of F2 (front/back position of the tongue) in a number of Arabic dialects” (see also Al-Ani 1970; Zawaydeh 1997). As such, “pharyngealized vowels are typically backer and lower than their counterparts [non-pharyngealized vowels], and the effect is most salient and consistent in the low vowel /a/” (Al-Hashmi 2016:43).

1997:232). Davis (1995:477) explains that there are no phonemes that block leftward emphasis spread. This means that in words like *xayya:t* ‘tailor’, where /a:/ immediately precedes an emphatic, there is leftward emphasis spread which can inhibit *imala* (see Habib 2012). In other words, when there is a pharyngealized consonant in the following phonological segment (after /a:/), *imala* is expected to be disfavoured. Al-Bataineh (2019) also states that the spread of the retracted tongue root (or the [+ RTR]) feature from [r] is similar to that from other pharyngealized consonants, which means that [r] generally has the same effect as emphatics (i.e., it affects all vowels except high front vowels).

Emphasis can also spread rightward to an adjacent low vowel (Davis 1995:485; see also Kulikov et al. 2020:16–17), such as the low vowel /a:/⁸³ (Altairi et al. 2017:10; Davis 1995:485–86; Habib 2012:67; Shahin 2011:9), e.g., *ʔa:za* ‘fresh’ (Davis 1995:485). Davis (1995:479) gives some examples of word-final /a/ in feminine nouns and adjectives, such as the words *salxa* ‘slap’, *fari:da* ‘obligation’ and *fallaḥa* ‘peasant (female)’, where the low vowel, or /a/, is pharyngealized if there is an emphatic in the root. However, in the absence of a preceding pharyngealized consonant, the allomorph is [i] (in the rural southern variety of PA that Davis (1995) examines), e.g., *samaki* ‘fish’, *farši* ‘mattress’ (Davis 1995:479; see also Herzallah 1990). This suggests that in word-medial *imala*, when there is a preceding pharyngealized consonant (before /a:/), the raising and fronting of /a:/ to [e:], or *imala*, is expected to be disfavoured.

Based on the discussion above, in the coding protocol, pharyngealized consonants are

⁸³ Davis (1995:476) argues that at least as far as the southern rural variety of PA he examined is concerned, “[i]t is only the nonback [+ high] sounds that are opaque to [or that block] rightward [emphasis] spread.” This means that when there is a pharyngealized consonant like an emphatic before [a:], the surfacing vowel should not be nonback and/or [+ high], so the fronting and raising of [a:] to [e:] cannot occur because of rightward emphasis spread (see Habib 2012). In other words, [a:], which is [+ low] and [+ back] (Habib 2012:67), is not opaque to rightward emphasis spread, i.e., it does not block rightward emphasis spread when a pharyngealized consonant like an emphatic occurs before [a:]. As such, when there is a pharyngealized/emphatic consonant before an [a:], then rightward emphasis spreads because [a:] does not block it, and the surfacing vowel cannot be the fronted and raised [e:] because of rightward emphasis spread (see Habib 2012). As Davis (1995:477) explains, “if [RTR] spreads onto a phoneme, that phoneme cannot be both [+ high] and [- back]” as in the fronted and raised [e:] variant.

distinguished from other consonants, like labials [m b f], palatals [dʒ ʃ j], velars⁸⁴ [k w], and dental-alveolars, e.g., [n t d s z]. Pharyngealized consonants are hypothesized to disfavour *imala* (or the [e:] variant) in word-medial position in the preceding and in the following phonological segments, while the remaining consonantal groups may favour *imala*.

It is worth noting that in addition to the effect of the preceding and following phonological segments on variant selection in *imala*, there are other linguistic factor groups which have been examined in the literature. Most notably, Habib (2012:65) examined the lexical conditioning of *imala*, arguing that words that are borrowed from Standard Arabic do not undergo *imala* in word-medial position, such as *ħa:dis* ‘accident’ (*ħa:diθ* in Standard Arabic). In the data for this study, a similar effect was observed, where some words borrowed from Standard Arabic categorically disfavoured *imala*, such as *ma:l* ‘money’ (vs. *maʃa:re* in PA and LA), *da:ʔiman* ‘always’ (vs. *da:yman* or *de:yman* in PA and LA), and *ʃada:ʔ* ‘hostility’ (vs. *ʃada:we* or *ʃade:we* in PA and LA). However, these tokens were infrequent in the data, so the factor group addressing lexical constraints was excluded from the analysis.

4.3 Results

4.3.1 Overall distribution of variants

Table 4.2 shows the distribution of the [e:] and the [a:] variants in the diachronic and synchronic datasets.

⁸⁴ It should be noted that in PA, the velar /k/ is said to be an underlying uvular if it is a variant of /q/, e.g., *waraqa* > *waraka* ‘paper’ (Abo Mokh and Davis 2020:44). In the present study, all instances of /k/ that were coded as ‘velar’ were not underlying uvulars (i.e., they were not variants of /q/, but actual velars, e.g., *ka:n* / *ke:n* ‘he/it was’; *maka:tib* ‘offices’; *maka:n* ‘place’).

Table 4.2: Overall distribution of the variants of word-medial *imala* in the diachronic and synchronic datasets

	[e:]		[a:]	
	%	N	%	N
PA synchronic dataset	15%	344/2285	85%	1941/2285
LA synchronic dataset	63%	891/1404	37%	513/1404
POHA	1%	66/4440	99%	4374/4440
LPTC	64%	2390/3741	36%	1351/3741

From the table above, we can see that [e:] in word-medial position seems to be the default variant in the LA synchronic dataset and in the LPTC, where variant rates are almost identical. This confirms previous claims in the literature that [e:] in word-medial position is the predominant variant in LA (Barkat 2008:678; Borg 2005:538; Hennessey 2011:59; Levin 2006:314; Naïm 2011:925). Additionally, from variant rates alone, we can see that the LPTC seems to approximate to the contemporary LA variety. However, the validity of the LPTC as an approximation to LA is assessed mainly through variant conditioning because variant rates can fluctuate for a variety of extra-linguistic reasons. As such, variant conditioning or the underlying grammar (including the hierarchy of constraints and the direction of effects)⁸⁵ is our key comparative measure because it is expected to be more constant and it reflects more stable constraints on variant selection (Poplack and Levey 2010:400; Poplack and Tagliamonte 2001:93–94). If patterns of the linguistic conditioning of variant choice are parallel across the LPTC and the LA synchronic dataset, we can confirm that the former is an approximation to LA with regards to word-medial *imala*.

Variant rates differ across the PA synchronic dataset and the POHA. In the latter, word-medial *imala* rarely occurs, indicating that non-raising in word-medial position is the default

⁸⁵ As explained in chapter 3, the constraint ranking of factors refers to the pattern of favouring and disfavouring effects on variant choice, and the direction of effects refers to the order from more to less of factors within a factor group (Tagliamonte 2006:235, 2012:122).

option in the POHA, which is considered as an approximation to pre-contact PA (see Chapter 3). However, *imala* is comparatively more frequent in the PA synchronic dataset, though it is not the default option, despite the fact that two generations in the contemporary dataset are Lebanese-born (i.e., the second- and the third-generation Palestinians). It remains to be determined who is leading the use of *imala* among Palestinian speakers, and whether this change is contact-induced due to the influence of LA. I explore this below, starting with distributional analyses of social factors, focusing initially on the diachronic datasets.

4.3.2 Distributional analysis of social factors

Diachronic data

In the POHA, we saw in Table 4.2 Table 4.2: Overall distribution of the variants of word-medial *imala* in the diachronic and synchronic datasets that *imala* in word-medial position occurs at a negligible rate of 1% (N=66/4440). Among the 80 speakers in the POHA sampled for this study, only 14 speakers used word-medial *imala*. Table 4.3 shows the individual rates of word-medial *imala* across these 14 speakers and their socio-demographic characteristics.

Table 4.3: Individual rates of word-medial *imala* among 14 speakers in the POHA

Speaker #	% Raised [e:]	Sex	Level of Education	Place of Residence (outside vs. inside camps)
1	1% (N=1/100)	M	Secondary	Outside
4	2% (N=1/55)	M	Primary	Outside
8	2% (N=1/51)	F	No school	Outside
10	2% (N=1/54)	M	Illiterate	Inside
17	49% (N=25/51)	F	No school	Outside
28	1% (N=1/79)	M	Primary	Outside
34	11% (N=6/55)	M	Illiterate	Outside
38	8% (N=5/61)	F	No school	Inside
41	2% (N=1/55)	M	Primary	Inside
43	2% (N=1/51)	M	Primary	Inside
45	2% (N=1/51)	M	No school	Inside
52	2% (N=1/52)	M	Primary	Inside
74	35% (N=18/52)	F	No school	Outside
76	6% (N=3/53)	M	Primary	Inside

Table 4.3 shows that two speakers use word-medial *imala* the most: speaker [17] at 49% and speaker [74] at 35%. The other speakers rarely use *imala*. However, it was further observed that in the speech of these remaining 12 speakers (excluding speakers [17] and [74]), *imala* was mainly limited to the same lexical items that occurred repeatedly, e.g.: *se:fa* ‘hour/watch’ or *se:fa:t* ‘hours/watches’; *me:khdi:n* ‘we/they have taken’ or *me:khith* ‘he has taken’; *we:hid* ‘one/someone, masculine’ or *we:hde* ‘one/someone, feminine’; and *me:t* ‘he died’, *me:tit* ‘she died’ or *me:to* ‘they died’. As such, in the case of these 12 speakers, word-medial *imala* seems to be lexically conditioned because it is not a productive phenomenon outside the restricted number of lexical items that favour its use.

Both speakers [17] and [74] are females who live outside the camps, and they did not attend school. What might explain the speech behaviour of these two speakers is their places of origin in pre-war Palestine. Speaker [17] is originally from a village called Qadas and speaker [74] is from a village called Saliha. I refer to the geographical location and the history of these two villages in explaining the speech behaviour of the two speakers. The map below shows the geographical location of Qadas and Saliha, which are adjacent to the Lebanese border.

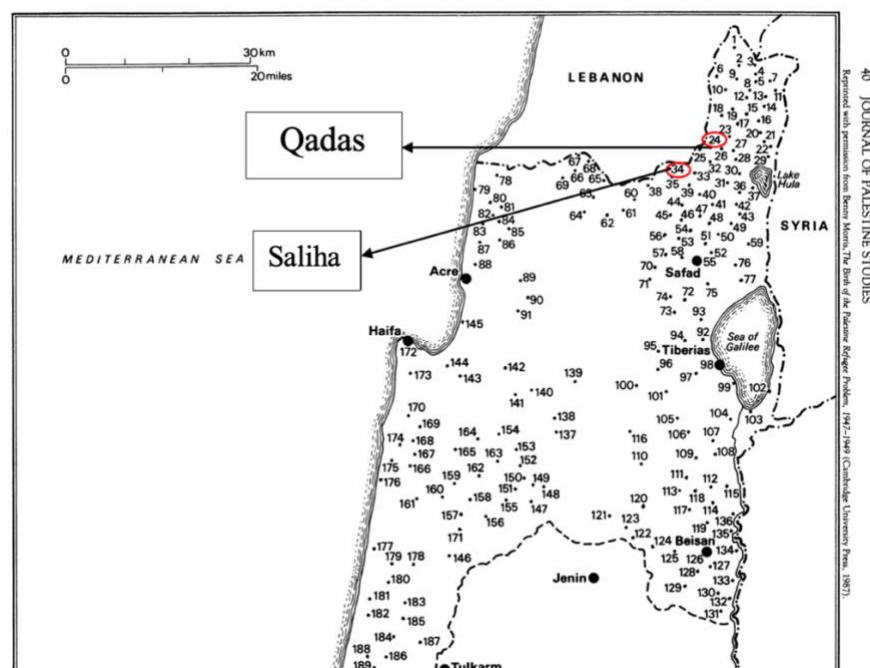


Figure 4.1: Arab settlements abandoned in 1948-1949 (Institute for Palestine Studies 1988; see also Morris 2003:xi)

Qadas and Saliha were among the only seven Shia villages in Palestine,⁸⁶ whose inhabitants were mainly Shia Muslims. Under the French mandate for Lebanon, these villages were originally part of Greater Lebanon in 1920,⁸⁷ and in 1921, France granted Lebanese citizenship to the residents of these villages (Kaufman 2006:689). In 1923, a border agreement was signed between France and Britain, which was granted a mandate for Palestine in 1920, and the seven Shia villages became part of Palestine. During the 1948 Arab-Israeli war, the majority of the inhabitants of these seven villages fled to Lebanon, and in 1994, they attained Lebanese citizenship. Because speakers [17] and [74] were born in an area that was formerly part of Lebanon, it is highly likely that they would have been exposed to prevailing Lebanese speech

⁸⁶ The seven villages are: Hunin, Qadas, Nabi Yusha', Saliha, Tarbikha, Al-Malikiyyah, Abl al-Qimh (Kaufman 2006:689).

⁸⁷ The establishment of Greater Lebanon was declared in 1920. Before that, present-day Lebanon was part of Bilad Al-Sham 'Land to the North' or Greater Syria, which also included Palestine, and present-day Jordan and Syria.

norms, which would explain their higher usage of word-medial *imala* compared to other Palestinian speakers in the POHA.

Going back to the overarching aim of this study, which is to examine whether PA is converging on LA, in terms of word-medial *imala* we can see from variant rates considered so far that in the speech of Palestinians in the POHA, which is assumed to approximate pre-contact PA, the [e:] variant in word-medial position is infinitesimally rare. As we saw above, the bulk of *imala* tokens in this dataset are attributable to two female speakers only, who are originally from villages that were formerly part of a Lebanese dialect area. Even the four speakers in the POHA who are originally from Jerusalem do not make use of word-medial *imala* (N=0/247), so their variety of Jerusalem Arabic does not seem to be characterized by this variant (contra Al-Wer 2007b:70).⁸⁸ In contrast, we see a vertiginous rise in the use of word-medial *imala* mainly in the third-generation in the PA synchronic dataset (see Figure 4.5 and Figure 4.7 below). Therefore, a first provisional conclusion we can draw from this analysis is that there is change in PA.

Turning to the second diachronic source, the LPTC, we see different patterns for word-medial *imala* than in the POHA. We saw in Table 4.2 that in the LPTC, word-medial *imala* occurs at a rate of 64% (N=2390/3741). Figure 4.2 shows the distribution of the raised [e:] variant in word-medial position by age and sex of the characters in the LPTC.

⁸⁸ Without systematically examining other varieties of PA outside the contact scenario confronted in the present study, I cannot be sure whether or not word-medial *imala* is a feature of other varieties of PA or of Jerusalem Arabic in particular.

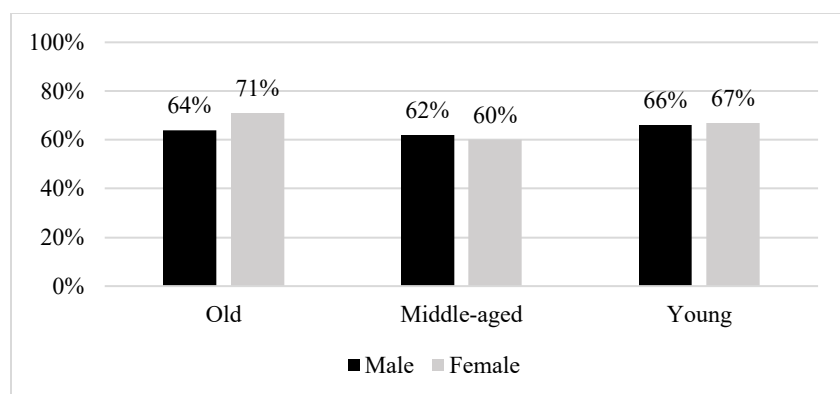


Figure 4.2: Distribution of the [e:] variant in word-medial position by age and sex of the characters in the LPTC

From the figure above, we can see that older, middle-aged and young characters use word-medial *imala* at similar rates. There is also little evidence of sex-differentiation in the LPTC in terms of word-medial *imala*. Again, this shows that the [e:] variant in word-medial position seems to be the default variant in LA.

Figure 4.3 shows the distribution of the raised [e:] variant in word-medial position by social status of the characters in the LPTC.

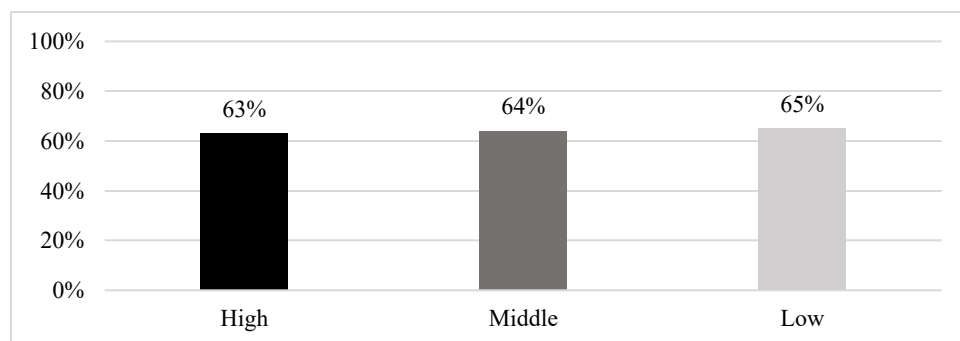


Figure 4.3: Distribution of the [e:] variant in word-medial position by social status of the characters in the LPTC

The figure above shows that the rates of word-medial *imala* are almost identical across characters of higher, middle and lower social status. This result adds to the burgeoning evidence that word-medial *imala* does not seem to be socially diagnostic in the LPTC. In the next section,

I examine whether this is also the case in the contemporary LA variety, and I also examine the contribution of social factors to variant choice in contemporary PA.

Synchronic data

Figure 4.4 and Figure 4.5 show the distribution of the raised [e:] variant in word-medial position by age (or generation for Palestinians) and speaker sex in the LA and PA synchronic datasets.

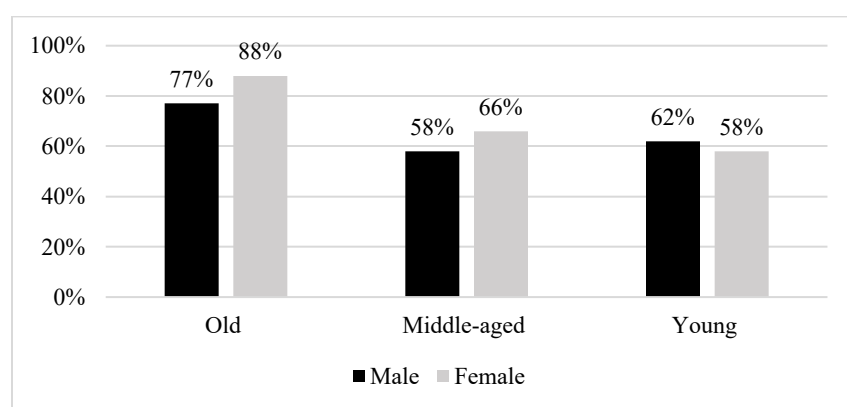


Figure 4.4: Distribution of the [e:] variant in word-medial position by age and sex in the Lebanese synchronic dataset

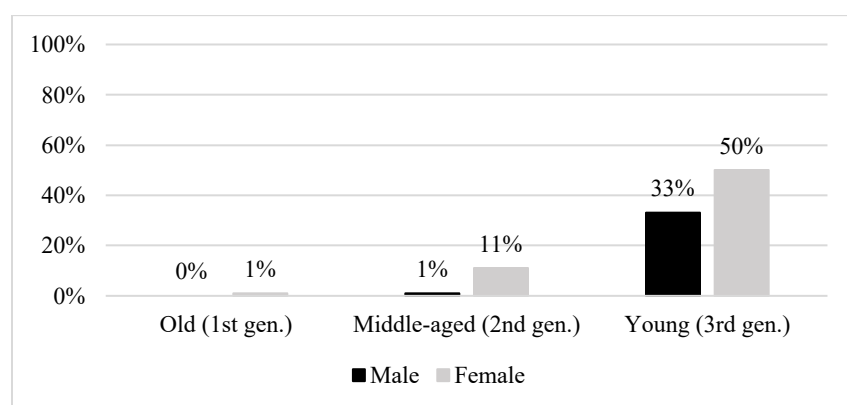


Figure 4.5: Distribution of the [e:] variant in word-medial position by age/generation and sex in the Palestinian synchronic dataset

In the LA synchronic dataset, we see in Figure 4.4 that *imala* rates are generally similar across the three age groups, though older Lebanese speakers use *imala* at slightly higher rates

than their younger counterparts. It should be noted that with the exception of one older Lebanese male from Beirut, the remaining older Lebanese speakers are from Mount Lebanon, from a town called Cornet Chahwan, which is around 20 km away from Beirut (see Chapter 5). Because the majority of the older Lebanese speakers are originally from an area outside Beirut, I cannot be sure that their usage of word-medial *imala* faithfully reflects metropolitan LA norms (characteristic of Beirut Arabic).

In terms of speaker sex, we observe in the LA synchronic dataset that males and females use *imala* at similar rates, so there is little evidence of sex-differentiation in speech. Accordingly, similar to what we saw in the LPTC, *imala* in contemporary LA does not seem to be socially diagnostic, and as far as variant rates and distributional analyses of social factors are concerned, *imala* seems to be stable in this variety.

In contrast to the distributional results in contemporary LA, in the Palestinian synchronic dataset, we see that word-medial *imala* is virtually absent in the speech of older Palestinians, where it only rarely occurs in the speech of older females (N=2/237). This is in line with the hypothesis that older speakers retain Palestinian vernacular norms that they acquired as children. Additionally, it again confirms that *imala* in word-medial position does not seem to be the default variant in PA as spoken in Beirut (Hennessey 2011:59; Naïm 2005:277, 2011:925; Versteegh 2009:482–83). The rate of *imala* slightly increases among the second-generation, most notably in the speech of middle-aged Palestinian females (N=57/533), while it rarely occurs in the speech of middle-aged males (N=3/532). Figure 4.5 shows that it is the younger Palestinians (N=282/654), especially females (N=197/396), who use *imala* much more frequently than their older counterparts. This is consistent with the earlier hypothesis that young speakers and females are more likely to use the supralocal feature associated with the prestige variety of Beirut Arabic

(Hennessey 2011:35, 92; see Versteegh 2001:153).

Figure 4.6 and Figure 4.7 show the effect of level of education and age on variant selection in the LA and PA synchronic datasets. These two factor groups have been cross-tabulated because they interact with each other, as previously mentioned.

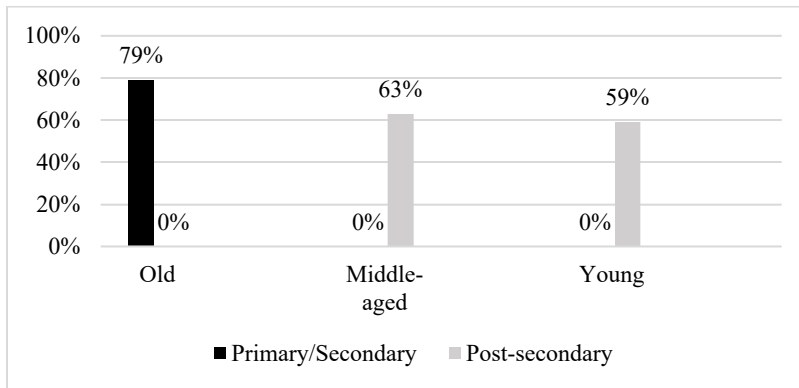


Figure 4.6: Distribution of the [e:] variant in word-medial position according to age and level of education in the Lebanese synchronic dataset

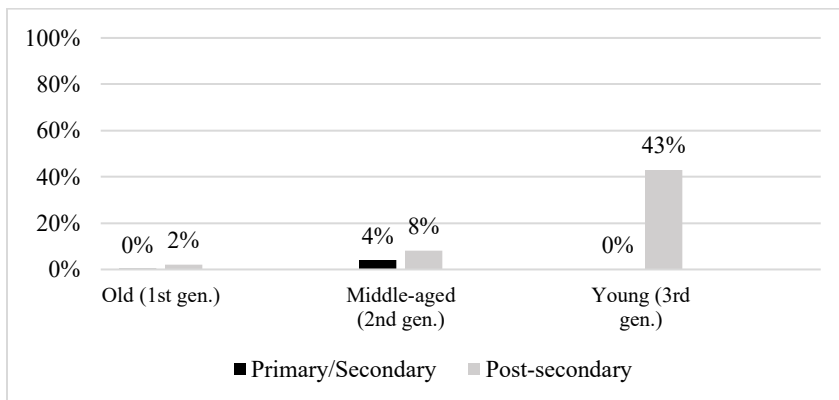


Figure 4.7: Distribution of the [e:] variant in word-medial position according to age and level of education in the Palestinian synchronic dataset

As observed earlier, in the Lebanese synchronic dataset, *imala* rates are generally similar across age groups/educational levels. However, in the Palestinian synchronic dataset, there are huge disparities in *imala* rates across age groups/educational levels. In the speech of older Palestinians, word-medial *imala* is used only once by those educated to primary or secondary level (N=1/506 or 0.19%), where it is traced back to one female (N=1/61), and it occurs only

once in the speech of the only older female educated to post-secondary level (N=1/60). Word-medial *imala* increases substantially in the speech of young Palestinians educated to post-secondary level. This is presumably because they are regularly exposed to LA speech features, especially in post-secondary institutions.

Figure 4.8 shows the effect of place of residence, which is cross-tabulated with age, on the use of word-medial *imala* among Palestinian speakers.

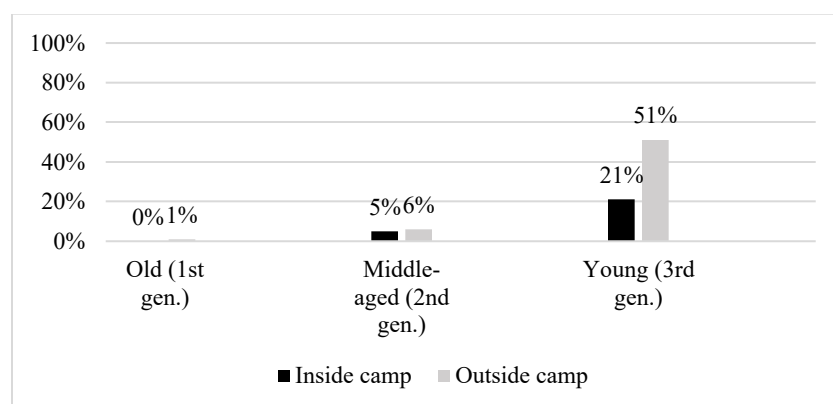


Figure 4.8: Distribution of the [e:] variant in word-medial position according to age and place of residence in the Palestinian synchronic dataset

Figure 4.8 shows that word-medial *imala* is especially predominant in the speech of young Palestinians outside the camps, but *imala* is still used by young (and educated) speakers inside the camps to some extent (N=35/169). As mentioned previously, this is because all of them have been presumably exposed to LA speech norms in post-secondary institutions outside the camps. However, the rate of *imala* use by young Palestinians inside the camps is still considerably lower than that in the speech of young Palestinians outside the camps. This is consistent with the hypothesis that speakers who live inside the camps may adhere more to Palestinian vernacular norms because of the dense social networks in the camps, which help maintain Palestinian speech patterns. On the other hand, speakers who live outside the camps are expected to use more word-medial *imala* presumably because of stronger ties or more frequent interactions with

Lebanese speakers. This hypothesis, though, is not entirely supported by the findings for the middle-aged speakers. For Palestinians in this age group, the rate of *imala* is similar across camp-dwellers (N=19/382 or 5%) and non-camp-dwellers (N=41/683 or 6%). Further inspection of the data reveals the effects of “outliers” or individual speakers whose linguistic behaviour falls outside that of the wider speech community (see Britain 2003:191).

To examine who the outliers are among Palestinian middle-aged speakers inside and outside camps, I narrow the analysis down to consider speakers in that age cohort. The table below shows the rates of word-medial *imala* across middle-aged Palestinian speakers recorded in this study, along with their socio-demographic characteristics.

Table 4.4: Individual rates of word-medial *imala* among Palestinian second-generation (middle-aged) speakers and their socio-demographic characteristics

Speaker #	% Raised [e:]	Sex	Level of Education	Place of Residence (outside vs. inside camps)
21	0% (N=0/52)	F	Post-secondary	Outside
22	0% (N=0/52)	M	Primary	Outside
22	0% (N=0/73)	M	Secondary	Outside
27	0% (N=0/51)	F	Post-secondary	Outside
30	0% (N=0/77)	M	Secondary	Outside
34	2% (N=1/60)	M	Secondary	Inside
36	3% (N=2/60)	M	Post-secondary	Outside
37	10% (N=6/60)	F	Secondary	Outside
38	0% (N=0/60)	M	Post-secondary	Outside
46	5% (N=2/41)	F	Post-secondary	Outside
47	7% (N=3/42)	F	Post-secondary	Outside
49	42% (N=27/64)	F	Post-secondary	Outside
53	31% (N=18/58)	F	Secondary	Inside
55	0% (N=0/63)	F	Secondary	Inside
56	0% (N=0/54)	M	Secondary	Inside
58	0% (N=0/59)	M	Secondary	Inside
59	0% (N=0/51)	F	Secondary	Inside
60	2% (N=1/51)	F	Post-secondary	Outside
64	0% (N=0/37)	M	Post-secondary	Inside

There are 19 Palestinian middle-aged speakers, 7 inside the camps (4 males and 3 females), and 12 outside the camps (5 males and 7 females). In the table above, we see that only 8 middle-aged speakers use word-medial *imala*, but 6 of these speakers only use it at negligible rates. The higher rates of *imala* are traced back to two females only (highlighted in grey in the table), one outside the camps and one inside the camps. I appeal to the individual characteristics and personal circumstances of these two outliers to explain their idiosyncratic speech behaviour.

Speaker [49] uses *imala* at a rate of 42% (N=27/64). She is educated to post-secondary level, and she works as a teacher in a private Lebanese school. As such, this speaker may be more inclined towards using word-medial *imala* characteristic of LA as spoken in Beirut (Hennessey 2011) due to the considerations of the linguistic marketplace (Bourdieu 1991).⁸⁹ As Wolfram and Schilling (2016:163) explain, the linguistic marketplace refers to the extent to which a person's economic activity necessitates using a particular language variety. In addition to her employment conditions, speaker [49] lives outside the camps in Beirut with her Lebanese husband. These characteristics may explain the higher rate of word-medial *imala* that she displays compared to speakers in her age group with some similar characteristics (i.e., non-camp dwellers, educated to post-secondary educational level).

The only other middle-aged Palestinian female who uses *imala* to any appreciable extent is speaker [53]. The fact that she lives in the camp and is educated to secondary level does not explain why she uses word-medial *imala* at a relatively high rate. However, for this speaker, her employment conditions may explain why she behaves differently from her cohort. This speaker

⁸⁹ Bourdieu (1991:17) explains that “[l]inguistic utterances or expressions are forms of practice and, as such, can be understood as the product of the relation between a linguistic habitus and a linguistic market. The linguistic habitus is a sub-set of the dispositions which comprise the habitus: it is that sub-set of dispositions acquired in the course of learning to speak in particular contexts (the family, the peer group, the school, etc.). These dispositions govern both the subsequent linguistic practices of an agent and the anticipation of the value that linguistic products will receive in other fields or markets - in the labour market, for example, or in the institutions of secondary or tertiary education.”

explained that she works as a cook in the house of a prominent Lebanese politician, and before that, she used to work in a company in the south of Lebanon, outside the camps. Her employment conditions outside the camps presuppose regular exposure to LA speakers, which may explain why she adopts Lebanese speech patterns by using word-medial *imala*.

The case of speaker [53] shows that the presumed social networks of speakers (in this case, the social networks in the camp) are not always deterministic of their speech behaviour. This ties in with Villena-Ponsoda's (2005:331) recognition that even within close-knit social networks, personal or attitudinal factors, as well as individual conditions of social life may produce linguistic differences. Such conditions of social life may include practical aspects of individual life (such as job or market conditions), which may affect a speaker's linguistic behaviour, as she may accommodate to other speakers' varieties "in line with more general psychosocial accommodation" (Villena-Ponsoda 2005:331).

In summary, the findings for the distributional analyses of the social factors among Palestinian and Lebanese speakers show that for the latter, *imala* in word-medial position is the predominant variant, and there is little evidence that social factors among Lebanese speakers contribute to variant selection. The case differs for Palestinian speakers, where word-medial *imala* is mainly used by the younger, educated generation, especially females. The fact that it is almost absent from the speech of older Palestinians, and it is also rare in the speech of the majority of middle-aged speakers, suggests that *imala* in young Palestinian speech is not likely to qualify as a case of transmission and incrementation of an internally-motivated development from the second- to the third-generation Palestinians.⁹⁰ In other words, it is not likely that the third-generation Palestinians acquired *imala* through transmission from their parents (the middle-

⁹⁰ Labov (2007:346) explains that *transmission* refers to the "unbroken sequence of native-language acquisition by children", and *incrementation* is when "successive cohorts and generations of children advance the change beyond the level of their caretakers and role models, and in the same direction over many generations."

aged speakers), and went on to advance this change. Instead, *imala* in the speech of young Palestinians may be an externally-motivated change due to contact with LA. I explore this possibility later in this chapter by conducting systematic comparisons of the linguistic conditioning of variant selection across the PA and LA synchronic datasets.

4.3.3 Multivariate Analysis

Comparison between the LPTC and the LA synchronic dataset

So far, we have seen that the raised [e:] variant in word-medial position seems to be the default choice in LA and in the LPTC, and it is used to some extent in the PA synchronic dataset, while in the POHA, it is virtually absent. For this reason, I exclude the POHA from Multivariate Analysis. I begin this section by comparing the patterns of underlying conditioning of word-medial *imala* in the LPTC and the LA synchronic dataset. The purposes behind this real-time comparison are: (i) to determine the factors that constrain variant choice in the LPTC, thus answering one of the research questions posed in Chapter 1;⁹¹ (ii) to confirm, following observations from variant rates and distributional analyses of social factors, that word-medial *imala* is stable in LA (this is confirmed if there are parallels in the patterns of conditioning of variant selection across the LPTC and the LA synchronic dataset); and (iii) to examine the extent to which the LPTC approximates to LA in terms of patterns of variant conditioning (see Chapter 3).

⁹¹ The research question that concerns the diachronic datasets and real-time comparisons is as follows: What are the linguistic and social constraints that determine the variable expression of the three targeted variables in antecedent states of PA and LA, and how can earlier stages of the varieties be used to illuminate patterns of variation in the contemporary varieties?

The data from the LA synchronic dataset and the LPTC were subjected to Rbrul (Johnson 2009) to examine the social and linguistic predictors of variant choice.⁹² The numbers in Table 4.5 below (and the tables that follow) are to be interpreted as follows: The input probability refers to the predicted probability of the response (dependent linguistic variable), averaged over all factor combinations (or cells), i.e., independent variables. In other words, the input probability is the average of the predicted values for each cell (Johnson 2010; Rbrul manual).⁹³ Total N refers to the total number of tokens extracted in each dataset. The deviance “is a measure of how well the model fits the data, or how much the actual data deviate from the predictions of the model. The larger the deviance, the worse the fit. When the model has many factors, this number increases” (Tagliamonte 2012:143).⁹⁴ The key numbers in the table are the (centred) factor weights⁹⁵ and/or the log-odd values. Factor weights are numbers between 0 and 1, and they indicate the probability that word-medial raised [e:] will be selected in the contexts listed in the left-hand side of the table. A factor weight above 0.5 favours word-medial raised [e:], while a factor weight below 0.5 disfavors this variant (see Johnson 2009:360–61). A factor weight of 0.5 expresses neutrality (i.e., the factor neither favours nor disfavors the raised [e:] variant). Log-odds can be any positive or negative number (Johnson 2009:361), and they convey the same information as factor weights.⁹⁶ A positive log-odd value favours word-medial raised [e:], while

⁹² Distributional analyses of linguistic factors contributing to variant choice in all datasets, excluding the POHA, can be found in Appendix B.

⁹³ In the Rbrul manual, Johnson (2010) explains that the input probability “reports the overall prediction of the model [...] All these models make a prediction for the mean (in linear regression) or the proportion (in logistic regression) in each cell, where a cell is defined as a given setting of the independent variable(s). The input probability is the average of the predicted values for each cell.”

⁹⁴ Johnson (2009:381) explains that the deviance reported by Rbrul “is defined as -2 times the log-likelihood, reported by GoldVarb. In both cases, values closer to zero represent better-fitting models.”

⁹⁵ Johnson (2010) explains that if centred weights were chosen in Rbrul settings, “then the correspondence would be exact. The default is uncentered weights [...] to match Goldvarb’s default behavior. In a group with two factors, if factor weights are not centered, the factor with more tokens comes out closer to 0.5, while the one with fewer tokens is shifted towards more extreme values.”

⁹⁶ Johnson (2009:361) explains that “[w]e obtain log-odds from probabilities by taking the natural (base e) logarithm of the odds, where the odds are the probability of an event occurring, divided by the probability of it not occurring.”

a negative log-odd value disfavors this variant. A log-odd value of 0 is equivalent to a factor weight of 0.5 (it indicates neutrality) (Johnson 2009:361). The numbers under the N column refer to the total number of tokens extracted for each of the factors in the left-hand side of the table, and the proportion of the application value refers to the proportion of the [e:] variant used in the contexts listed in the left-hand side of the table.

Table 4.5: Contribution of social and linguistic factors to the selection of word-medial raised [e:] in the LPTC and the Lebanese synchronic dataset (all age groups)

(Notes: shaded areas indicate statistically significant factor groups)

LPTC					Lebanese synchronic data (all age groups)				
Input prob.	0.824				Input prob.	0.869			
Total N	3741				Total N	1404			
Deviance	3606.15				Deviance	818.787			
	Log-odds	N	Proportion App. Value	FW centered		Log-odds	N	Proportion App. Value	FW centered
Sex					Sex				
Female	.004	950	.654	.501	Female	.006	847	.614	.502
Male	-.004	2791	.634	.499	Male	-.006	557	.666	.498
Age					Age/Education				
Young	.171	1381	.666	.543	Older / non-PS	1.105	261	.793	.751
Older	.144	503	.654	.536	Middle-aged / PS	-.414	258	.628	.398
Middle-aged	-.315	1857	.614	.422	Young / PS	-.691	885	.590	.334
Social status									
Middle	.115	1254	.640	.529					
Low	.020	981	.654	.505					
High	-.135	1506	.628	.466					
Preceding Phonological Context					Preceding Phonological Context				
Palatal	.679	483	.828	.664	Palatal	1.725	156	.917	.849
Velar	.601	384	.781	.646	Dental-Alveolar	1.707	392	.906	.846
Dental-Alveolar	.538	1041	.802	.631	Labial	.445	195	.810	.609
Labial	.281	554	.771	.570	Velar	.078	224	.817	.519
Pharyngeals	-2.099	1279	.335	.109	Pharyngeals	-3.955	437	.119	.019
Following Phonological Context					Following Phonological Context				
Velar	1.376	170	.865	.798	Velar	1.082	24	.750	.747
Dental-Alveolar	.345	1634	.735	.585	Dental-Alveolar	.685	721	.753	.665
Labial	.182	457	.676	.545	Labial	.408	172	.645	.601
Palatal	-.469	446	.590	.385	Palatal	-.219	115	.574	.445
Pharyngeals	-1.434	1034	.455	.192	Pharyngeals	-1.956	372	.411	.124

In terms of social factors, in both the LPTC and the LA synchronic dataset, only age (collapsed with education in latter) is selected as significant. In the LPTC, *imala* is weakly favoured by younger (FW=.543) and older (FW=.536) characters, and weakly disfavoured by middle-aged characters (FW=.422). These results don't point in favour of change in progress as the oldest and youngest characters pattern similarly. In LA, it is the older speakers who use *imala* the most (FW=.751), but in comparison, the [e:] variant is disfavoured by middle-aged (FW=.398) and younger speakers (FW=.334). It was observed in the distributional analysis of social factors that older Lebanese speakers use word-medial *imala* the most, but as mentioned previously, because they are from an area outside Beirut (except for one older man from Beirut), I cannot be sure that their usage patterns (in terms of word-medial *imala*) faithfully reflect contemporary metropolitan norms. As observed in the distributional analyses, there is little evidence of sex-differentiation in the use of word-medial *imala* across the LPTC and the LA synchronic dataset, and in former, the raised [e:] variant is used at similar rates across higher, middle and lower status characters.

In terms of the preceding phonological segment, there are minor differences between the LPTC and the LA synchronic dataset. In the LPTC, *imala* is favoured by velars (FW=.646), followed by dental-alveolars (FW=.631), while it is weakly favoured by labials (FW=.570). In the LA synchronic dataset, *imala* is favoured by dental-alveolars (FW=.846), followed by labials (FW=.609), and it is weakly favoured by velars (FW=.519). Despite these minor differences, we can see broad parallels in the effect of the preceding phonological segment on variant choice, such that in both comparison groups, palatals strongly favour word-medial *imala* (FW=.664 in the LPTC and FW=.849 in the LA synchronic dataset), while pharyngeals strongly disfavour the [e:] variant (FW=.109 in the LPTC and FW=.019 in the LA synchronic dataset).

In terms of the following phonological context, we observe the same constraint hierarchy and direction of effects in the LPTC and in the LA synchronic dataset: Velars strongly favour *imala*, followed by dental-alveolars and labials. Palatals have a disfavoured effect in both datasets (FW=.385 in the LPTC and FW=.445 in the LA synchronic dataset), and pharyngeals strongly disfavour *imala* (FW=.192 in the LPTC and FW=.124 in the LA synchronic dataset).

From these findings, we can conclude that not only do we find almost identical variant rates across the LPTC and the LA synchronic dataset, but there are also structural similarities in the linguistic conditioning of word-medial *imala*. What this confirms is that *imala* is the default choice in LA, and it is relatively stable in this variety. It further suggests that the LPTC, despite being based on a fictional source, approximates to spoken LA, as gauged from broad parallels across the LPTC and the LA synchronic dataset in terms of the fine-grained linguistic conditioning of word-medial *imala*.

Comparison between the LA and PA synchronic datasets

To examine whether Palestinians are reproducing the same linguistic constraints on variant choice as those operating in LA (i.e., whether PA is converging on LA in terms of word-medial *imala*), I compare the constraint ranking of linguistic factors and the direction of their effects across LA and PA. Because word-medial *imala* is only productive in the speech of the third-generation Palestinians, I limit the analysis in this section to young Palestinians and compare their speech patterns with those of young Lebanese speakers. Table 4.6 shows the socio-demographic characteristics of young Palestinian and Lebanese speakers.

Table 4.6: Characteristics of the younger Palestinian and Lebanese sample populations

Younger Palestinians		Younger Lebanese	
Sex	N	Sex	N
Male	5	Male	5
Female	6	Female	12
Place of residence		(Camp residence is not applicable for Lebanese speakers)	
Inside camp	3		
Outside camp	8		
Total	11	Total	17

Table 4.7 shows the overall distribution of the [e:] and [a:] variants in young Palestinian and Lebanese speech.

Table 4.7: Overall distribution of the variants of word-medial *imala* in the PA and LA synchronic datasets (younger speakers)

	[e:]		[a:]	
	%	N	%	N
PA synchronic dataset (Younger speakers)	43%	282/654	57%	372/654
LA synchronic dataset (Younger speakers)	59%	522/885	41%	363/885

The synchronic datasets (young Lebanese and Palestinian speech) were subjected to Rbrul (Johnson 2009). Table 4.8 shows the results of multivariate analysis of social and linguistic predictors of variant choice across the two datasets.

Table 4.8: Contribution of social and linguistic factors to the selection of word-medial raised [e:] in younger Lebanese and third-generation Palestinians

(Notes: shaded areas indicate statistically significant factor groups; K.O.= invariant context)

Younger Lebanese					Younger (third-generation) Palestinians				
Input prob.	0.830				Input prob.	0.284			
Total N	885				Total N	654			
Deviance	504.300				Deviance	425.938			
	Log-odds	N	Proportion App. Value	FW centered		Log-odds	N	Proportion App. Value	FW centered
Sex					Sex / Camp Residence				
Male	.077	258	.624	.519	Female / Outside camp	1.146	331	.595	.759
Female	-.077	627	.576	.481	Male / Outside camp	-.320	154	.325	.421
					Male / Inside camp	-.826	104	.337	.304
					Female / Inside camp	K.O.	65	.000	K.O.
Preceding Phonological Context					Preceding Phonological Context				
Palatal	2.355	86	.953	.913	Dental-Alveolar	2.001	162	.667	.881
Dental-Alveolar	1.576	248	.883	.829	Palatal	1.202	80	.625	.769
Labial	.337	122	.779	.583	Labial	.370	97	.680	.591
Velar	.112	137	.796	.528	Velar	-.333	88	.420	.418
Pharyngeals	-4.380	292	.058	.012	Pharyngeals	-3.240	227	.093	.038
Following Phonological Context					Following Phonological Context				
Velar	1.819	12	.667	.860	Palatal	.993	37	.432	.730
Dental-Alveolar	.295	456	.711	.573	Dental-Alveolar	.444	334	.482	.609
Palatal	.104	73	.466	.526	Labial	.164	70	.429	.541
Labial	-.163	109	.578	.459	Velar	-.134	29	.517	.467
Pharyngeals	-2.055	235	.396	.114	Pharyngeals	-1.467	184	.326	.187

In the table above, we can see that multivariate analysis confirms that while there is little evidence of sex-differentiation in terms of the use of word-medial *imala* in the speech of young Lebanese speakers, in young Palestinian speech, it is only females who live outside camps who favour this variant (FW=.759), although this effect is non-significant.

In terms of the preceding phonological segment, the direction of effect for this factor group is generally similar across the two datasets. The only minor difference is that in LA, palatals favour word-medial [e:] the most (FW=.913), followed by dental-alveolars (FW=.829), while the direction of effects is reversed in young PA speech. However, we see that in both datasets, these two consonantal groups have a strong favouring effect on variant selection, followed by labials, then velars. Preceding pharyngeals (gutturals, [r], emphatics and uvulars) in both datasets strongly disfavour word-medial [e:], in line with the earlier hypothesis and previous claims in the literature, especially in relation to emphasis spread (Al-Nassir 1985, 1993; Barkat 2008; Davis 1995; Habib 2012; Levin 2006; Owens 1993, 2006)

In terms of the effect of the following phonological segment on variant selection, there are slight differences between LA and PA. For example, a following velar in young LA speech (FW=.860) strongly favours word-medial *imala*, but in young PA speech, it weakly disfavors this variant (FW=.467). Following labials weakly disfavour variant choice in LA (FW=.459), while they have a weak favouring effect in PA (FW=.541). However, even though the direction of effects differs across both datasets, the pattern of favouring and disfavouring effects on variant choice is generally similar, where in the two comparison groups, following pharyngeals have a disfavouring effect on variant selection (especially [r] and emphatics in young LA and PA speech; see distributional analysis of linguistic factors in Appendix B).⁹⁷ By contrast, other

⁹⁷ Note that only uvulars in the following phonological segment highly favoured word-medial *imala* in young LA speech (N=14/15 or 93%) and in young PA speech (N=19/21 or 90.5%), but because it was collapsed with other

consonantal groups, such as dental-alveolars and palatals, have a generally favouring effect on variant choice. These similarities, coupled with the broad parallels we already saw in terms of the effect of the preceding phonological segment on variant choice in both datasets, show that the patterns of the linguistic conditioning of word-medial *imala* are broadly similar across young LA and PA speech. This result, taken together with salient intergenerational differences in the Palestinian synchronic dataset in terms of the use of the target feature, bolsters the inference that we are dealing with an externally-motivated change in PA as spoken in Beirut (young speech) due to the influence of LA.

4.4 Excursus on word-final *imala*

So far, we have seen that mainly young Palestinians converge on LA speakers in terms of word-medial *imala*. But why did (contact-induced) change occur in this component of the grammar (phonology)? There are several social factors that are favourable to contact-induced change in the context investigated here (the intensity and length of contact between LA and PA; the majority status of LA vs. the minority status of PA; and the prestige attributed to Beirut Arabic). However, there may also be linguistic motivations facilitating the diffusion of word-medial *imala* to (young) Palestinians. The adoption of this feature does not seem to have systemic phonological consequences for young Palestinians because it does not require extensive internal restructuring of the grammar. As highlighted in Chapter 2, Palestinian speakers variably raise the [a] at the end of feminine nouns and adjectives (Al-Wer 2020:561; Naïm 2011:925; Shahin 2007:530; Versteegh 2009:482–83). This word-final *imala* has been observed in the LA and PA synchronic and diachronic datasets, as the examples below show:

pharyngeals, this effect is not shown in the multivariate analysis (see distributional analysis of linguistic factors in Appendix B).

- (7) ʕaty-u:-na rukhʂa, ʕmel-na **naʔe:be.** [40/L/M/83/07:26]
 gave.3p-us license did.3p syndicate
 ‘They gave us a license, we formed a syndicate.’
- (8) ktir **helwe** ʂ-ʂoura. [49/P/F/44/54:28]
 very nice.f the-picture
 ‘The picture is very nice.’
- (9) abu:-y ka:n ʕend-o **maʂbane.**
 father.my was.3sm has.3sm soap.factory
 ‘My father had a soap factory.’ [POHA/02/05.07.2004/F/76/01:18]
- (10) l-ḥaye:t **ʂafbe.**
 the-life difficult.f
 ‘Life is difficult.’ [LPTC/Bi-n-Nesbeh La Bokra...Chou?/1978/184/M/02:27:50]

Variable word-final *imala* was extracted and coded for in the Palestinian synchronic dataset (for all age groups) to examine the conditioning of the [e] variant in word-final position. A total of 1747 tokens of word-final *imala* were extracted. Variant distributions are presented in the table below.

Table 4.9: Overall distribution of the variants of word-final *imala* in the Palestinian synchronic dataset

	[e]		[a]	
	%	N	%	N
PA synchronic dataset (all age groups)	68%	1191/1747	32%	556/1747

From the table above, we can see that word-final *imala* seems to be the default choice in PA as spoken in Beirut. Multivariate analysis was conducted on word-final *imala* in the PA synchronic dataset, where I only considered the effect of the preceding phonological segment on variant selection because for some tokens, the following phonological segment was either a pause or an intervention from another speaker (in the case of group interviews). The results of the multivariate analysis are presented in the table below.

Table 4.10: Contribution of the preceding phonological segment to the selection of word-final raised [e] in the PA synchronic dataset (all age groups)
(Notes: shaded areas indicate statistically significant factor groups; K.O.= invariant context)

PA Synchronic Dataset				
Input prob.	0.937			
Total N	1675 ⁹⁸			
Deviance	462.448			
	Log-odds	N	Proportion App. Value	FW centered
Preceding Phonological Context				
Velar	K.O	72	1.000	K.O.
Palatal	2.802	443	.995	.943
Labial	2.003	211	.991	.881
Dental-alveolar	.612	449	.962	.648
Pharyngeals	-5.417	572	.065	.004

In terms of the preceding phonological segment, several affinities are observed between its effect on word-final *imala* and its effect on word-medial *imala* in young PA speech. Preceding pharyngeals strongly disfavour word-final *imala* (FW=.004) as well as word-medial *imala* (FW=.038). Palatals strongly favour word-final (FW=.943) and word-medial *imala* (FW=.769), followed by labials, which strongly favour word-final *imala* (FW=.881) and weakly favour word-medial *imala* (FW=.591). Dental-alveolars also favour word-final and word-medial *imala*, though they have a stronger favouring effect on the latter (FW=.881). A minor difference between the conditioning of word-final and word-medial *imala* is that velars categorically favour word-final *imala* (mostly in words where /w/ was the consonant preceding *imala*, such as *helwe* ‘pretty/nice’ and *ʔahwe* ‘coffee’), while velars weakly disfavour word-medial *imala* (FW=.418). Setting this minor difference aside, we generally see parallels between the linguistic conditioning of word-final and word-medial *imala* in PA. This could be interpreted to mean that PA speakers in Beirut may be extending the linguistic conditioning of word-final *imala* in PA to word-medial

⁹⁸ The total N is 1675 because velars (a total of 72 tokens) were excluded, as they are invariant (K.O.).

imala. One interpretation of the shift to the raised [e:] variant in word-medial position by young Palestinians in Beirut is that it qualifies as an off-the-shelf change (see Milroy 2007:152, 159) because, as a conspicuously salient variable (see Chapter 1), it is accessible for dialect accommodation (see Milroy 2007:159) and it is superficial from a phonological perspective. This means that even though the use of word-medial *imala* by young Palestinians may have required regular social interaction with speakers of LA, the diffusion and acquisition of this supralocal form by young Palestinians is assumed to have been relatively straightforward in the sense that it is governed by a familiar set of constraints, and it did not necessitate extensive systemic phonological restructuring.

4.5 Discussion

I am now in a position to review the contribution of the findings in this chapter to elucidating the research questions posed in Chapter 1. The research questions concerned the social and linguistic predictors of variant choice in the synchronic and diachronic datasets, how diachronic datasets can illuminate patterns of variation in the contemporary varieties, and evidence of PA converging on LA. Drawing on the framework of comparative variationist sociolinguistics, I examined the social and linguistic conditioning of word-medial *imala* in the LA and PA synchronic datasets, as well as in the LPTC, while the POHA was later excluded from the analysis because the overwhelming majority of older Palestinian speakers in that corpus either did not use the [e:] variant in word-medial position or they rarely used it.

In terms of the social predictors of variant choice, we saw in this chapter that in LA and the LPTC, word-medial *imala* does not seem to be socially diagnostic. On the other hand, in the PA synchronic dataset, it is constrained by social factors, where apparent-time comparisons showed that the bulk of the tokens of word-medial *imala* are produced by young (and educated)

Palestinian speakers, especially (young) females who live outside camps. Among the first- and second-generation Palestinians, what little word-medial *imala* there is can be explained as the product of idiosyncratic or exceptional patterns rather than robust community trends within the respective generational cohorts. This shows that PA is undergoing change in terms of word-medial *imala*, and this change is not an internally-motivated change, in the sense that it is not taking place over generations in a regular, incremental pattern.

To examine the grammar underlying the variable raising of /a:/ to [e:] in word-medial position, we conducted systematic comparisons of the patterns of linguistic conditioning of variant choice across the synchronic and diachronic datasets, where applicable. What do we take from these comparisons? First, by comparing the constraint hierarchy and direction of effects associated with linguistic factors in the LA synchronic dataset and the LPTC, we confirmed that the latter seems to approximate to colloquial LA. We also observed that there does not seem to be an ongoing change in terms of word-medial *imala* in LA, whether in terms of variant rates or conditioning of variant selection. In both the LPTC and the LA synchronic dataset, we saw that preceding and following pharyngeals (in addition to following palatals) disfavoured word-medial *imala*, while other consonants generally had a favouring effect on variant selection.

Second, to address the overarching objective of the present study, which is to examine whether PA is converging on LA, I compared the constraint hierarchy and direction of effects associated with linguistic factors in the LA and PA synchronic datasets (young speech only). The comparison showed that despite minor differences in conditioning patterns, there are broad parallels between the two varieties, especially in terms of the contribution of the preceding phonological segment to variant selection. This confirmed that the change in terms of word-medial *imala* in (young) PA speech seems to be induced by contact with LA. This change seems

to be a relatively superficial linguistic change in Beirut PA from a phonological perspective. In addition this change seems to be from above the level of conscious awareness, because as mentioned previously, some Palestinian participants in this study explicitly associate word-medial *imala* with Lebanese speech patterns.

Trudgill (2008) explains that accommodation or convergence is a mechanistic, largely automatic consequence of face-to-face interaction.⁹⁹ In the present study, this might have been the case for the majority of younger Palestinians who converged on LA speech norms (in terms of word-medial *imala*), as this convergence may have been facilitated by the presumed regular interaction with Lebanese speakers in post-secondary institutions. However, it is important to note that there were a few outliers among the third-generation Palestinians who did not accommodate to LA speech patterns, at least as far as word-medial *imala* is concerned, possibly because of attitudinal or ideological factors.

Among the 11 younger Palestinians interviewed, two speakers failed to produce the [e:] variant in word-medial position (speaker [51]: N=0/52; speaker [57]: N=0/65), and one speaker used word-medial *imala* at a negligible rate (speaker [65]: N=2/41). Speakers [51] and [57], one male and one female, live in refugee camps in Beirut (one in Shatila and one in Burj El-Barajneh), which I already identified at the beginning of this chapter as favourable to maintaining Palestinian speech patterns (see Hennessey 2011:1, 12, 19–20). The female speaker (speaker [57]) explained that she has lived in the camps since birth, and she used to live in Shatila camp with her husband before moving to Burj El Barajneh camp to live with her extended family. But for speaker [51], it is not only camp residence that may have affected his speech patterns. During the interview, he stated that he attended an UNRWA-run school (before

⁹⁹ Tuten (2008:260) states that Trudgill points out that “social and cognitive psychologists have called into question intentionality as the dominant explanation of convergent behavior and have argued increasingly in favor of a more mechanical or automatic view of processes of accommodation.”

going to university outside the camp), and he explained how important it is to maintain his dialect in the following excerpt:

- (11) *...bas hiyye juzu?, inno l-luya ġenna- t-tamassuk bi l-luya walla bi l-lahje, hay juzu? la?anno tħa:fze ġan ħa:- ġa ħa:lek, ġa tura:sek, ġala hawiytek.*

[51/P/M/26/13:43]

‘...but it is a part, I mean the language for us- hanging on to the language or the dialect, this is part (of our identity) because it means maintaining yourself, your (cultural) heritage, your identity.’

Similar to speaker [51], speaker [65], who lives outside the camps but attended an UNRWA-run school (inside a refugee camp), also highlighted the importance of maintaining his Palestinian dialect for the sake of his identity:

- (12) *I’ve been asked actually kti:r inno inte baġdak bet?ulle inno inte xal?a:n hon w ġa:yeš hon [...] yeah I- I’m proud of- of being Palestinians and of choosing to be Palestinian [...] I- I- actually I did not choose- I- I- I chose to speak Palestinian, bħiss it’s not hard at all to speak Lebanese for me, cos I’ve been there and I live there and I- you know, miš safbē, bas why? I’m Palestinian and I’m prou- la?anno ana- ana baħa:reb kerma:l my identity, and language is big part of my identity.*

[65/P/M/24/10:48]

‘I’ve been asked actually a lot, that “you have just told me that you were born here and you live here (in Lebanon)” [...] yeah I- I’m proud of- of being [Palestinian] and of choosing to be Palestinian [...] I- I- actually I did not choose- I- I- I chose to speak Palestinian, I feel it’s not hard at all to speak Lebanese for me, because I’ve been there and I live there [in Lebanon] and I- you know, it’s not hard, but why? I’m Palestinian and I’m prou- because I- I fight for my identity, and language is [a] big part of my identity.’

The excerpts above show that for speaker [51] and speaker [65], there is a personal commitment to maintaining vernacular allegiances, resulting in the resistance to (linguistic) influence from the Lebanese community for the sake of preserving their identity.

To sum up, although the majority of the younger Palestinians accommodate to LA in terms of word-medial *imala*, we saw that in the case of the three aforementioned speakers, there might be attitudinal or ideological factors that may have caused them to resist accommodation

towards LA speech norms (at least in terms of the salient word-medial *imala*).¹⁰⁰ In other words, the urge to converge (Britain 2018) may be hindered by affective allegiance and orientations towards Palestinian speech patterns.

In the next chapter, I turn to the morpho-syntactic variable of verbal negation to examine the extent to which it is susceptible to convergent change.

¹⁰⁰ As Labov (2010:244) notes, “the most convincing and demonstrable determinants of language change are structural and mechanical, but we must be alert to the possibility that ideology is a driving force behind change, as well as a barrier to its further expansion.”

5 Verbal Negation

The second linguistic variable which is targeted in the present study to examine whether PA is converging on LA in Beirut is verbal negation. This chapter presents the results addressing this variable in the Lebanese and Palestinian synchronic and diachronic datasets.

This chapter is structured as follows. Section 5.1 defines the variable context, and section 5.2 explains how the social and linguistic factors were coded, as well as the hypotheses they test. Section 5.3 presents the results regarding verbal negation in the synchronic and diachronic datasets. Section 5.4 discusses the major findings and their contributions to elucidating the possibility of convergent change.

5.1 Circumscribing the variable context

All instances of verbal negation were extracted from the synchronic and diachronic datasets, including negation of pseudo-verbs (3) and participles (4).¹⁰¹ The main variants of verbal negation are illustrated in the examples below, which are reproduced verbatim from the data:

(1) MA
 sent l-ma:de, **ma** se:faret. [04/L/F/24/02:49]
 year the-past NEG traveled.ls
 ‘Last year, I didn’t travel (anywhere).’

(2) MA-š
ma akal-t-š ana. [24/P/M/87/33:07]
 NEG ate.ls-NEG I
 ‘I didn’t eat.’

(3) -š
 fi-š laḥme illa kašer. [POHA/42/02.08.2004/F/68/11:54]
 there.is-NEG meat except Kosher
 ‘There isn’t any meat except for Kosher (meat).’

¹⁰¹ See Chapter 2 for details pertaining to pseudo-verbs and participles in (spoken) Arabic.

- (4) *miš*
 ana ʕa l-xali:j **miš** ra:yih.
 I to the-Gulf NEG going.1s
 ‘I am not going to the Gulf (Arabian Gulf).’
 [LPTC/Bi-n-Nesbeh La Bokra...Chou?/1978/174/M/1:45:15]

Additional variants that occurred mainly in the LPTC (and rarely in the POHA and the LA synchronic dataset; see Table 5.2) are illustrated below:

- (5) *A-š*
 ana a-byanni-š. [LPTC/Bi-n-Nesbeh La Bokra...Chou?/1978/180/M/1:05:23]
 I NEG-sing.1s-NEG
 ‘I don’t sing.’
- (6) *A-*
 ana kamen baʕlif¹⁰² ašya w ešya ktir a-btenkašaf.
 I also know.1s things and things many NEG-be.revealed.3sf
 Intended: ‘I also know many things that cannot be revealed.’
 [LPTC/Wadi Al Ghazzar/1964/59/M/15:41]
- (7) *LA-š*
 la tuɖrub-ni-š. [LPTC/Dawaleeb Al Hawa/1965/172/M/1:24:28]
 NEG hit.2sm-me-NEG
 ‘Don’t hit me.’

All instances of non-verbal negation were excluded, e.g.:

- | | |
|--|---|
| <p>(8) huwwe miš metʕallam.¹⁰³
 he NEG educated
 ‘He is not educated.’
 [39/L/F/53/0:21]</p> | <p>(9) ma meškke.
 NEG problem
 ‘(It’s) not a problem.’
 [23/P/M/54/02:13]</p> |
|--|---|

In addition to excluding instances of non-verbal negation, I excluded instances where *ma* was used as an Indefinite Relative (IR)¹⁰⁴ (see Al-Sayyed and Wilmsen 2017:155) rather than as a negative element, e.g.:

¹⁰² This should be *baʕrif* ‘I know’, but the speaker stuttered and said *baʕlif*.

¹⁰³ There is no overt copula in this example.

¹⁰⁴ Al-Sayyed and Wilmsen (2017:155) explain that *ma* can function as a relative pronoun. For example, the utterance in (10) can be literally translated into ‘Before (*that*) I travel, I want to see her’, and the utterance in (11) can be translated into ‘He was granted citizenship after (*that*) he (graduated) from university’.

- (10) ʔabel **ma** se:fir bidde šu:f-a. [13/P/F/28/12:24]
 before IR travel.1s want.1s see.1s-her
 ‘Before I travel, I want to see her.’
- (11) axad jinsiyye baʕed **ma** xallaş mn j-ja:mʕa.
 got.3sm citizenship after IR finished.3sm from the-university
 ‘He was granted citizenship after he (graduated) from university.’ [59/P/F/50/58:38]

Table 5.1 shows the total number of tokens of verbal negation extracted from each dataset.

Table 5.1: Number of tokens of verbal negation extracted from the synchronic and diachronic datasets

Dataset	N
PA synchronic dataset	1570
LA synchronic dataset	1394
POHA	4015
LPTC	1210
Total	8189

5.2 Factor groups

Several factors were coded to test hypotheses about the social and linguistic predictors of variant selection in verbal negation. These factors are explained in the following sub-sections.

5.2.1 Social and stylistic factor groups

Synchronic datasets

Each speaker in the synchronic datasets was coded for the same social factor groups as implemented in the case of in word-medial *imala*, including age (and immigrant generation for Palestinians only), speaker sex, level of education (primary/secondary versus post-secondary), and place of residence for Palestinians (distinguishing speakers living inside and outside refugee camps in Beirut). In addition, I coded for a stylistic factor group, topic of discourse.

In terms of age or immigrant generation for Palestinians, as mentioned in Chapter 4, it is hypothesized that older Palestinians will be more likely to retain conservative Palestinian speech patterns, which means that they might be expected to use *ma-š* and *-š* more than younger, educated Palestinians. On the other hand, the latter might be inclined to favour *ma*. The preverbal negative particle *ma* is said to be common to all Arabic dialects (Lucas 2010:165, 2020:648), and it is the (supralocal) verbal negator par excellence in spoken Arabic (see e.g., Lucas 2007:403; Mousa 2019:97). In other words, it is the default, socially unmarked variant in most vernacular varieties of Arabic. Mousa (2019:88) explains that in PA, *ma* is more common in urban varieties. By contrast, the variants *ma-š* and *-š* are said to be stigmatized in major Levantine urban centres¹⁰⁵ (see e.g., Fleisch 1974:349, 353; Khairallah 2014:77; Wilmsen 2014:152). In PA, these two variants are claimed to be features of mainly rural varieties (see e.g., Awwad 1987:111; Mousa 2019:88, 97), and there are indications that their use indexes provinciality or lack of cultivation. For example, Wilmsen (2014:152) claims that negation with *-š* is stigmatized as rural speech in Levantine capitals (see also Mousa 2019:19, 97), and Khairallah (2014:77) reports anecdotal evidence depicting negation with *-š* (in LA) as typical of a “hillbilly accent,” again suggesting that it indexes rusticity or provinciality. Khairallah (2014:77) also states that one of her participants (from Mount Lebanon) said that she avoids using *-š* when in Beirut, instead negating the ‘normal way’ (*sic*) in that city by using preverbal *ma* only. Evidence of an urban-rural divide in terms of variant choice in verbal negation in PA is also explicitly invoked by one of the speakers recorded for the present study:

¹⁰⁵ Wilmsen (2014:152) argues that “the loss of *-š* in negation came about after Muslim speakers of central and eastern Arabic dialects—carrying a prestigious new religion into the region, in the process of forming an empire—became the new urban elites, and their dialects, that had never negated with *-š*, became the objects of imitation. The *-š* in the speech of the earlier inhabitants, adherents of the older religions, became stigmatized, just as it is stigmatized as rural speech in Levantine capitals of today, and was easily abandoned. That it had become scarcely more than a clitic associated with the same negator as that used in the eastern dialects, *mā*, facilitated its loss.”

- (12) *Kul balad fi ila lahje xa:šša [...] halla? ahelna ihna falla:hi:n [...] kullo- masalan bašrefiš, ma ruḥtiš, ma ji:tiš. Dayman š-ši:n la:zem tku:n mertebta bi l- bi l-kelme [...] halla? ihna, masalan, mni:je minʔu:l ma bašref [...] halla? betʔarbe inte šala l-mudun.* [58/P/M/45/1:09:49]

‘Every area (in Palestine/Israel) has its own dialect [...] Now our parents are peasants [...] all of them- for example, *bašrefiš* ‘I don’t know’, *ma ruḥtiš* ‘I didn’t go’, *ma ji:tiš* ‘I didn’t come’. The -š always has to be attached to the- to the word [...] Now for us, for example, we say *ma bašref* ‘I don’t know’[...] Now you are close to *the cities/urban areas*’ [emphasis mine]

Further evidence of the marked social meaning with which negation with -š seems to be imbued comes from the LPTC in the excerpt below, reproduced from Chapter 3. The excerpt is taken from the 1977 play *Bint El Jabal* ‘Mountain Girl’ which is based on George Bernard Shaw’s 1913 stage play *Pygmalion*. In the Lebanese play, the protagonist, Lisa, is a flower seller who is depicted as speaking a “vulgar” variety of Lebanese Arabic characteristic of her village.¹⁰⁶ She meets a linguist, Henry Adib, who teaches her to speak “proper” Lebanese Arabic, which does not include the stigmatized forms she uses, as seen in the excerpt below:

- (13) [Lisa] *Walaw, a-betxalli:ne ʔu:la-š marra te:ne?* [Henry] *ma betxalli:ne, ma betxalli:ne.* [Lisa] *Mi:niyy a-bexhallik-š?* [Henry] *Lek ya Lisa, ana šam bšaḥḥlik, šam bʔellik beʔu:lo “ma betxalline ʔu:la marra te:nye?”* [Lisa] *ʔaš šley, ma hayde yeir ši, haydi:k šelem, w hayde ḥdi:s. ʔena bʔu:l “a-betxalli:ne ʔu:la-š,” w inta betʔu:l “ma betxalli:ne ʔu:la” [...] Ma kollo ʔawl b-ʔawl, ma hak-š?* [Henry] *W lek ya Lisa, inte halla? je:ye tetšallame l-laḥz l-maḥbu:t, ma hek? Ntabahte inno ʔeltillik “ma hek?” Ma ʔeltillik “ma hak-š?” Yašne inte le:zim tašrfe tmayze bein l-laḥz š-šaḥi:h, w l-laḥz l-mubtazal. Halla? iza baddik, ʔuli:ha marra te:nye, w iza betri:de nsi:le kelmit “ma hak-š.”* [LPTC/Bint El Jabal/1977/185, 189/F, M/1:04:44]

‘[Lisa] Won’t you let me say it [*a-betxalli:ne ʔoula-š*] one more time? [Henry] Won’t you let me [*ma betxalli:ne*], won’t you let me [*ma betxalli:ne*]. [Lisa] Who’s the one who won’t let you [*a-betxallek-š*]?’ [Henry] Lisa, I am correcting you, I am telling you they say ‘won’t you let me [*ma betxalli:ne*] say it one more time?’ [Lisa] It’s okay, this is different, this one (the way you say it) is for teaching, and the other one (the way I say it) is speech. I say *a-betxalli:ne ʔu:la-š* ‘won’t you let me say it’ and you say *ma betxalli:ne ʔu:la* ‘won’t you let me say it’ [...] It’s all speech, isn’t it [*ma hak-š*]?’ [Henry] Lisa, now you are coming to learn proper pronunciation, aren’t you [*ma hek*]?’ Did you notice that I

¹⁰⁶ In the play, Lisa claims that she is from a village called Kfar Yanouh, but this appears to be a fictional village. In Lebanon, there is a village in Mount Lebanon called Yanouh. It is not clear if Kfar Yanouh is based on the actual village of Yanouh.

said *ma hek* ‘aren’t you?’ I didn’t say *ma hak-š* ‘aren’t you?’ So you have to know how to distinguish between proper pronunciation and vulgar pronunciation. Now if you want, say it one more time, and please forget the word *ma hak-š* ‘aren’t you’.

Because of the apparent stigma accorded to *ma-š* and *-š*, and their association with rural (Levantine) varieties (Wilmsen 2014:152), younger, educated speakers are hypothesized to disfavour them. This is consistent with Mousa’s (2019:97) claim that in the villages in the northwest part of the West Bank, younger rural PA speakers tend to avoid using *ma-š* and *-š* in favour of the urban form, *ma* (see also Abdel-Jawad 1987:361–62; Amara 2005:885–86, 890–91). On the other hand, older Palestinians tend to preserve emblematic dialect features (see e.g., Hennessey 2011:29), so as mentioned above, they may be more likely to retain *ma-š* and *-š* than their younger counterparts.

In addition to the stigma associated with *ma-š* and *-š*, it should be noted that neither variant is allegedly characteristic of indigenous Beirut Arabic, which is said to use mainly *ma* for verbal negation (Wilmsen 2014:150; see also Behnstedt and Woidich 2005:101; Fleisch 1974:349, 353; Khairallah 2014:76). The working hypothesis adopted in this chapter is that younger Palestinians in Beirut would be more likely to conform to Beirut Arabic speech patterns (including *ma*), following Fityan’s (1981) findings. Fityan (1981) examined patterns of variation and change among older and younger Palestinian refugees living in a camp outside Beirut, targeting a phonological variable, the *qaf* variable (q).¹⁰⁷ He found that older speakers preserved conservative speech norms in relation to this variable, while younger Palestinians favoured socially acceptable forms characteristic of Lebanese speech norms. Fityan (1981:48–49) explains

¹⁰⁷ The variants of the *qaf* variable (q) are [q], [q^h], [ʔ], [k] and [g] (Fityan 1981:24). The [q] variant is reportedly acceptable by the community in the camp and is characteristic of Classical Arabic (Fityan 1981:64), while [ʔ] is said to be characteristic of urban PA speech patterns as well as Lebanese speech patterns in Baalbeck (Fityan 1981:56–57, 102), where Fityan conducted his study. These two variants – [q] and [ʔ] – were found to be favoured by younger speakers (Fityan 1981:101), while the remaining variants, described as ‘relic’ forms by Fityan (1981:102), were disfavoured or avoided altogether by younger speakers.

his findings by arguing that older (and less educated) Palestinians generally have no social aspirations that motivate them to modify their speech to fit in the wider Lebanese society. On the other hand, younger (and more educated) Palestinians may be concerned about the social acceptability of their speech, and their aspirations for upward social mobility lead them to adopt Lebanese speech patterns which they acquire from their peers in school (see Fityan 1981:58). This is another reason why younger, educated Palestinians in the present study will be likely to favour *ma*, especially considering that they are in more intense contact with Beirut Arabic speakers than older Palestinians, mainly in post-secondary institutions (see below).

Speaker sex is also predicted to be a key determinant of variant selection. It is anticipated that females will favour *ma*, while males will use *ma-š* and *-š* more than females, in line with the well-attested sociolinguistic gender pattern (Fasold 1990:92), where females tend to use fewer stigmatized features than males (Labov 1972b:243), and are more inclined to adopt supralocal features instead. This is because they are reported to show a greater sensitivity to socially evaluative linguistic forms than do males (Chambers and Trudgill 2004:85; Cheshire 2002:430; Labov 1966; Milroy 1980; Milroy et al. 1994:352; Trudgill 1972:180, 1983:161; Wolfram 1969:78; Wolfram and Fasold 1974:93). This pattern of sex-differentiation in speech has been attested in several studies in the Arabic-speaking world, including previous treatments of dialect contact situations, where women were found to favour urban (supralocal) forms more than men (Al-Wer 1997:261; see e.g., Abdel-Jawad 1981, 1987:360–61; Al-Wer 2007b:62, 66–67; Amara 2005:898; Cotter 2013; Cotter and Horesh 2015; Haeri 1987; Hennessey 2011:92–93).

With regards to educational level, as mentioned previously, this independent variable interacts with age because all young Palestinians are educated to post-secondary level, while

older Palestinians are educated to primary level.¹⁰⁸ The expectation is that young and educated speakers will favour *ma*, while their older and less educated counterparts will favour *ma-š* and *-š*. As explained in Chapter 4, educational level functions as a proxy measure of degree of contact with LA speakers (Hennessey 2011:111–12; see Al-Wer 2002:42). This means that younger and educated Palestinians are believed to interact regularly with Lebanese speakers in post-secondary institutions (outside the camps), which favours the diffusion of Beirut Arabic speech features to younger Palestinians, including *ma*. On the other hand, less educated (and older) Palestinians are less likely to have similar opportunities to interact with LA speakers on a regular basis (see Fityan 1981:102). Further evidence to that effect emerges from the questionnaire administered to participants at the end of sociolinguistic interviews, including a frequency metric to assess how often Palestinians come into contact with Lebanese speakers (i.e., ‘very often,’ ‘quite often,’ or ‘occasionally’). As would be expected, several younger Palestinians (6 out of 11)¹⁰⁹ stated that they come into contact with Lebanese speakers very often, while 6 out of the 8 older, less educated speakers stated that they come into contact with the local Lebanese population only occasionally. Even among middle-aged speakers, 7 of the 9 speakers educated to post-secondary level reported that they regularly interact with Lebanese speakers, while 6 of the 10 less educated middle-aged speakers reported only occasional contact with Lebanese speakers. These self-reported data corroborate our earlier hypothesis that younger and more educated Palestinians are in more intense contact with Lebanese speakers than older and less educated Palestinians. Because the sampling population includes less educated middle-aged Palestinians, as well as middle-aged speakers educated to post-secondary level, this affords additional insights into the

¹⁰⁸ The only older female educated to post-secondary level was excluded because she did not produce enough tokens of verbal negation.

¹⁰⁹ It should be noted that 4 out of the 11 younger speakers did not answer the question about their degree of contact with their Lebanese counterpart.

role that educational level may play in the selection of variable expressions to mark verbal negation.

In terms of place of residence for Palestinians, we conjecture that insofar as *ma-š* and *-š* are used, they will be comparatively more frequent in the refugee camps than in PA spoken outside the camps. As explained in Chapter 4, Palestinian camp-dwellers are assumed to be part of dense, multiplex networks with other Palestinians in the camp where they live (Hennessey 2011:1, 12, 19–20), and these networks are believed to favour the maintenance of Palestinian vernacular norms, which may involve the use of *ma-š* and *-š*. On the other hand, Palestinians in Beirut who live outside the camps tend to have stronger ties with LA speakers (Hennessey 2011:33–34, 103, 131, 140), which means that they are more likely to use *ma* characteristic of Beirut Arabic, and less likely to use socially stigmatized features in Beirut, like *ma-š* and *-š*. However, it is possible that some Palestinian non-camp-dwellers would approximate to Palestinian speech norms because these norms may have covert prestige that may render them more resistant to change (Hennessey 2011:12–13). At the same time, it is conceivable that some Palestinian camp-dwellers, particularly the younger (and educated) ones, may show a preference for the socially unmarked variant, *ma*, as camp residence is claimed not to act as a complete barrier to the assimilation of LA speech patterns (Hennessey 2011:34). This was already seen in Chapter 4, where young, educated camp-dwellers used word-medial *imala* to a certain extent (N= 35/169 or 21%) because they are in regular contact with LA speakers in post-secondary institutions outside the camps.

For the LA synchronic dataset, because all young and middle-aged speakers are from Beirut (and are educated to post-secondary level), we envisage that they will favour *ma*, characteristic of Beirut Arabic. However, for older (and less educated) speakers, only one of

them is from Beirut, while the remaining four are from Cornet Chahwan in Mount Lebanon, which is around 20 km away from Beirut. Figure 5.1 shows Cornet Chahwan on the map.

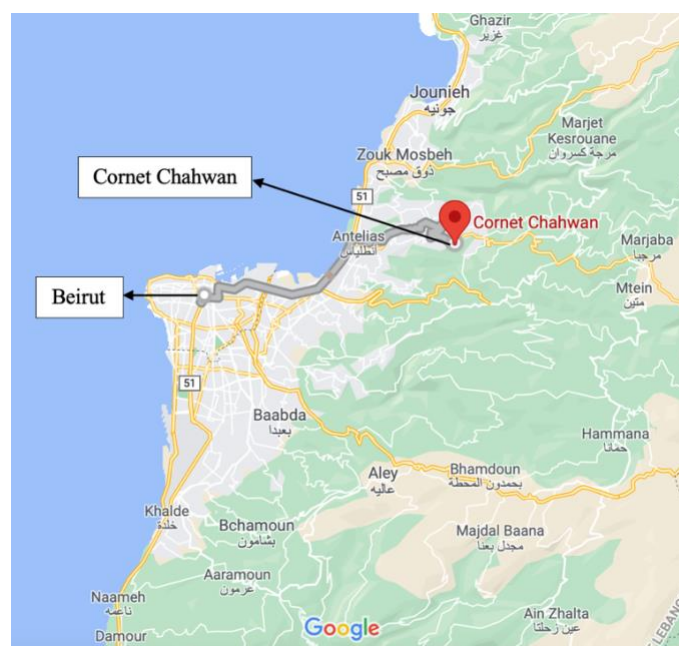


Figure 5.1: Location of Cornet Chahwan (and Beirut)
(Source: Google Maps)

Because the variants *ma-š* and *-š* are said to be characteristic of LA as spoken in Mount Lebanon (Abu-Haidar 1979:109–10; Fleisch 1974; Khairallah 2014:17), we speculate that older speakers who are from Cornet Chahwan will favour these two variants.

For both the Lebanese and Palestinian synchronic datasets, I also coded for topic of discourse to examine the potential effect of stylistic factors on variant selection in verbal negation. Originally, the coding protocol distinguished between different topics of discourse, such as religion (14), language (15), education (16), moralizing or soapbox speech (17), politics (18), miscellaneous topics, and narratives.

- (14) **ma** fi masi:he w muslim ŋen-na ka:n.
NEG there.is Christian and Muslim have.3p was.3sm
Intended: ‘There wasn’t any difference between Christians and Muslims (in our village).’
[32/P/M/79/50:19]

- (15) **ma** mnaʃref nʃabbir bi l-fuʃha. [05/L/F/31/29:51]
 NEG know.1p express.1p in the-Standard
 Intended: ‘We don’t know how to express ourselves in Standard Arabic.’

- (16) **ma** ka:n hada l-ʃelem metl l-yōm. [24/P/M/87/31:32]
 NEG was.3sm this the-education like the-day
 ‘There wasn’t (much) education like today.’

- (17) **ma** fi ʃi byedman huʔu:ʔ l-mara; **ma** fi
 NEG there.is thing guarantee rights the-woman NEG there.is
 ʃi yeir l-madane.
 thing other.than the-civil
 ‘There is nothing that guarantees women’s rights; there is nothing except civil (marriage).’
 [02/L/F/26/26:13]

- (18) **ma** beʃerfo la Sykes-Picot wala- wala waʃd Balfour.
 NEG know.3p NEG Sykes-Picot NEG NEG promise Balfour
 ‘They don’t know about Sykes-Picot (Agreement) nor about Balfour Declaration.’
 [08/P/F/20/41:40]

Because of a limited number of tokens within each topic of discourse exemplified above, these topics were later collapsed together into soapbox speech (e.g., topics on religion, language, education, politics, etc.), miscellaneous topics and narratives in order to conduct meaningful quantitative analyses. The working assumption is that because *ma-š* and *-š* are reportedly stigmatized and associated with rural (Levantine) varieties, more careful or soapbox speech may inhibit the use of these two variants, and alternatively favour the use of *ma*.

Diachronic datasets

In the POHA, speakers were coded for sex, educational level and place of residence (inside versus outside camps). It should be noted that there is extensive interaction between these three social factor groups. For example, among the 80 speakers in the POHA sampled for this study, the only speakers who are educated to post-secondary level are six males, all of whom live outside refugee camps. Due to this interaction, it was necessary to collapse the three social factor

groups together as follows: male/ post-secondary/ outside camp; male/ non-post-secondary/ outside camp; male/ non-post-secondary/ inside camp; female/ non-post-secondary/ outside camp; female/non-post-secondary/inside camp. Because of their higher educational levels and their residence outside the camps, the six educated males are projected to use *ma* more frequently than *ma-š* and *-š*. We also hypothesize that females, especially those who live outside the camps, will disfavour the latter two variants, while less educated males, especially those who live inside the camps, will use *ma-š* and *-š* more frequently.

In the LPTC, each token was coded for sex, age and social status (higher, middle and lower social status). We hypothesize that the variants *ma-š* and *-š* will be used more frequently by males in the LPTC than by females, as men are reported to be more favourably disposed than women to low-status speech forms (see e.g., Chambers and Trudgill 2004:85; Tagliamonte 2012:33; Trudgill 1972:182–83). Additionally, *ma-š* and *-š* might be used more frequently by older characters and by those who have a lower social status. The hypothesis for social status is consistent with findings of previous sociolinguistic studies which have taken social class into consideration, where higher social classes tend to use prestige (or socially unmarked) forms, while lower social classes tend to use more stigmatized features (Habib 2008; Labov 1972b:113).¹¹⁰

¹¹⁰ It is important to note that this might not always be the case. For example, in Labov's (1972b) study of speech patterns in New York City, he found that lower social classes sometimes use the prestige form in an attempt to appear prestigious (see also Trudgill 1974). So speakers from lower social classes try to imitate those from higher social classes because they aspire to be like them. Speakers from lower social classes (especially females) may even use the prestige form more frequently than those from higher social classes, which Labov (1966, 1972b) calls 'hypercorrection' towards the more prestigious form, i.e., "speakers who exhibit this patterning are speaking in a hyper-'correct' or hyper-standard style" (Wolfram and Schilling-Estes 2016:288). Such hypercorrection is also described by Labov (1966:227) as 'apparent deviation' and a 'cross-over pattern' (Labov 1972b; see Wolfram and Schilling-Estes 2016:289). This cross-over pattern is most likely to be observed with features that are undergoing change, such as the increasing use of *r*-ful pronunciations in New York City in the 1960s (Wolfram and Schilling-Estes 2016:289; see Labov 1972b). Based on previous findings, it may be the case that lower social status characters will use *ma* at a relatively high rate as a form of 'hyper-correction'.

As in the case of the synchronic datasets, I also coded for topic of discourse in the diachronic datasets (the POHA and the LPTC), including soapbox speech, miscellaneous topics and narratives. The hypothesis is that *ma-š* and *-š* will be less likely to be used in soapbox speech in the diachronic datasets.

5.2.2 Linguistic factor groups

In all datasets, each token was coded for several linguistic factor groups hypothesized to contribute to variant choice. These hypotheses were abstracted from the existing relevant literature (see e.g., Blau 1960; Farwaneh 2021; Hoyt 2006, 2010; Lucas 2007, 2010; Obler 1975), as well as from informal observations of the data. The first factor group relates to verb form. I coded for imperfect verbs,¹¹¹ including verbs in the present tense (19) and present progressive tense (20); perfect verbs (past tense verbs) (21); future tense verbs (22); participles (23); and pseudo-verbs (24).

(19) ana ma **bru:h** ʃand l-coiffeur. [02/L/F/26/15:43]
 I NEG go.1s to the-coiffeur/hairdresser
 ‘I don’t go to the coiffeur/hairdresser.’

(20) ma **ʃam** **beħke** ʃan l-musi:ʔa l-pop. [09/L/M/21/12:42]
 NEG PROG talk.1s about the-music the-pop
 ‘I am not talking about pop music.’

(21) ma **ru:h-t-š** ʃa l-beit. [45/P/F/78/42:31]
 NEG went.1s-NEG to the-house
 ‘I didn’t go home.’

(22) ma **rah** **yʔul-l-ak** ši. [POHA/29/15.05.2003/M/72/37:43]
 NEG FUT tell.3sm-to-you.m thing
 ‘He will not tell you anything.’

(23) miš **de:ris** ši. [18/L/M/33/27:34]
 NEG studied.1sm thing
 ‘I haven’t studied anything.’

¹¹¹ “Imperfect verb forms typically express present time reference and are mainly prefixing in their inflection” (Lucas 2010:167).

- (24) ana ma **badde** etjawwaz. [LPTC/Mais El Reem/1975/98/F/1:20:07]
 I NEG want.1s marry.1s
 ‘I don’t want to marry.’

Verb form is hypothesized to be a key linguistic determinant of variant choice in verbal negation. In PA, *ma* is said to negate present tense and past tense verbs (Mousa 2019:86), future tense verbs (Mousa 2019:100–101), as well as pseudo-verbs (Lucas 2010). *Ma-š* is said to be favoured with past tense verbs because it is claimed that *-š* alone cannot negate such verbs in PA (Abu-Haidar 1979:109; Hoyt 2006, 2010; Lucas 2010; Obler 1975; Shahin 2000:37). Rather, *-š* is said to be favoured mainly with present tense verbs in PA (Lucas 2010:171), though in LA, Thackston (1996:145) claims that *-š* can be suffixed to all verbs, including past tense verbs (see also Khairallah 2014:36). In PA, *-š* is also said to be favoured with the majority of pseudo-verbs, except with the pseudo-verb *šand-* ‘have’ because it begins with a voiced pharyngeal fricative, not with a labial consonant (see below) (Lucas 2010:171). Both *ma-š* and *-š* are said to be disfavoured with future tense verbs, at least as far as PA is concerned (Mousa 2019:100–101). Instead, it is claimed that *miš* is mainly favoured with future tense verbs in PA, as well as with participles (Brustad 2000:289; Mousa 2019:100–101).

Extrapolating from previous claims in the relevant literature, it is conjectured in the present study that *ma* will be favoured with all verb forms except for participles, *ma-š* will be favoured with past tense verbs, and *-š* with present tense verbs and with pseudo-verbs in PA. However, Lebanese speakers (mainly older speakers from Mount Lebanon) may show a preference for *-š* with past tense verbs (based on Thackston’s (1996) claim). We posit that *miš* will be favoured with participles and with future tense verbs, which is why these two verb forms were collapsed together in the coding protocol.

Another linguistic factor group coded for is initial consonant of the negated verb (or pseudo-verb), which is also widely believed to contribute to variant choice in verbal negation (Bauer 1926; Blau 1960; Driver 1925:197; Hoyt 2006; Obler 1975:107; Palva 2004; Wilmsen 2014). In this factor group, two factors were included: labial-initial verbs, as in those beginning with the labial obstruents [b] and [f] (see e.g., Hoyt 2006:16), and non-labial-initial verbs. Hoyt (2006:16) claims that omission of *ma* is only possible with labial-initial verbs, and Blau (1960:198) explains that *ma* is deleted because of its phonological similarity to the *b*-prefix that is characteristic of most imperfect or present tense verbs (in Levantine Arabic). Accordingly, we believe that in PA, *-š* will be favoured with labial-initial verbs, while *ma-š* (or *ma*) will be favoured with non-labial-initial verbs (see e.g., Lucas 2010). However, it should be noted that there are some scholars who claim that *-š* alone can be used with non-labial-initial verbs (Bauer 1926; Driver 1925:197; Obler 1975:107; Palva 2004; Wilmsen 2014). In LA, for example, Thackston (1996) claims that *-š* can negate non-labial-initial verbs in the past tense, as in *kuntⁱ-š* ‘I wasn’t’.

The third linguistic factor group coded for is intervening material between the negative element and the negated verb. Intervening material could be a progressive marker (25), a future tense marker (26), copula (27) or clitic(s) (28).

(25)	miš	šam	yle:ʔe	šiyil.	[40/L/M/83/25:53]
	NEG	PROG	find.3sm	job	
	‘He is not finding a job.’				

(26)	miš	rah	ašuf-a	ana.	[45/P/F/78/1:44]
	NEG	FUT	see.1s-her	I	
	‘I will not see her.’				

(27)	ma	kan-š	fi	maṭṣam	bi	ḍ-ḍayʕa.	[42/L/M/89/09:37]
	NEG	was.3sm-NEG	there.is	restaurant	in	the-village	
	‘There wasn’t any restaurant in the village.’						

- (28) ma ʔebl-**o-wa**-š. [22/P/M/66/52:08]
 NEG accept-3p-her-NEG
 ‘They did not accept her (in her visa application).’

Additional intervening material included serial verb constructions.¹¹² Versteegh (2009:195) explains that serial verbs “indicate a verbal syntagm consisting of two (or more) finite verbs without a formal coordinating marker but with the same argument structure, one of which is semantically demoted, often grammaticalized, and lexically restricted” (see also Altakhaineh and Zibin 2018:47; Haspelmath 2016:296). In these constructions, “[n]egation is always marked on the first verb in the string and applies to the whole string” (Hussein 1990:343; see also Altakhaineh and Zibin 2018:47; Muysken and Veenstra 1995), and this was observed in the data for the present study. Some examples of serial verb constructions from the speech data are found below:

- (29) ma **ʕad** šef-t-o. [13/P/F/28/03:19]
 NEG returned saw.1s-him
 ‘I didn’t see him anymore.’

- (30) yefḏa l-maḥal, ma **yʕud**-š fi ši. [41/L/F/72/16:23]
 emptied.3sm the-store NEG return-NEG there.is thing
 Intended: ‘The store gets empty, there wouldn’t be anything anymore.’

Other intervening material includes the verb *baʔa/baʔ* (lit. ‘remained’), though it is not clear from the literature if constructions with this verb constitute serial verb constructions. Some examples of this verb are found below:

- (31) ma **baʔa** nšu:f-on. [27/P/F/61/04:07]
 NEG remained see.1p-them
 ‘We don’t see them anymore.’

¹¹² Woidich (2002) states that in Egyptian Arabic, the so-called ‘serial verb constructions’ have nothing in common with the serial verbs in other languages, so he prefers to call them instead ‘verbal phrases with asyndetic perfect’ (as cited in Versteegh 2009:196).

(32) ma **baʔ** etzakkar. [07/L/M/43/04:05]
 NEG remained remember.1s
 ‘I don’t remember anymore.’

(33) ma **baʔ-u:-š** yeʔbal-o yxall-o ɥada yɣallem.
 NEG remained.3p-NEG accept.3p allow.3p someone teach.3sm
 ‘They didn’t allow anyone to teach anymore.’ [POHA/18/20.02.2003/F/78/02:09]

(34) ma **baʔa-š** fiyye etjawwaz. [LPTC/Bint El Jabal/1977/194/M/02:07:05]
 NEG remained-NEG can.1s marry.1s
 ‘I cannot marry anymore.’

The effect of intervening material between the negative element and the negated verb on variant selection has not been examined previously in the literature, so here it is considered an exploratory manoeuvre. Based on informal observations of patterns in the data, however, it is possible that intervening material between the first negator and the verb may favour a second negator to reinforce the first negative element, as shown in examples (27), (30), (33) and (34). Additionally, cursory inspection of the data suggests that intervening material, such as future tense markers or progressive markers, favours the use of *miš*, as in (25) and (26).

5.3 Results

5.3.1 Overall distribution of variants

Table 5.2 shows the overall distribution of the variants of verbal negation in the four datasets.

Table 5.2: Overall distribution of the variants of verbal negation in the synchronic and diachronic datasets

(Note: shaded areas indicate that the variant(s) in question was not found in the dataset listed in the left-hand side of the table)

	ma		ma-š		-š		miš		a-š		Other	
	%	N	%	N	%	N	%	N	%	N	%	N
PA synch.	85	1330/1570	4	67/1570	6	91/1570	5	82/1570				
LA synch.	92	1282/1394	2	29/1394	1	8/1394	5	74/1394	0.07	1/1394		
POHA	61	2451/4015	22	895/4015	13	505/4015	4	163/4015	0.02	1/4015		
LPTC	80	967/1210	4	47/1210	1	14/1210	10	121/1210	4.7	57/1210	0.3	4/1210

From the table above, we can see that *ma* is the majority variant in the four datasets; however, it occurs at a considerably lower rate in the POHA, although it is still the default variant in that dataset. This is because *ma-š* and *-š* occur more frequently in the POHA than in the other datasets.

The table above also shows that the LA synchronic dataset, the POHA and the LPTC have additional variants of verbal negation that were not found in the PA synchronic dataset. The LA synchronic dataset has only one token of *a-š* that was produced by an older male from Cornet Chahwan. Similarly, the POHA has one token of *a-š* that was produced by a male speaker. The LPTC also has tokens of *a-š*, as well as other variants that comprise two tokens of pre-verbal *a-* and two tokens of the variant *la-š*. The two tokens of *a-* were produced by one middle-aged male belonging to the middle social status and one older male belonging to the lower social status. The two tokens of *la-š* were produced by one middle-aged male and one older male, both of whom have lower social status.

Khairallah (2014:17) states that the variant *a-š* is found in Lebanese varieties spoken in southern and mid-mountainous regions of Lebanon, such as in Keserwan District in the Mount Lebanon Governorate (see also Cowell [1964] 2005:383), but as the table above shows, *a-š* occurred only once in the LA synchronic dataset by a speaker from Cornet Chahwan, which is also Mount Lebanon Governorate. Palva (2004:227) also claims that in northern Lebanon, verbs can be negated with the preverbal *a-*. As mentioned in Chapter 2, this variant is reportedly a reduced form of *ma* (see Lucas 2010:186), as Abu-Haidar (1979:110) states that the [m] in *ma* can be elided when it precedes a labial consonant, becoming *a-*.¹¹³ The variant *la-š* is used for the imperative, as we saw in (7), and Abu-Haidar (1979) found that in Baskinta (a village in Mount Lebanon, located around 43 km north east of Beirut), the variant *la-š* is used for the imperative

¹¹³ In the two tokens of *a-* in the LPTC, this variant did occur before the labial consonant [b].

as well. In PA, Cowell ([1964] 2005:383) claims that *a-š* is used, but as the table above shows, this variant occurred only once in the POHA, and it never occurred in the PA synchronic dataset.¹¹⁴

In terms of convergence of PA on LA, the working hypothesis is that superficially similar structures across these two varieties in terms of verbal negation should render them hospitable to contact-induced change. Evidence of convergence of PA on LA should be manifested in terms of similar variant distributions in both varieties, and more crucially, in terms of parallels in the ranking of linguistic constraints on variant choice across the two datasets. In the section below, I examine the social and stylistic constraints on variant choice in the synchronic and diachronic datasets.

5.3.2 Social and stylistic factor groups

Before examining the social and stylistic factors contributing to variant choice in verbal negation, a note is in order. We saw in the previous section that even though *ma* is quantitatively preponderant in the four datasets, the POHA is the only dataset where *ma-š* and *-š* occur to a certain extent (22% and 13% respectively), whereas in the remaining datasets, variant distributions show more restricted variability in the expression of verbal negation. That is, in the PA and LA synchronic datasets, as well as in the LPTC, there is a great imbalance in overall variant distributions, especially in terms of *ma-š* and *-š* (as well as *a-š*, *a-* and *la-š* in the LPTC), which occur at rate of around 5% or less, compromising the reliability of multivariate analysis in these datasets, if not making it unfeasible (Guy 1988:125; Tagliamonte 2006:84, 86–87).¹¹⁵ For

¹¹⁴ Lucas (2010:186) explains that “certain Lebanese varieties [...] have a reduced form of *mā* – ‘a – in similar contexts to those in which the postverbal construction [-š] is found in Palestinian, raising the possibility that this represents a stage that Palestinian passed through before it became possible to omit *mā* entirely in these contexts.”

¹¹⁵ Tagliamonte (2006:84) explains that “if the data at your disposal is near categorical (either 100 per cent or 0 per cent), then there is little room for quantitative investigation.” This means that if variability hovers at very low levels, such as around 5% or under, or at very high levels, such as around 95% or over (Tagliamonte 2006:87), “differences

this reason, from this section forward, I only present distributional results for the PA synchronic dataset, the LA synchronic dataset and the LPTC, while I present multivariate analyses for the POHA where data distributions are amenable to statistical analysis.

Synchronic datasets

Table 5.3 shows the distributional analyses of social and stylistic factors in the LA and PA synchronic datasets. There is only one token of *a-š* in LA, so it is excluded from the distributional analyses.

between variants in independent contexts may be too small to achieve statistical significance” (Tagliamonte 2006:84). It is for this reason that I exclude PA, LA and the LPTC from multivariate analysis.

Table 5.3: Overall distribution of variants according to social and stylistic factors in the LA and PA synchronic datasets (all age groups)

LA synchronic dataset (all age groups)								PA synchronic dataset (all age groups)									
ma			ma-š		-š		miš		ma			ma-š		-š		miš	
	%	N	%	N	%	N	%	N		%	N	%	N	%	N	%	N
Total N	92	1282/1394	2	29/1394	1	8/1394	5	74/1394	Total N	85	1330/1570	4	67/1570	6	91/1570	5	82/1570
Sex									Sex								
Male	88	405/462	4	19/462	1	6/462	7	31/462	Male	78	547/700	8	51/700	10	72/700	4	30/700
Female	94	877/932	1	10/932	0.2	2/932	4.8	43/932	Female	90	783/870	2	16/870	2	19/870	6	52/870
Age/ Education									Age/ Education								
Young/PS	95	883/931	0	0/931	0	0/931	5	48/931	Young/PS	94.8	440/465	0.2	1/465	2	9/465	3	15/465
Mid-aged/PS	94	311/332	0	0/332	0	0/332	6	21/332	Mid-aged/ PS	92	321/349	0.3	1/349	0.7	2/349	7	25/349
									Mid-aged/ non-PS	77	310/405	8	33/405	9	38/405	6	24/405
Old/non-PS	67	88/131	22	29/131	6	8/131	4	5/131	Old/non-PS	74	259/351	9	32/351	12	42/351	5	18/351
									Residence								
									Outside camp	90	944/1049	3	26/1049	2	22/1049	5	57/1049
									Inside camp	74	386/521	8	41/521	13	69/521	5	25/521
Topic									Topic								
Soapbox speech	94	101/107	0	0/107	0	0/107	6	6/107	Soapbox speech	80	171/213	4	8/213	11	24/213	5	10/213
Narrative	85.5	177/207	7	15/207	0	0/207	7	14/207	Narrative	81	281/346	7	25/346	7	25/346	5	15/346
Misc.	93	1004/1080	1	14/1080	1	8/1080	5	54/1080	Misc.	87	878/1011	3	34/1011	4	42/1011	6	57/1011

In the LA synchronic dataset, *ma* is used at a slightly higher rate by Lebanese females than by males, while it is notable that (older) males use *ma-š* and *-š* at a higher rate than the only female from Cornet Chahwan. Similarly, in the PA synchronic dataset, females use *ma* for verbal negation more than males, while the latter favour *ma-š* and *-š*. Because these two variants are stigmatized, we assumed that females will eschew them in favour of the unmarked default exponent of verbal negation, *ma*, while men, especially less educated ones, will use low-status linguistic forms like *ma-š* and *-š* more frequently. Unlike the case for these two variants, *miš* is used by males and females at similar rates in the LA and PA datasets, which means that *miš* may not be socially diagnostic, at least as far as speaker sex is concerned.

In terms of age, which is collapsed with education, in the LA dataset, *ma* is the majority variant for young and middle-aged speakers (all of whom are educated to post-secondary level). These speakers don't produce any instance of *ma-š* and *-š*. This would seem to confirm that these two variants are not productive features of contemporary Lebanese Arabic as spoken in Beirut (Behnstedt and Woidich 2005:101; Fleisch 1974:349, 353; Khairallah 2014:76; Wilmsen 2014:150). Additionally, the only older male from Beirut uses *ma* quasi-categorically to express verbal negation (N=59/60). *Ma-š* and *-š*, then, were only used by older speakers from Cornet Chahwan. This would provisionally seem to suggest that these two variants are characteristic of LA as spoken in Mount Lebanon (Abu-Haidar 1979:109–10; Fleisch 1974; Khairallah 2014:17). Therefore, the differences between age groups in LA in terms of variant rates are heavily influenced by speakers' place of origin (i.e., Beirut vs. Mount Lebanon). For *miš*, the results present further evidence that this variant is not socially diagnostic because it is distributed at similar rates across age groups/educational levels in the LA synchronic dataset.

The PA synchronic dataset offers key insights into the contribution of educational level to variant choice in verbal negation. As Table 5.3 shows, *ma* is favoured not only by younger speakers, but also by middle-aged speakers who are educated to post-secondary level. On the other hand, *ma-š* and *-š* are favoured by less educated middle-aged and older Palestinians, especially males (out of the 74 tokens of *ma-š* and *-š* produced by older speakers, only 17 tokens came from women; and out of the 71 tokens of *ma-š* and *-š* produced by less educated middle-aged speakers, only 11 tokens came from women). This is in line with the hypothesis that not only younger speakers, but also more educated speakers, even among the middle-aged Palestinians, will be less inclined to use *ma-š* and *-š* because these variants seem to be burdened with negative social freight in Beirut, possibly denoting provinciality and uneducatedness. On the other hand, less educated middle-aged and older speakers, who are presumed to be in less intense contact with Lebanese speakers, adhere more closely to *ma-š* and *-š*. As we saw in the LA dataset, *miš* in PA is distributed at similar rates across age groups/educational levels, though younger speakers use it at a slightly lower rate than their older counterparts.

A first provisional conclusion that we can draw based on the distributional results for age/educational level in the PA synchronic dataset is that the system of verbal negation in PA is becoming gradually streamlined, with *ma* sweeping the system, especially in the speech of younger and more educated speakers. Accordingly, it seems that the system of verbal negation in PA used by younger and more educated speakers is aligned with that used by younger and middle-aged educated speakers of LA.

The results for place of residence in the PA synchronic dataset confirm the earlier hypothesis that speakers outside the camps, by virtue of their presumed stronger ties to the Lebanese community, are more exposed to Beirut Arabic speech patterns, so they favour the

socially desirable variant *ma* over the stigmatized variants *ma-š* and *-š*. On the other hand, Palestinian camp dwellers, who have multiplex social networks with other Palestinians in the camps, use *ma-š* and *-š* more frequently than non-camp-dwellers. Additionally, it is possible that *ma-š* and *-š* are favoured among camp-dwellers because they may have covert prestige, as both of these variants are associated mainly with rural PA varieties (Mousa 2019), and many Palestinians in the camps are originally from rural areas in Palestine¹¹⁶ (Bocco 2009:239; see also Amiri 2016; Doraï 2010). For *miš*, we observe that it is used at the same rate (5%) by camp-dwellers and non-camp-dwellers, which further suggests that it is not socially diagnostic.

In terms of topic of discourse, *ma* in LA occurs at a slightly higher rate in soapbox speech and miscellaneous topics than in narratives, while *ma-š* and *-š* are categorically disfavoured in soapbox speech, as conjectured. Instead, *ma-š* is favoured in narratives, while *-š* occurs mainly in miscellaneous topics. The fact that *ma* is used more in soapbox speech (and miscellaneous topics) while *ma-š* and *-š* are avoided in careful speech bolsters the inference that variant choice in verbal negation in LA is stylistically sensitive. However, this does not seem to be the case for *miš*, which is distributed almost equally across different discourse topics.

In PA, *ma* occurs at similar rates in soapbox speech, narratives and miscellaneous topics, while *ma-š* occurs mostly in narratives. However, what is surprising is that *-š* in PA occurs mostly in soapbox speech. Closer inspection of the data revealed that out of the 24 tokens of *-š*

¹¹⁶ Bocco (2009:239) explains that most camp residents are Palestinians of rural origins (i.e., peasants and Bedouins). In the present study, several camp residents stated that they are originally from different villages in Palestine, such as Sa'sa', Kweikat, Al-Kabri, Nahf, Deir el-Qasi and Amqa. As such, the possibility of dialect levelling taking place in the camps cannot be ruled out. Fityan (1981:22) describes the situation in Al-Jalil camp in Baalbek (located in the north of Lebanon), explaining that most camp residents have retained Palestinian linguistic features from their areas of origin, while some have acquired new speech features from inside and outside their community. Accordingly, "the linguistic development of this camp has seemed to go through internal camp dialect leveling, in some instances preserving Palestinian speech features and in others incorporating Lebanese features, but variation among different groups of Palestinians, often continuations of old distinctions, persist" (Hennessey 2011:38). Similar to Al-Jalil camp, it is possible that some dialect levelling had taken place in the refugee camps targeted in this study; however, the possibility of dialect levelling in the camps cannot be addressed in any detail because that would require historical evidence that is currently inaccessible.

used in soapbox speech, 23 tokens were produced by (mainly less educated older and middle-aged) males, who spoke about politics in relation to the 1948 Arab–Israeli War. Accordingly, even though *-š* occurred more frequently than expected in soapbox speech, this is due to a hidden sex (and age) effect on the choice of this variant. For *miš* in PA, the results are similar to those in LA, as this variant does not seem to be stylistically sensitive.

Diachronic datasets

Table 5.4 shows the results of the multivariate analysis of social and stylistic predictors of variant choice in the POHA. Note that because the variant *a-š* occurred only once in the POHA, it was excluded from the analysis.

Table 5.4: Contribution of social and stylistic factors to variant selection in the POHA
(Note: shaded areas indicate statistically significant factor groups)

ma					ma-š				-š				miš				
Sex/education/ residence	Input prob.		0.486		Input prob.		0.144		Input prob.		0.0338		Input prob.		0.00486		
	Total N		4014		Total N		3807 ¹¹⁷		Total N		3807		Total N		4014		
	Deviance		3618.616		Deviance		2833.929		Deviance		2002.852		Deviance		371.693		
	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	
	M/PS/outside	.606	303	.716	.647	-.810	278	.169	.308	-.881	278	.058	.293	.886	303	.076	.708
	F/non-PS/inside	.187	971	.654	.547	.004	927	.198	.501	-.001	927	.142	.500	-1.437	971	.021	.192
	M/non-PS/inside	.080	1033	.643	.520	.124	999	.230	.531	-.033	999	.114	.492	-.273	1033	.024	.432
	F/non-PS/outside	-.279	1031	.573	.431	-.140	957	.220	.465	.642	957	.167	.655	.636	1031	.067	.654
	M/non-PS/outside	-.594	676	.509	.356	.822	646	.345	.695	.273	646	.128	.568	.188	676	.039	.547
	Topic																
Soapbox speech	1.013	138	.841	.734	-.580	131	.099	.359	-.991	131	.046	.271	-1.464	138	.022	.188	
Miscellaneous	-.454	2884	.614	.388	.287	2747	.234	.571	.471	2747	.131	.616	.547	2884	.038	.633	
Narrative	-.559	992	.568	.364	.293	929	.257	.573	.520	929	.151	.627	.917	992	.050	.714	

¹¹⁷ There were 207 tokens including future tense verbs/participles that were excluded from the analysis (4014 – 207 = 3807) because of a K.O. value.

In the POHA, *ma* is favoured the most by males with post-secondary education, all of whom live outside the camps (FW=.647). This is consistent with the earlier hypothesis that this group of speakers, by virtue of their higher level of education, will favour the socially unmarked form *ma*. The 19 females inside the camps with primary or secondary education weakly favour *ma* (FW=.547), followed by male camp-dwellers (FW=.520). However, the 21 females outside the camps weakly disfavour *ma* (FW=.431), and male non-camp-dwellers also disfavour this variant (FW=.356). This is somewhat surprising because we assumed that female non-camp-dwellers will favour *ma* as the default variant. This was not the case possibly because these females favour *-š* more than other speakers, which I will return to later.

In terms of *ma-š*, we also observe some differences across speaker groups. This variant is highly favoured by less educated males who live outside the camps (FW=.695), followed by male camp-dwellers (FW=.531). Female camp-dwellers neither favour nor disfavour *ma-š* (FW=.501), while female non-camp-dwellers weakly disfavour this variant (FW=.465). Unlike the case for *ma*, here we see that post-secondary educated males who live outside the camps highly disfavour *ma-š* (FW=.308). These findings are in line with the earlier hypothesis that the stigmatized variant *ma-š* will be disfavoured by educated males in the POHA, while it will be favoured by less educated speakers, especially males.

The variant *-š* is highly favoured by females who live outside the camps (FW=.655), followed by male non-camp-dwellers (FW=.568). For camp-dwellers, females (FW=.500) and males (FW=.492) have a neutral effect on *-š*. As speculated, males with post-secondary education highly disfavour *-š* (FW=.293). But the surprising finding concerns female non-camp-dwellers who favour *-š*. As far as is known, it does not seem that the 21 women in the POHA who live outside the camps have any distinctive personal circumstances that can explain why

they favour -š. In terms of education, only 5 of these women attended school, 4 of whom have secondary education. However, the 5 women who attended school used -š at a rate of 16% (N=41/260), similar to the 16 women who never attended school, and who used -š at an almost equivalent rate of 15% (N=119/773). Additionally, the urban-rural divide may not explain why female non-camp-dwellers favour -š. Out of the 21 women, 10 are from urban areas, and they used -š at a rate of 17% (N=83/478), while 11 are from rural areas, and they used -š at a rate of 14% (N=77/555).

Education and place of origin may not explain the women's use of -š, but there may be other possible explanations for their linguistic behaviour. Out of the 21 women living outside the camps, 16 left Palestine when they were 17 years old or above (their ages ranged between 17 and 35 years old, while the remaining 5 women were between 11 and 15 years old when they left). This means that they were around, or beyond, the stage of vernacular stabilization (see Chapter 3) when they arrived in Lebanon, so it is possible that they have acquired negation with -š in pre-war Palestine. In addition, it is possible that these older women have been largely isolated from urban vectors that disseminate socially unmarked or supralocal forms. Romaine (2003:108) explains that socially mobile persons tend to increase their use of the (local) standard more than others. Al-Wer (2002:50) also highlights the role of education in linguistic variation and change in the Arabic-speaking world, arguing that it is a prerequisite for mobility and contact.¹¹⁸ Based on that, speakers with lower educational levels, especially those who are unemployed, have limited outside contact, so they are thought to be the most conservative linguistically (Al-Wer 2002:50). The older women in the POHA not only had lower levels of education, but the

¹¹⁸ Al-Wer (2002:42) states that in the Arabic-speaking world, "access to education, especially at the higher level, and often even beyond primary schooling, involves significant alterations to individuals' socialization patterns. It involves leaving one's home town, changes in familial links, expansion in social contacts, interaction with speakers of other dialects, exposure to different social values, shifting of one's loyalties and attachments to various social groups, changes in priorities and ambitions, etc."

majority of them (18 out of the 21 female non-camp-dwellers) also did not work, so perhaps they had traditional roles as housewives. As a result, these women may have been less socially mobile than men, and may have had social networks limited to relatives (rather than including workmates or friends) (see Villena-Ponsoda 2005:326), which could explain their use of conservative features like *-š*.

Turning to *miš*, we see that the factor group of sex/education/residence was selected as significant. *Miš* is favoured by educated males (FW=.708), followed by female non-camp-dwellers (FW=.654). Less educated males who live outside the camps weakly favour this variant (FW=.547), while male camp-dwellers weakly disfavour it (FW=.432) and female camp-dwellers strongly disfavour *miš* (FW=.192). To my knowledge, there is no prestige (or stigma) attached to *miš* in spoken (Levantine) Arabic, so it is not clear why educated males and female non-camp-dwellers favour this variant as opposed to other groups of speakers. However, these findings from the POHA contrast with the findings from the synchronic datasets, where *miš* was not socially diagnostic, as it was distributed similarly across speaker groups.

In terms of the effect of discourse topic on variant choice in the POHA, we observe in Table 5.4 that all variants seem to be stylistically constrained. As hypothesized, *ma* is only favoured in soapbox speech (FW=.734), while *ma-š* and *-š* are disfavoured in this context. Instead, these two variants are mainly favoured in narratives and in miscellaneous topics. Unlike in the synchronic datasets, in the POHA, *miš* is stylistically constrained, as this variant is favoured in narratives (FW=.714), followed by miscellaneous topics (FW=.633), while it is strongly disfavoured in soapbox speech (FW=.188).

Turning to the LPTC, Table 5.5 shows the distribution of *ma*, *ma-š*, *-š*, *miš* and *a-š* according to social and stylistic factors in this dataset (the variants *a-* and *la-š* are excluded from

the distributional analyses because they rarely occurred in the LPTC).

Table 5.5: Overall distribution of variants according to social and stylistic factors in the LPTC

	ma		ma-š		-š		miš		a-š	
	%	N	%	N	%	N	%	N	%	N
Total N	80	967/1210	4	47/1210	1	14/1210	10	121/1210	4.7	57/1210
Sex										
Male	80	751/935	5	46/935	1.5	14/935	9	87/935	3.5	33/935
Female	78.5	216/275	0.5	1/275	0	0/275	12	34/275	9	24/275
Age Group										
Old	69	100/145	11	16/145	1	1/145	10	15/145	8	11/145
Middle-aged	88	504/570	1.5	9/570	0.5	2/570	7	42/570	2	11/570
Young	73	363/495	4	22/495	2	11/495	13	64/495	7	35/495
Social status										
Higher status	93	452/486	0	0/486	0	0/486	7	34/486	0	0/486
Middle status	80	245/305	3	8/305	2	6/305	12	36/305	3	9/305
Lower status	64	270/419	9	39/419	2	8/419	12	51/419	11.5	48/419
Topic										
Soapbox speech	94	80/85	0	0/85	0	0/85	6	5/85	0	0/85
Narrative	64	7/11	27	3/11	0	0/11	9	1/11	0	0/11
Miscellaneous	79	880/1114	4	44/1114	1	14/1114	10	115/1114	5	57/1114

In the LPTC, there is no evidence of sex-differentiation in the use of *ma*. However, it seems that the sex of characters in the LPTC affects the choice of *ma-š* and *-š*, because as hypothesized, these variants are used mainly by males, while only one (lower status) female used *ma-š*. The results for *miš* are similar to those of *ma*, as *miš* is also used at generally similar rates by males and females. The surprising results concern the effect of sex on the selection of *a-š*, which is favoured by females. However, closer inspection of the data revealed that only one female produced the 24 tokens of *a-š*, and this female is Lisa from the 1977 play *Bint El Jabal* ‘Mountain Girl’. As mentioned before, Lisa is a lower status character who speaks a “vulgar”

variety of LA and tries to learn “proper” LA from her mentor, linguist Henry Adib. Accordingly, there is a hidden status effect that explains why Lisa is the only female in the LPTC who used *a-š*. It is also important to remember that the LPTC is a fictional corpus, so this source should be used cautiously in drawing comparisons with the modern spoken LA data source. As such, the playwright may be deliberately manipulating linguistic features for parodic or didactic purposes (see Ayres-Bennett 2004:31, 74, 81), so there might be an element of caricature or stereotyping in the presentation of non-standard features (Ayres-Bennett 2004:74).¹¹⁹

In terms of the age of characters in the LPTC, *ma* is used mostly by middle-aged characters (88%), followed by younger characters (73%) and then older ones (69%). It was anticipated that *ma* would be used mostly by younger characters, but this was not the case possibly because in younger speech, as well as in older speech, *ma-š*, *-š* and *a-š* compete (see below), while middle-aged characters used these variants the least. In terms of social status, *ma* is used most frequently by higher status characters (93%), followed by middle-status characters (80%) and then lower status characters (64%). This confirms the hypothesis that because *ma* is the socially unmarked variant, it will be favoured mostly by higher status characters.

As predicted, older speakers use *ma-š* at a higher rate than the other age groups (11%). However, as mentioned above, younger speakers use *ma-š* (4%) more than middle-aged speakers (1.5%). Additionally, younger speakers use *-š* more than the other age groups, and they use *a-š* at a rate similar to older speakers. These findings possibly derive from the fact that age interacts with social status in the LPTC. In the plays, there are many lower status younger characters (out of the 81 lower status characters, 45 are young). Out of the 22 tokens of *ma-š* produced by younger speakers, 16 were produced by those with a lower social status and 6 tokens were

¹¹⁹ Ayres-Bennett (2004:31) explains that linguistic features may be chosen by playwrights as being so strongly non-standard that there was no doubt in the audience’s mind that the speech was intended to be ridiculous.

produced by middle-status younger characters. Additionally, 7 of the 11 tokens of *-š* produced by younger characters were actually produced by lower status characters, while the remaining four were produced by middle-status characters. Similarly, out of the 35 tokens of *a-š* produced by younger characters, 30 tokens were produced by lower status characters and the rest were produced by middle-status characters. Among the 15 older characters in the LPTC, 7 are higher status characters, while the remaining are middle-status or lower status characters. Only these middle-status or lower status speakers used *ma-š*, *-š* and *a-š*, with the latter being traced back to only one older (lower status) male, who portrayed Lisa's father in *Bint El Jabal* 'Mountain Girl', and who worked as a dustman. This reflects the findings for social status, where lower status characters favoured *ma-š* and *a-š*, while *-š* was distributed equally across middle-status and lower status characters. On the other hand, higher status characters never used *ma-š*, *-š* or *a-š* (N=0/486), using mainly *ma* (N=452/486) and to a lesser extent *miš* (N=34/486) for verbal negation. This shows that the socially marked forms in the LPTC were exclusively used by middle-status and lower status characters. Thus, while the social status of *ma-š* and *-š* (and *a-š*) has only been dealt with anecdotally or cursorily in the relevant literature (see e.g., Khairallah 2014; Wilmsen 2014), the distributional analysis in Table 5.5 presents quantitative evidence which shows that *ma-š*, *-š*, and *a-š* are socially diagnostic. Accordingly, even though the LPTC is fictional in nature, we can still leverage what variation there is in this fictional source to elucidate the social status of the variants of verbal negation. Examining the distribution of variants according to extra-linguistic constraints suggests that the LPTC seems to capture aspects of variation that we find in natural language data (where *ma-š* and *-š* are socially diagnostic in the synchronic datasets, at least in terms of sex and age/educational level, as well as place of residence for Palestinians). By contrast, in terms of *miš*, further evidence that it does not seem to

be socially diagnostic in the LPTC comes from the fact that it is used at generally similar rates across different age groups, as well as across higher, middle- and lower status characters. These findings highlight the utility of complementary data sources (here the LPTC and the LA synchronic dataset) in showing that *ma* (and *miš*) is unmarked, while *ma-š* and *-š* (as well as *a-š*) are marked forms (see Ayres-Bennett 2000:329, 342, 2004:31, 59, 226).¹²⁰

In terms of topic of discourse, *ma* was used mostly in soapbox speech, while *ma-š*, *-š* and *a-š* were categorically disfavoured in this context. Instead, *ma-š* was used in narratives and in miscellaneous topics, while *-š* and *a-š* only occurred in the latter. These findings reflect the results from the other datasets, where soapbox speech generally inhibited or disfavoured the use of *ma-š* and *-š*. *Miš* occurred at similar rates across discourse topics, though it occurred slightly less frequently in soapbox speech, possibly because mainly *ma* is used in this context.

Summarizing, the results for the extralinguistic factors affecting variant choice in verbal negation show that while *miš* was generally distributed at similar rates across different speaker groups and discourse topics (except in the POHA), *ma*, *ma-š* and *-š* (and *a-š* in the LPTC) seem to be socially and stylistically constrained in all the datasets. As we saw in the synchronic datasets, speaker sex, age/educational level and place of residence for Palestinians seem to play a key role in variant choice. The effect of age/educational level on variant selection in particular shows evidence of *ma* becoming the predominant form in apparent time. However, it is not only age, but also educational level that seems to play a decisive role in variant choice. As we saw previously, Palestinian speakers educated to post-secondary level, both younger and middle-aged speakers, are more likely than those with lower levels of education to use *ma*, while generally avoiding *ma-š* and the negative enclitic *-š*. The POHA displayed a similar trend to the PA

¹²⁰ According to Ayres-Bennett (2000:342), it is important to seek confirmation of the status of a variant by comparing sources of different kinds and seeking convergence.

synchronic dataset, where the only males educated to post-secondary level also favoured *ma* and disfavoured the socially marked variants, *ma-š* and *-š*. These findings suggest that the verbal negation system in PA as spoken in Beirut constitutes a classic case of dialect levelling, where socially/regionally marked forms are being reduced in favour of supralocal variants with broader socio-spatial currency (see e.g., Britain 2010a:151–52, 2010b; Kerswill 2003; Trudgill 1986:98), especially in the speech of younger, educated Palestinians.¹²¹ However, to the extent that LA has influenced the verbal negation system in PA, this should be manifested not only in terms of variant rates and social factors, but also in terms of the linguistic conditioning of variant selection, which I turn to in the next section.

5.3.3 Linguistic predictors of variant choice

To examine whether PA is converging on LA in terms of verbal negation, it is important to compare patterns of the internal conditioning of variant selection across the relevant synchronic and diachronic benchmarks. In this section, I first conduct apparent-time comparisons across age groups in the LA and PA synchronic datasets. If young Palestinians are diverging from their elders, and they are simultaneously behaving like (young) Lebanese speakers, this would constitute compelling evidence that we may be dealing with an externally-motivated change in PA due to the influence of LA. I also conduct real-time comparisons across the diachronic data sources and the contemporary varieties. If it is additionally found that contact PA (young speech) features constraint hierarchies different from those of the POHA (which approximates to pre-contact PA), while simultaneously exhibiting parallels with those of LA (young speech), this would strengthen the inference of contact-induced change in contemporary PA.

¹²¹ Britain (2010b:193) explains that dialect levelling refers “to the process by which, as a result of mobility and dialect contact, linguistic variants with a wider socio-spatial currency become more widespread at the expense of more localised forms.” So in the process of levelling, one (majority) variant – here *ma* – “emerges victorious from the mixing of [...] different dialect variants of the same variable” (Britain 2010b:194–95).

LA and PA synchronic datasets

Table 5.6 presents the results of the apparent-time comparisons across the LA and PA synchronic datasets.

Table 5.6: Overall distribution of variants according to linguistic factors across different age groups in the LA and PA synchronic datasets¹²²

Lebanese Speakers									Palestinian Speakers										
Older ¹²³					Middle-aged		Younger		Older				Middle-aged				Younger ¹²⁴		
	ma	ma-š	-š	miš	ma	miš	ma	miš	ma	ma-š	-š	miš	ma	ma-š	-š	miš	ma	-š	miš
	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%
Total %	67	22	6	4	94	6	95	5	74	9	12	5	84	4.5	5	6.5	95	2	3
Verb																			
Imperfect (Present)	75	20	2.5	2.5	98	2	98	2	77	8	13	2	86	5	6	3	97	2	1
Perfect (Past)	32	54	14	0	100	0	96	4	78	22	0	0	87	9	0	4	100	0	0
Future or Participle	60	0	0	40	56	44	58	42	40	0	0	60	49	0	0	51	70	0	30
Pseudo-verb	91	2	5	0	98	2	100	0	74	4	22	0	88	3	9	0	94	5	0
Initial consonant																			
Labial	85	7.5	4.5	1.5	97	3	97.5	2.5	76	5	17	2	85	4	8	3	96	3	1
Non-labial	48	38	8	6	89	11	90	10	71	16	4	9	82	6	0	12	93	0	6
Intervening material																			
None	83	5	8	3	94.5	5.5	96	4	85	4	7	4	89	2	4	5	97	1	2
Present	21	73	0	6	88	12	83	17	32	29	31	8	42	25	19	14	79	5	14

¹²² The tables including apparent-time comparisons with the raw numbers in the LA and PA synchronic datasets are found in Appendix C.¹²³ There is also one token of *a-š* in older LA speech, which occurred with a pseudo-verb.¹²⁴ There is only one token of *ma-š* in young PA speech, so it has been excluded from the distributional analysis of linguistic factors.

Because evidence of convergence of PA on LA in terms of verbal negation most likely resides in the speech of young speakers, I first compare the patterns of variant conditioning across the speech of young LA and young PA speakers. Both cohorts use mainly *ma* for verbal negation, which is used at the same rate in young LA and young PA speech (95%). They also use *miš* at similar rates. A primordial issue, however, is whether young Palestinians reproduce the same linguistic constraints on variant choice as those operating in young LA speech. As we see in Table 5.6, the pattern of favouring and disfavouring effects conditioning *ma* is similar across both cohorts. This variant is favoured with all verb forms listed, but it occurs at a considerably lower rate with future tense verbs or participles, as *miš* is favoured in this context in both varieties. In both young LA and PA speech, *ma* also occurs more with labial-initial verbs and in the absence of intervening material between this negative element and the negated verb.

Regarding *miš*, for all factor groups, the constraint hierarchies are parallel across young LA and young PA speech. As mentioned above, *miš* mainly occurs with participles and future tense verbs across the two comparison groups, consistent with previous claims in the literature that *miš* is used to negate these verbs (see e.g., Brustad 2000:289; Farwanah 2021:171–72). *Miš* also occurs more frequently with non-labial-initial verbs, as well as in the presence of intervening material. In the case of young LA and PA speakers, the intervening material between *miš* and the negated verb was either the progressive marker *šam*, as we saw in example (25), or a discourse marker, e.g.:

- (35) w yḍallon integrated bi l-work environment,
 and remain.3p integrated in the-work environment
 miš **inno** yetlašo min šiył-un.
 NEG DISCOURSE MARKER leave.3p from work-their
 ‘And they remain integrated in the work environment, not leave their jobs.’
 [13/P/F/28/10:57]

- (36) *tfadḍayte, miš inno štayalte.*
 get.free.2sf NEG DISCOURSE MARKER worked.2sf
 Intended: ‘You had free time, you did not work.’

[17/L/F/34/18:52]

The examples above also show *miš* being used to negate present tense verbs, as in *yeṭlaṣo* ‘(they) leave’ in (35), as well as past tense verbs, as in *štayalte* ‘(you) worked’ in (36). In these two cases, as well as in several other tokens where *miš* was used to negate present tense or past tense verbs, this variant occurred in instances of what Al-Sayyed and Wilmsen (2017:159–60) call contrastive negation, which involves making a contrast between two sets of conditions.

The comparison of the patterns of linguistic conditioning of *ma* and *miš* across younger LA and younger PA speakers show that there are striking similarities between the two groups. However, I cannot be sure that there has been (contact-induced structural) change in PA unless I look at the conditioning of variant choice in older LA and older PA speakers. To identify change, we need clear evidence that the linguistic conditioning of variant choice differs across older LA and older PA speech. In addition, identifying change rests on the demonstration that patterns of variant conditioning in older PA speech differ from those in younger speech (i.e., change would be detected if younger speakers are diverging from their elders).

By comparing the patterns of conditioning of variant selection across older LA and older PA speech, we see that both cohorts generally behave similarly, with a few minor differences. In older LA speech, *ma* occurs most frequently with pseudo-verbs, followed by present tense verbs, while it occurs less with future tense verbs or participles because *miš* is favoured in this context. *Ma* in older LA speech occurs the least with past tense verbs, where mainly *ma-š* competes. In older PA speech, *ma* occurs at similar rates across past tense and present tense verbs, as well as pseudo-verbs, while it occurs less with future tense verbs or participles, as we saw in LA. In terms of other factor groups, the constraint hierarchies and direction of effects are similar across

the two comparison groups, where *ma* occurs more with labial-initial verbs and without any intervening material between this negative element and the verb.

For all factor groups, the direction of effects conditioning *ma-š* and *miš* are parallel across older LA and PA speech, but the main difference between the two cohorts concerns the conditioning of *-š*. In older PA speech, *-š* is favoured with pseudo-verbs the most, followed by present tense verbs. This variant is also favoured with labial-initial verbs and with intervening material (the only possible intervening material between the negated verb and *-š* includes one or more clitics). The results for PA confirm the majority of claims in the literature that *-š* is disfavoured with past tense and non-labial-initial verbs in this variety (Abu-Haidar 1979; Blau 1960; Hoyt 2006; Lucas 2007, 2010; Obler 1975). Nevertheless, in older LA speech, *-š* is favoured with past tense and non-labial-initial verbs (mainly with the copula *kan* ‘he/it was’), and it only occurs in the absence of intervening material (i.e., *kan-š* ‘he/it wasn’t’). This is consistent with Thackston’s (1996) claim that in LA, *-š* negates non-labial-initial verbs in the past tense. Apart from the differences in the conditioning of *-š*, there are general parallels across older LA and older PA speech. Based on these affinities, we cannot rule out the possibility of coincidental parallels.

After comparing the patterns of linguistic conditioning of variant selection across older LA and older PA speakers and finding similarities, we need to examine whether change has taken place from older to younger speakers in the same variety. Within LA, change does not seem to have taken place in the conditioning of variant selection, because both *ma* and *miš* are constrained similarly across older, middle-aged and younger speakers. In terms of variant rates, the only difference is that older speakers use *ma-š* and *-š*, while their younger counterparts do not. However, this is not taken as evidence of change in LA (in terms of variant distributions)

because we are comparing educated younger and middle-aged LA speakers from Beirut with less educated older speakers mainly from Mount Lebanon.

Within PA, we also see structural correspondences across the three age groups. With respect to *ma*, we observe that it has a tendency to occur at similar rates with different verb forms, except with future tense verbs or participles. As we saw in LA, *ma* in PA also occurs more frequently with labial-initial verbs, and without intervening material between this variant and the negated verb. The constraint hierarchies and direction of effects conditioning *-š* and *miš* are also generally parallel across the three age groups in PA. These findings don't point in favour of contact-induced structural change in the verbal negation system in PA as spoken in Beirut (at least as far as *ma*, *-š* and *miš* are concerned). To put it in another way, even though we saw before that younger PA speakers behave like younger LA speakers, we cannot infer structural convergence of PA on LA with respect to verbal negation because younger Palestinians behave like their older counterparts in terms of the conditioning of variant choice.

Even though there does not seem to be contact-induced structural change in the verbal negation system in PA as spoken in Beirut, there does seem to be a change in progress in this variety in terms of variant rates, as *ma-š* and *-š* decrease progressively from older speakers to younger ones. This change seems to be determined mainly by education, because recall (from Table 5.3) that *ma-š* and *-š* are mainly used by (less educated) older speakers and less educated middle-aged speakers, while these two variants considerably decrease in the speech of educated middle-aged and young Palestinians. This reduction in socially marked variants in the speech of younger, educated Palestinians may have been prompted by their regularity of contact with Beirut LA speakers via post-secondary institutions.

It is worth exploring who among the young Palestinians in the present study uses *ma-š* and *-š*. The only token of *ma-š* was produced by a 33-year-old woman who lives outside the camps. This speaker also used *-š* a few times (N=3/74). During the interview, she explained that before pursuing a university degree, she studied at a vocational centre established by UNRWA, which is open for Palestinians because access to Lebanese government vocational training schools is restricted exclusively to Lebanese nationals (Suleiman 2010:11). Given her education background, the speech of this female speaker may have been slightly influenced by the Palestinian network she was part of in the vocational centre.

The other tokens of *-š* were produced by a 20-year-old female (N=3/80) and a 26-year-old male (N=3/61). The 20-year-old female lives outside the camps, and she was enrolled at a university in Lebanon at the time of the interview. She explained that she is a member of a Palestinian cultural club at the university, and the members of this club are mainly (though not exclusively) Palestinians, so again, we suppose that ties with other Palestinians in this club may have influenced the speech of this young woman. The young male who used *-š* is also a university student, but he lives in Shatila camp. He is actually the same speaker who never used word-medial *imala* (speaker [51]; see Chapter 4). One possible explanation for why he never used *ma-š* and rarely used *-š* comes from the following excerpt:

- (37) *Masalan niḥna bas kunna bi l-madrise bašdna, bi s-sanawiyye, ka:n fi na:s yšayru:na min l-burj yʔulu:lna “into ya libna:niyye ya taṇtat”, laʔanno niḥna ma mneḥke nafs lahjeton. Ahl l-burj aylabun min Kweika:t w Taršiḥa w l- Ka:bri [...] fa kullun ki:f ḥakyun? Henne falla:ḥi:n [...] niḥna ma kunna mneḥke heke, niḥna mneḥke- laʔanno šati:la ašlan niḥna ʔaʔrab ša Beirut, yašne henne bʔalb d-da:hye [...] taʔri:ban niḥna madxal Beirut, fa la:zem- la:zem yku:n ahehna tʔassaro w niḥna tʔassarna.*
[51/P/M/26/27:36]

‘For example, when we were still in school, in high school, there were people from El-Burj (Burj El-Barajneh camp in Beirut) who used to bully us, they used to say “you

(residents of Shatila camp) are Lebanese, you are *tantat*”,¹²⁵ because we don’t speak the same dialect as theirs. The majority of people in El-Burj are from Kweikat, Tarshiha and Al-Kabri (villages in Palestine/Israel) [...] so how do all of them speak? They are (originally) peasants (they speak the rural dialect) [...] But we didn’t speak like that, we speak- because Shatila is closer to Beirut, but they (people in Burj El Barajneh) are in Dahieh (southern suburb of Beirut) [...] we are almost at the outskirts of Beirut, so it is inevitable- it is inevitable that our parents have been influenced and we have been influenced.’

In the excerpt above, speaker [51] admits that his speech and that of his parents have been influenced by contact with metropolitan norms of LA (Beirut Arabic), which can explain why he uses mainly *ma* for verbal negation. This is in contrast to his linguistic behaviour in terms of word-medial *imala*, which he avoids. We can countenance the possibility that non-raising of the word-medial vowel and *ma-š* and *-š* index different social meanings. Non-raising is an option in both LA and PA (i.e., the realization of the word-medial vowel is variable), and it may have covert prestige for speaker [51]. By contrast, for this speaker, as well as for other educated speakers of PA (and LA), *ma-š* and *-š* can index rurality (see e.g., Wilmsen 2014) and “uncultivated” speech, which are at variance with the social aspirations of the educated younger generation (see Fityan 1981), who may wish to ally themselves with urban modernity (see Britain 2010b:202; Kerswill 2002:198, 2003:224, 226, 230).

To sum up, in this section, I conducted apparent-time comparisons across the LA and PA synchronic datasets in order to examine whether PA is converging on LA in terms of verbal negation. The comparison of the patterns of variant conditioning across younger Palestinian and Lebanese speakers showed that both cohorts behave similarly in the conditioning of *ma* and *miš*. However, I could not infer convergence of PA on LA because it was also found that in each variety, different age groups generally display similar patterns of conditioning of variant choice.

¹²⁵ *Tantat* is the plural of *tant* ‘aunt’ (from French *tante*), which is used to refer to someone who is not tough (Al-Wer 2007b:67). The term can imply effeminacy in a boy, without reference to sexuality (Al-Wer 2007b:74).

In addition, the fact that even older PA and older LA speakers share general structural correspondences, except in terms of the conditioning of *-š*, is further evidence that the structural similarities we saw between LA and PA may not be the result of convergence, but the result of possible coincidental (structural) parallels. The only change in progress we could detect in PA concerned the decrease in the use of *ma-š* and *-š* among educated middle-aged and younger speakers, which may have been induced by regular contact with metropolitan norms of LA. In the next section, I explore the possibility of a coincidence site in terms of verbal negation between LA and PA by comparing the synchronic datasets with the diachronic data sources. If the underlying grammar associated with verbal negation is similar across the POHA and the PA synchronic dataset, this presents further evidence that there is no contact-induced structural change in this variable system in PA as spoken in Beirut.

The POHA and young PA speech

In this section, I examine if there is a change in the patterns of linguistic conditioning of variant choice in verbal negation across the POHA and younger, educated speakers in the PA synchronic dataset. Table 5.7 compares between these two datasets.

Table 5.7: Contribution of linguistic factors to variant selection in the POHA and young PA speech
(Notes: shaded areas indicate statistically significant factor groups; K.O.= invariant context)¹²⁶

POHA								Younger Palestinians ¹²⁷							
ma			ma-š		-š		miš		ma		-š		miš		
Verb	Input prob.	0.486	Input prob.	0.144	Input prob.	0.0338	Input prob.	0.00486	%	Total N	%	Total N	%	Total N	
	Total N	4014	Total N	3807 ¹²⁸	Total N	3807 ¹²⁹	Total N	4014	95	440/465	2	9/465	3	15/465	
	Deviance	3618.616	Deviance	2833.929	Deviance	2002.852	Deviance	371.693							
	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	%	N	%	N	%	N	
	Perfect (Past)	1.265	.780	-.466	.386	-.794	.311	-1.959	.124	100	105/105	0	0/105	0	0/105
	Imperfect (Present)	.503	.623	-.140	.465	.547	.633	-1.924	.127	97	231/238	2	5/238	1	2/238
	Pseudo-verb	.286	.571	.606	.647	.247	.561	-2.121	.107	94	74/79	5	4/79	0	0/79
	Future or Participle	-2.054	.114	K.O.	K.O.	K.O.	K.O.	6.004	.998	70	30/43	0	0/43	30	13/43
	Initial consonant														
	Non-labial	.163	.541	1.043	.739	-1.736	.150	-.060	.485	93	180/193	0	0/193	6	12/193
	Labial	-.163	.459	-1.043	.261	1.736	.850	.060	.515	96	260/272	3	9/272	1	3/272
	Intervening material														
None	1.442	.809	-1.091	.251	-.912	.287	.309	.577	97	395/408	1	6/408	2	7/408	
Present	-1.442	.191	1.091	.749	.912	.713	-.309	.423	79	45/57	5	3/57	14	8/57	

¹²⁶ The complete table for the contribution of linguistic factors to variant selection in the POHA (with Total N and proportion of application value) can be found in Appendix D.

¹²⁷ The results regarding the linguistic constraints on *ma-š* in young PA speech are not reported because, as mentioned previously, there is only one token of *ma-š* in young PA speech.

¹²⁸ There were 207 tokens (4014 – 207 = 3807) including future tense verbs/participles that were excluded from the analysis because of a K.O. value.

¹²⁹ There were 207 tokens (4014 – 207 = 3807) including future tense verbs/participles that were excluded from the analysis because of a K.O. value.

In Table 5.7, we find several parallels between speakers in the POHA and younger Palestinians in the synchronic dataset. For example, the two groups share the same constraint hierarchy and direction of effects conditioning *ma* in the verb factor group. A slight difference in the conditioning of *ma* relates to the effect of the initial consonant of the negated verb. Although in both datasets *ma* is distributed similarly across labial-initial and non-labial-initial verbs, in the POHA, *ma* is weakly favoured with the latter (FW=.541), while in young PA speech, *ma* occurs more frequently with labial-initial verbs. But in terms of intervening material between *ma* and the negated verb, in both datasets, this variant is favoured without any intervening material.

In terms of the variant *ma-š*, as mentioned previously, it was only used once in younger PA speech (N=1/465 or 0.2%), occurring with the pseudo-verb *fī* ‘there is’ (i.e., *ma fī-š*... ‘there isn’t...’). In the POHA, *ma-š* occurs much more frequently than in younger PA speech, at a rate of 22% (N=895/4015). As we see in the table above, the speakers in the POHA favour *ma-š* only with pseudo-verbs (FW=.647), while *ma-š* is categorically disfavoured with future tense verbs or participles. This is in line with Mousa’s (2019:100–101) claim that this variant cannot negate future tense verbs in PA. Older speakers in the POHA also strongly favour *ma-š* with non-labial-initial verbs. This is consistent with previous claims in the literature that *ma* (whether with or without *-š*) is retained with non-labial-initial verbs and can only be omitted with labial-initial verbs in PA (Blau 1960:198; Hoyt 2006:16). *Ma-š* in the POHA is also strongly favoured in the presence of intervening material, in line with our earlier hypothesis that intervening material between the first negator (*ma*) and the verb may favour a second negator (*-š*) to reinforce the first negative element.

Turning to *-š*, we find that the POHA and young PA speech do not share the same direction of effects associated with the verb factor group. In the POHA, *-š* is favoured the most

with present tense verbs (FW=.633), followed by pseudo-verbs (FW=.561), while the reverse direction of effects is observed in young PA speech. However, in both datasets, we can see that *-š* is disfavoured with past tense verbs, in line with previous claims in the literature (Abu-Haidar 1979; Blau 1960; Hoyt 2006; Lucas 2007, 2010; Obler 1975). In addition, *-š* in both datasets is categorically disfavoured with future tense verbs or participles. Consistent with previous claims in the literature, *-š* in the POHA and young PA speech is clearly favoured with labial-initial verbs (see e.g., Blau 1960; Hoyt 2006). The two comparison groups also share the same constraint hierarchy and direction of effects associated with the last factor group, where *-š* is favoured with intervening material (as mentioned previously, the intervening material between the negated verb and *-š* includes one or more clitics).

Moving to *miš*, we observe that in the POHA and young PA speech, *miš* is favoured with future tense verbs or participles. In terms of the initial consonant of the negated verb, the two groups do not share the same constraint hierarchy and direction of effects. In the POHA, the effect of this factor group on the use of *miš* is almost neutral, while in young PA speech, *miš* seems to show a preference for non-labial-initial verbs. Moreover, with respect to the effect of intervening material between *miš* and the negated verb, this variant in the POHA is weakly favoured without intervening material (FW=.577), while in young PA speech, it occurs more frequently in the presence of such material.

Notwithstanding the minor differences across the POHA and young PA speech, we observe broad parallels between the two datasets, especially in terms of the linguistic conditioning of *ma* and *-š*. This suggests that in terms of the underlying grammatical conditioning of variant selection in verbal negation, there does not seem to be a change in Beirut PA. However, it is important to note that one possible confound in the comparison between the

POHA and young PA speech is that we are comparing between the speech of mostly less educated speakers in the POHA (with the exception of six male speakers who live outside the camps) with post-secondary educated younger PA speakers. For that reason, I narrow the analysis down to educated speakers in the POHA and compare their speech patterns with those of younger, educated third-generation Palestinians in Table 5.8.

Table 5.8: Overall distribution of variants according to linguistic factors across post-secondary educated speakers in the POHA and younger Palestinians in the synchronic dataset

POHA (educated speakers only)									Younger Palestinians						
ma			ma-š		-š		miš		ma		-š		miš		
	%	Total N	%	Total N	%	Total N	%	Total N	%	Total N	%	Total N	%	Total N	
	72	217/303	15	47/303	5	16/303	8	23/303	95	440/465	2	9/465	3	15/465	
Verb	%	N			%	N	%	N	%	N	%	N	%	N	
	Perfect (Past)	79	105/133	21	28/133	0	0/133	0	0/133	100	105/105	0	0/105	0	0/105
	Imperfect (Present)	76	66/87	11	10/87	13	11/87	0	0/87	97	231/238	2	5/238	1	2/238
	Pseudo-verb	76	44/58	15	9/58	9	5/58	0	0/58	94	74/79	5	4/79	0	0/79
	Future or Participle	8	2/25	0	0/25	0	0/25	92	23/25	70	30/43	0	0/43	30	13/43
	Initial consonant														
	Non-labial	71	126/179	22	40/179	0	0/179	7	13/179	93	180/193	0	0/193	6	12/193
	Labial	73	91/124	6	7/124	13	16/124	8	10/124	96	260/272	3	9/272	1	3/272
	Intervening material														
	None	82	193/234	3	7/234	5	11/234	10	23/234	97	395/408	1	6/408	2	7/408
Present	35	24/69	58	40/69	7	5/69	0	0/69	79	45/57	5	3/57	14	8/57	

In Table 5.8, we observe several similarities in the patterns of linguistic conditioning of variant choice across the speech of educated speakers in the POHA and younger Palestinians in the synchronic dataset. For all factor groups, the constraint hierarchy and direction of effects conditioning *ma* are parallel across the two comparison groups. As mentioned above, *ma-š* is

only productive in the POHA in contrast to younger PA speech, and it occurs at a rate of 15% (N=47/303) in the speech of the six educated males. Table 5.8 shows that these educated speakers use *ma-š* most frequently with past tense verbs, followed by pseudo-verbs and then present tense verbs. As seen previously, *ma-š* seems to be favoured with non-labial-initial verbs and in the presence of intervening material in the POHA.

In terms of *-š*, even though the constraint hierarchy and direction of effects associated with the verb factor group are not parallel across the two groups, we can see that *-š* is mainly used with present tense verbs and pseudo-verbs, while it is categorically disfavoured with past tense and future tense verbs or participles. Both groups also use *-š* only with labial-initial verbs, as hypothesized. The constraint hierarchy and direction of effects associated with the last factor group are also parallel across educated speakers in the POHA and younger PA speakers, where *-š* is preferred in the presence of intervening material between the negated verb and this negative enclitic.

Across the two comparison groups, we observe that *miš* is mainly favoured with future tense verbs or participles. However, there are a few differences between the educated speakers in the POHA and younger Palestinians in terms of the effect of the remaining factor groups on *miš*. In the speech of older, educated Palestinians in the POHA, *miš* is distributed at similar rates across labial-initial and non-labial-initial verbs. On the other hand, *miš* in young PA speech is preferred with non-labial-initial verbs. In terms of the last factor group, in the speech of educated Palestinians in the POHA, *miš* only occurs in the absence of intervening material between this variant and the negated verb, while in young PA speech, the direction of effect is reversed. Closer inspection of the data revealed that in the POHA, the six educated males use *miš* only with participles, as in the examples below, where there is no material intervening between this

variant and the negated verb (this is in contrast with future tense verbs, where the intervening material would be a future tense marker like *rah* or *ha-*).

- (38) **miš** ʕa:ref ki:f bidd-e aħmel-hum.
 NEG know.1sm how FUT carry.1s-them
 ‘I don’t know how I will carry them.’

[POHA/03/11.12.2003/M/85/23:02]

- (39) fi mana:teq bi beit-na ana **miš** ra:yeh-l-a
 there.is areas in house-our I NEG gone.1sm-to-them
 b-ħaya:t-i.
 in-lifetime-my
 Intended: ‘There are areas in our house that I have never been to (describing a big house).’

[POHA/24/12.11.2003/M/72/36:24]

The comparison between educated Palestinians in the POHA and younger Palestinians in the synchronic dataset shows that there are general parallels in the patterns of linguistic conditioning of *ma*, *-š* and to a lesser extent, *miš*. This again suggests that there does not seem to be a change in the underlying linguistic constraints governing variant choice in the verbal negation system in Beirut PA. Nevertheless, we observe from Table 5.8 that there are some differences between the two comparison groups in terms of variant rates, where *ma* occurs at a rate of 72% in the speech of the six educated males in the POHA, while it increases to a rate of 95% in young PA speech. Conversely, *ma-š* and *-š* (and *miš*) decrease from older, educated speech in the POHA to younger PA speech. This bolsters the earlier inference that there is a change in terms of variant rates in the system of verbal negation in Beirut PA, where socially marked variants are being whittled away in real-time (and in apparent-time, as we saw in Table 5.6), while *ma* is the pre-eminent variant in the verbal negation system in this variety. As mentioned previously, this change seems to be motivated mainly by education.

In the next section, I compare the patterns of linguistic conditioning of variant choice across the LPTC and the LA synchronic dataset to check if there is no change in the verbal negation system in LA, as we saw in the apparent-time comparisons within the LA synchronic dataset in Table 5.6.

The LPTC and the LA synchronic dataset

Table 5.9 shows the overall distribution of variants according to linguistic factors in the LPTC and the LA synchronic dataset.

Table 5.9: Overall distribution of variants according to linguistic factors in the LPTC and the LA synchronic dataset

LPTC							Lebanese speakers							
Higher status			Lower status				Older				Middle-aged		Younger	
	ma	miš	ma	ma-š	-š	miš	ma	ma-š	-š	miš	ma	miš	ma	miš
	%	%	%	%	%	%	%	%	%	%	%	%	%	%
Total %	93	7	64	9	2	12	67	22	6	4	94	6	95	5
Verb														
Imperfect (Present)	96	4	64	5	2	6	75	20	2.5	2.5	98	2	98	2
Perfect (Past)	100	0	69	23	2	1	32	54	14	0	100	0	96	4
Future or Participle	44	56	30	2	0	68	60	0	0	40	56	44	58	42
Pseudo-verb	100	0	84	7	3	1	91	2	4.5	0	98	2	100	0
Initial consonant														
Labial	99.5	0.5	67	3	4	7	85	7.5	4.5	1.5	97	3	97.5	2.5
Non-labial	87	13	62	15	0	17	48	38	8	6	89	11	90	10
Intervening material														
None	97	3	73	4	2	13	83	5	8	3	94.5	5.5	96	4
Present	72	28	41	25	3	10	21	73	0	6	88	12	83	17

As mentioned previously, the LPTC is a diachronic surrogate of orality which is considered to be an approximation to LA, rather than a faithful facsimile of LA speech. For this reason, the comparison between the LPTC and the LA synchronic dataset in the table above is intended as a broad-brush, exploratory manoeuvre. In the table above, I compare the distribution of variants in the LPTC by social status with the age distributions in the LA synchronic dataset. Even though I coded for age in the LPTC, it should be noted that that age groups in the plays may not correspond to the age groups in the LA synchronic dataset because of the factor of social status. Several older characters (7 out of 15 older characters) in the LPTC belong to a higher social status (e.g., pasha, mayor, lawyer, etc.), and they do not use the socially marked variants (*ma-š* and *-š*), unlike the (less educated) older speakers in the LA synchronic dataset. On the other hand, the majority of younger characters in the LPTC belong to a lower or middle social status (101 out of 115 characters), as they portrayed occupations like waiter, newspaper seller, mechanic, butler, chef, firefighter, etc. These characters may not reflect the post-secondary educated young (and middle-aged) Lebanese speakers in the synchronic dataset, who avoid the two stigmatized variants, possibly because they are influenced by Beirut Arabic speech patterns and their higher level of education.

In Table 5.9, we can observe several similarities between higher status characters and post-secondary educated middle-aged and younger speakers. The first striking similarity is in terms of variant rates. As mentioned above, higher status characters and educated middle-aged and young Lebanese speakers do not use *ma-š* and *-š*. Instead, *ma* is the majority variant, where it ranges between 93% and 95% across the three comparison groups, while *miš* ranges between 5% and 7%. There are also broad parallels in terms of the patterns of linguistic conditioning of variant choice across the three comparison groups. In terms of the effect of the verb form on *ma*,

higher status characters, as well as middle-aged and younger educated Lebanese speakers, favour *ma* the least with future tense verbs or participles, as *miš* is mainly used in this context, as we saw in the POHA and the PA synchronic dataset. The three comparison groups also share the same constraint hierarchy and direction of effects associated with the initial consonant of the negated verb and intervening material between *ma* and the negated verb. Similarly, across the three comparison groups, *miš* occurs more frequently with non-labial-initial verbs, and in the presence of intervening material between this variant and the negated verb.

Moving to lower status characters in the LPTC and less educated, older speakers in the synchronic dataset, we also see some broad parallels between these two comparison groups. While *ma* is the majority variant in both groups, it ranges between 64% and 67%, so it occurs at a considerably lower rate than what we saw in the speech of higher status LPTC characters and educated middle-aged and younger speakers in the synchronic dataset. Lower status characters and older speakers in the contemporary variety also use *ma-š* and *-š*, though these two variants occur more frequently in the speech of older Lebanese speakers, especially *ma-š*. On the other hand, *miš* is more frequent among lower status characters in the LPTC than among older Lebanese speakers.

In terms of the effect of verb form on *ma*, lower status characters in the LPTC and older speakers in the synchronic dataset do not share a similar constraint hierarchy and direction of effects associated with this factor group. In lower status speech, *ma* occurs most frequently with pseudo-verbs, followed by past tense verbs and then present tense verbs. As we saw before, *ma* occurs the least with future tense verbs or participles. On the other hand, in older speech in the synchronic dataset, *ma* occurs mostly with pseudo-verbs, followed by present tense verbs, and then future tense verbs or participles, while it occurs the least with past tense verbs. However,

the two comparison groups share the same constraint hierarchy and direction of effects associated with the remaining factor groups, where *ma* occurs more frequently with labial-initial verbs and in the absence of intervening material between this variant and the negated verb.

Turning to *ma-š*, we find that lower status characters and older speakers show a preference for this variant with past tense verbs, while they disfavour it with future tense verbs or participles. In addition, the two comparison groups use *ma-š* more frequently with non-labial-initial verbs, as we saw in the POHA. The two comparison groups also share the same constraint hierarchy and direction of effects associated with the last factor group, where *ma-š* is preferred in the presence of intervening material, which was likewise observed in the POHA.

There are some differences between lower status characters and older speakers in terms of the patterns of linguistic conditioning of *-š*. In the speech of lower status characters, *-š* occurs at similar rates with pseudo-verbs, present tense verbs and past tense verbs, while it is categorically disfavoured with future tense verbs or participles. In older LA speech, *-š* is also avoided with future tense verbs or participles, but as we saw before, it shows a preference for past tense verbs, mainly the copula *ke:n* ‘he/ it was’. This is perhaps why *-š* is favoured with non-labial-initial verbs and it only occurs in the absence of intervening material between this negative enclitic and the negated verb. By contrast, lower status characters use *-š* only with labial-initial verbs, and they use it slightly more frequently in the presence of intervening material.

Apart from the differences between the two comparison groups in terms of the patterns of conditioning of *-š*, there are general parallels in terms of the constraint hierarchies and direction of effects conditioning *miš*. As we saw previously, this variant occurs most frequently with future tense verbs or participles and it is preferred with non-labial-initial verbs. A minor

difference between the two comparison groups concerns the effect of intervening material between *miš* and the negated verb, where this variant occurs slightly more frequently without intervening material in lower status speech, while the reverse direction of effects is displayed in older LA speech.

To summarize, the comparison between the LPTC and the LA synchronic dataset shows that there are general parallels in the patterns of linguistic conditioning of variant choice, especially in terms of *ma* and *miš* across higher status speech and educated middle-aged and younger LA speech, and in terms of *ma*, *ma-š* and *miš* across lower status speech and older LA speech. These general parallels suggest that there does not seem to be ongoing change in terms of the linguistic conditioning of variant selection in the verbal negation system in LA. Nevertheless, the comparison between the LPTC and the LA synchronic dataset can elucidate the social connotations that may attach to variant usage. The fact that higher status speakers in the LPTC and (middle-aged and younger) educated speakers in the synchronic dataset behave similarly in terms of variant rates (and variant conditioning) bears further testimony to the status of *ma* as the socially unmarked default exponent of verbal negation. In addition, the fact that higher status LPTC characters and middle-aged and younger speakers in the synchronic dataset avoid *ma-š* and *-š*, while lower status characters and less educated, older Lebanese speakers use them, is also further evidence of their status as stigmatized variants which conflict with both the higher social status of the LPTC characters as well as higher levels of education of the speakers in the synchronic dataset.

5.4 Discussion

Drawing on the framework of comparative variationist sociolinguistics, I examined the social and linguistic conditioning of variant choice in verbal negation in the synchronic and diachronic

datasets to answer the research questions posed in Chapter 1. The research questions relate to the social and linguistic constraints on variant choice in the synchronic and diachronic datasets, how the latter can be used to illuminate patterns of variation in the contemporary varieties, and evidence of PA converging on LA (here in terms of verbal negation).

The present study leverages new data to elucidate the social diagnosticity of variant choice in the Levantine verbal negation system, a topic that has previously been treated in an ad hoc, impressionistic fashion in the literature, and which has not relied on quantitative reasoning. In the LA synchronic dataset, *ma* was used mainly by post-secondary educated young and middle-aged speakers from Beirut, while less educated, older speakers from Mount Lebanon also used *ma-š* and *-š*. In the PA synchronic dataset, *ma* was favoured by females, educated young and middle-aged speakers, and non-camp-dwellers, while *ma-š* and *-š* were used more frequently by males, less educated middle-aged and older speakers, and by Palestinians who live inside the camps. Further evidence that *ma-š* and *-š* seem to be stigmatized comes from the diachronic datasets, as in the POHA, educated males (who live outside the camps) disfavoured these variants, and in the LPTC, these two variants were mainly disfavoured by females and higher status characters.

In terms of stylistic factors, in all the datasets, the variants *ma-š* and *-š* were generally disfavoured or avoided altogether in more careful (soapbox) speech. However, in the PA synchronic dataset, only *-š* occurred at a considerably higher rate in soapbox speech about politics because these topics were discussed mainly by (less educated) older and middle-aged males.

Turning to linguistic factors, I conducted apparent-time and real-time comparisons to assess whether Beirut PA is undergoing contact-induced change in the verbal negation system

due to the influence of LA. Variant distributions were generally similar across the LA and PA synchronic datasets, especially in terms of the rates of *ma* and *miš*, while *ma-š* and *-š* occurred slightly more frequently in the PA synchronic dataset. Change over time was investigated by comparing patterns of variant conditioning across older, middle-aged and younger speakers in LA and PA. Even though younger LA and PA speakers behaved similarly, I could not infer convergence because when comparing across age cohorts, I found that all speakers within each variety generally share a similar underlying grammar in terms of verbal negation. This shows that in terms of linguistic constraints on variant selection in this variable system, there does not seem to be a change in either variety.

Real-time comparisons across synchronic and diachronic datasets confirmed the earlier observations from the apparent-time comparisons. Across the POHA and young PA speech, it was found that there were general parallels in the linguistic conditioning of variant choice (at least as far as *ma*, *-š* and *miš* are concerned). In addition, there were several structural affinities between the speech of educated speakers in the POHA and educated, younger Palestinians in the synchronic dataset, especially in terms of the linguistic conditioning of *ma* and *-š*.¹³⁰ This confirmed that there does not seem to be structural change (whether contact-induced or otherwise) in the verbal negation system in PA as spoken in Beirut, and a similar conclusion was reached when comparing the patterns of linguistic conditioning of variant selection across the LPTC and the LA synchronic dataset.

While structural (contact-induced) change was not observed in the linguistic constraints governing the verbal negation system in PA as spoken in Beirut, we did observe that this variety

¹³⁰ The results in this study suggest that variant exponents of negation are phonologically conditioned, as in *ma-š* being favoured with non-labial-initial verbs and *-š* with labial-initial verbs in PA. These variants appear to be grammatically conditioned too, as indicated by the aspectual constraints on variant selection (e.g., *-š* being favoured with present tense verbs and pseudo-verbs in PA, as well as *miš* occurring mainly with future tense verbs or participles in the PA and LA synchronic and diachronic datasets).

is undergoing change in progress in terms of overall variant distributions. This is because in both the contemporary PA variety and in the POHA, which is considered to be an approximation to the pre-contact PA variety (see Chapter 3), we observed that as educational level increases, the rates of *ma-š* and *-š* decrease. However, it is notable that *ma-š* and *-š*, though disfavoured by educated males in the POHA, still occurred at noticeably higher rates in their speech than in the speech of the younger post-secondary educated Palestinians in the synchronic dataset. This shows that the change in terms of variant rates (increase in the rate of *ma* and decrease in the rates of *ma-š* and *-š*) has advanced among younger speakers, and it adds to the burgeoning evidence that the system of verbal negation in PA as spoken in Beirut is becoming streamlined over time, such that *ma* is colonizing the variable system in Beirut PA, similar to Beirut Arabic.

The findings in this chapter confirm that the system of verbal negation in PA as spoken in Beirut is undergoing dialect levelling (see Kerswill and Williams 2002:88) as a result of contact with Beirut Arabic, with socially marked variants being reduced over time, especially in the speech of educated Palestinians. Kerswill (2003:223) explains that dialect levelling results from the socio-psychological process of speech accommodation (Giles and Powesland 1997; Trudgill 1986:1–4), where interlocutors tend to converge linguistically (see also Britain 2010b:197; Hinskens 1996:457). In other words, dialect levelling is a product of convergent linguistic accommodation (Britain 2010b:197).¹³¹ This change in variant rates in PA as spoken in Beirut may have been abetted by young Palestinians' presumed regular contact with (young) LA speakers in post-secondary institutions in Beirut. As Britain (2010b:197) explains, the process of dialect levelling is a result of increased mobility and contact between speakers of different varieties (see also Kerswill 2003:230). In terms of verbal negation, then, convergence in PA lies

¹³¹ Convergence and dialect levelling are not distinct processes, i.e., they are not mutually incompatible. Dialect levelling is particularly used for the attrition of unusual or regionally marked forms in favour of unmarked variants (Trudgill 1986:98).

in the use of forms typical of Beirut Arabic, and in the levelling away of minority forms, especially in the speech of younger and educated Palestinians.

As Table 5.1 shows, this chapter was based on a substantial amount of data (a total of 8189 tokens of verbal negation), so the empirical aspect of this work gives us enhanced confidence in the representativeness and reliability of the results. This chapter brought new findings to bear on verbal negation in spoken (Levantine) Arabic which were not previously available in the relevant literature. As discussed in Chapter 2, there are little to no empirically accountable studies on verbal negation in spoken Arabic, and the variable context has not been circumscribed in accountable ways (i.e., the focus has been on one variant or another). The data and findings in this chapter present empirical evidence that elucidates the possible social meanings attached to the variants of verbal negation, especially *ma*, *ma-š* and *-š*. In addition, the results are congruent with those of other studies on dialect contact situations in the Arabic-speaking world, where rural or stigmatized speech features tend to be shunned by (young) post-secondary educated speakers in favour of the local urban features (see e.g., Al-Wer 1997, 2002). As Al-Wer (2002:46) explains, in the Arab world, the higher the educational level of the speaker, the more advanced the change is in her speech towards innovative, socially unmarked or urban forms (see also Al-Wer 1997:259). Apart from the Arabic-speaking world, the results in this chapter are also consistent with previous findings that in dialect levelling, unmarked forms (here *ma*) often survive at the expense of marked forms (Britain 2010b:195).

6 Future Temporal Reference

The third linguistic variable targeted in the present study to examine whether PA is converging on LA in Beirut is future temporal reference. This chapter presents the results for this variable in the Lebanese and Palestinian synchronic and diachronic datasets.

This chapter is structured as follows. Section 6.1 defines the variable context, and section 6.2 explains how the social and linguistic factors were coded, as well as the hypotheses they are designed to test. Section 6.3 presents the quantitative results regarding future temporal reference in the synchronic and diachronic datasets, and assesses whether there is (contact-induced) change in Beirut PA in this sector of the grammar. Section 6.4 discusses the results and presents a summary of the major findings, examining their contribution to elucidating the major research questions posed at the outset of this investigation.

6.1 Circumscribing the variable context

The variable context included any unambiguous reference to events or states occurring after speech time (Poplack and Tagliamonte 1999:325; Poplack and Turpin 1999:143). Several variants compete in the expression of futurity in LA and PA: *rah* and the proclitic *ha-*, which are said to derive from the verb of movement *ra:yih* ‘going’ (Jarad 2014:101; Leddy-Cecere 2018:49–50, 2020:606), as well as *bidd-* lit. ‘want’ and the *b*-prefix, which is cliticized to the imperfect verb in Levantine Arabic, including LA (Mitchell and El-Hassan 1994; Thackston 1996) and PA (AbuAmsa 2016).¹³² The *b*-prefix is a multifunctional particle that can also be used to denote non-future temporal meanings, such as the progressive and habitual aspect (see

¹³² Mitchell and El-Hassan (1994:13) refer to the *b*-prefix as the *b*-nonpast.

Jarad 2013:71; Owens 2018:211, 219). The aforementioned exponents of futurity are illustrated in the examples below, which are reproduced verbatim from the data:

- (1) RAḤ
raḥ xabrik xabriyye. [07/L/M/43/39:35]
 FUT tell.1s-you.f news/story
 ‘I am going to tell you a story.’
- (2) ḤA-
 fa halla? hiyye **ḥa**-tru:h ʕa ʔamerika. [13/P/F/28/42:28]
 so now she FUT-go.3sf to America
 ‘So now she is going to go to America.’
- (3) BIDD-
 niḥna bukra **bidd**-na nekbar. [POHA/02/05.07.2004/F/76/26:54]
 we tomorrow FUT.1p grow.up
 Intended meaning: ‘We will grow older tomorrow.’
- (4) B-
 šaḥer zame:n w **b**-erjaʕ. [LPTC/Faraman/1970/32/M/1:30:00]
 month time and FUT-return.1s
 ‘I will be back in a month’s time.’

There were several exclusions from the variable context. For example, I excluded instances of the future in the past, as in (5).

- (5) ke:net raḥ tku:n tajrebe ḥelwe. [10/L/M/19/44:04]
 was.3sf FUT be.3sf experience nice
 ‘It was going to be a nice experience.’

I also excluded instances where *bidd-* was used as a verb of volition (without reference to the future), as in (6), or where *b-* was used to express non-future temporal meanings, such as the (habitual) present, as in (7).

- (6) ma **bidd**-e l-maša:re. [15/P/F/20/06:56]
 NEG want.1s the-money
 ‘I don’t want the money.’
- (7) ma **b**-aʕmul research. [04/L/F/24/19:18]
 NEG IMPF.do.1s research
 ‘I don’t do research.’

In addition, I excluded instances where *rah* was used with a long vowel (i.e., *ra:h*), where it functions as a motion verb, and can be inflected for subject agreement (see Jarad 2014:106), e.g.:

- (8) lamma **ra:h** ʕa l-ḥarb, ka:n metzawwej.
 when went.3sm to/on the war was.3sm married.3sm
 ‘When he went to the war, he was married.’ [POHA/29/15.05.2003/M/72/02:37]
- (9) **ra:h-o** la huni:k, ʕa turkiyya.
 went.3p to there to/on Turkey
 ‘They went there, to Turkey.’ [POHA/12/06.07.2004/F/82/03:57]

In examples (8) and (9), *ra:h* conveys a sense of going or moving towards a location or a goal, but it does not encode future temporal reference, justifying its exclusion from the variable context. Only the future marker *rah* (with a short vowel /a/) was retained for analysis.¹³³

Table 6.1 shows the number of tokens of future temporal reference extracted from each dataset following exclusions.

Table 6.1: Number of tokens of future temporal reference extracted from the synchronic and diachronic datasets

Dataset	N
PA synchronic dataset	571
LA synchronic dataset	447
POHA	634
LPTC	1342
Total	2994

¹³³ Leddy-Cecere (2018:49–50) argues that *rah* is derived from *ra:yih* ‘going,’ which is originally the active participle of the lexical verb *ra:h* (with a long vowel /a/). Leddy-Cecere (2018:50) explains that while in some Arabic varieties *ra:yih* or *ra:h* may be used to refer to the future, in other varieties, such as some Levantine Arabic varieties (like Beirut Arabic), *ra:yih/ra:h* shows loss of agreement inflection along with phonetic reduction and loss of morphosyntactic autonomy, becoming the (future) particle *rah* (with a short vowel /a/) (see also Jarad 2014:106, 111).

6.2 Factor groups

6.2.1 Social factor groups

Synchronic datasets

Each speaker in the synchronic datasets was coded for age (and immigrant generation for Palestinians), speaker sex, level of education (primary/secondary versus post-secondary), and place of residence for Palestinians only (distinguishing speakers living inside and outside refugee camps in Beirut). Based on the small body of scholarship addressing the variable expression of future temporal reference in spoken Arabic, it is not clear if any exponents of futurity are subject to social evaluation, as no sustained attention has been paid to the extralinguistic conditioning of variant choice (see e.g., Jarad 2013, 2014; Leddy-Cecere 2020). However, because we are dealing with a relatively infrequent morpho-syntactic variable in natural discourse, the future temporal reference variable may be less amenable to social conditioning because it is not socially salient, or the target of prescriptive commentary (see e.g., Cheshire 1999:61; Hinskens et al. 2005; Milroy and Gordon 2003:63, 170–71; Rickford and Wasow 1995:106; Winford 1996:188). Exponents of futurity may therefore be less socially indexical or less associated with the speech of a particular social group (Hinskens et al. 2005:139, 143; see also Cheshire 1999:61; Milroy and Gordon 2003:171).¹³⁴

In the light of the preceding discussion, the effects of speaker sex, as well as place of residence for Palestinians on variant choice in future temporal reference, will be treated here as exploratory manoeuvres. The remaining social factor groups, speaker age and educational level, were collapsed together because, as explained in Chapter 4, all young and middle-aged LA

¹³⁴ Though Cheshire (1999:62) explains that some morpho-syntactic variables that may occur more frequently in speech, such as (verbal) negation, may be more susceptible to social evaluation. This is what we saw in Chapter 5, where the stigmatized variants *ma-š* and *-š* were generally disfavoured by younger (and educated) speakers, in contrast to their older (and less educated) counterparts.

speakers, as well as young Palestinians, are educated to post-secondary level, while their older counterparts are educated to primary or secondary level. The factor group of speaker age/education is hypothesized to play a key role in variant choice, especially in PA because of a reported change in progress in the expression of future temporal reference in this variety (AbuAmsa 2016). In PA as spoken in Gaza city, AbuAmsa (2016:14) claims that the future marker *ħa-* constitutes a change in progress, as it is favoured by younger speakers, while older speakers categorically disfavour it. Accordingly, if Beirut PA is undergoing a parallel independent change, we conjecture that *ħa-* will be favoured by younger (and educated) Palestinians, but it will be disfavoured by older Palestinians.

Diachronic datasets

In the POHA, speakers were coded for speaker sex, educational level and place of residence (inside versus outside camps). As mentioned in Chapter 5, because of a possible interaction between these three factor groups (i.e., there are only six post-secondary educated speakers, all of whom are males who live outside the camps), they were collapsed together as follows: male/post-secondary/ outside camp; male/ non-post-secondary/ outside camp; male/ non-post-secondary/ inside camp; female/ non-post-secondary/ outside camp; female/ non-post-secondary/ inside camp. Because *ħa-* may be a change in progress in PA led by younger speakers, we speculate that older speakers in the POHA may disfavour *ħa-*, as AbuAmsa (2016:9) found for older Palestinian speakers in Gaza city. However, it remains to be seen whether speaker sex, educational level, or place of residence among the POHA speakers will play a significant role in variant choice in future temporal reference.

In the LPTC, each token was coded for sex and age group. In addition, using the taxonomy adumbrated in Chapter 4, each token in the LTPC was coded according to the

characters' inferred social status, distinguishing characters having higher, middle or lower social status. With regard to tracking linguistic change, it is not clear in LA whether there is a change in progress in the future temporal reference sector, as is suggested for PA (AbuAmsa 2016), but, if this is the case, the expectation is that older characters will disfavour *ha-*, while younger characters will favour it. I explore later in this chapter whether the other social factors (speaker sex and social status) contribute to variant choice in the LPTC.

6.2.2 Linguistic factor groups

Several elements historically implicated in the grammaticalization of future markers cross-linguistically were operationalized as linguistic factors in the present study (see e.g., Bybee et al. 1994), following previous variationist studies on European languages (see e.g., Poplack and Dion 2009; Poplack and Malvar 2007; Poplack and Tagliamonte 1999; Poplack and Turpin 1999; Tagliamonte et al. 2014; Torres Cacoullos and Walker 2009). The first linguistic factor group is polarity, which includes two factors: negative contexts (10) and affirmative contexts (11).¹³⁵

(10) ma ḥa-tru:h maʕ-na. [30/P/M/56/04:15]
 NEG FUT-go.3sf with-us
 'She will not go with us.'

(11) ana baʕden b-šu:f-a. [16/L/F/19/51:38]
 I later FUT-see.1s-her
 'I will see her later.'

There are no previous claims in the literature on spoken Arabic about the effect of polarity on the choice of *rah* and *ha-*, but cursory observation of the data in the present study revealed that these two variants appear to be favoured in negative contexts, while *bidd-* appears to be favoured in affirmative contexts. Pereira (2009:555) found that in Libyan Arabic, the *b-* future marker is also

¹³⁵ Interrogatives were rare in the data, so they were re-coded into sentence polarity (i.e., whether they were in affirmative form or in negative form).

favoured in affirmative contexts, and the existence of the same correlation is explored here in LA and PA as well.

The second linguistic factor group is clause type, which includes two factors: main clauses and subordinate clauses. In Gulf Arabic, Persson (2008:33) observes that *rah* is most frequently used in main clauses. For the remaining variants, Jarad (2013:83) surmises that in Syrian Arabic, *bidd-* and the *b-* future marker may occur in main and subordinate clauses, though informal inspection of the data in the present study revealed that the *b-* future marker shows a preference for subordinate clauses. Bybee et al. (1994:234) explain that in the earlier stages of grammaticalization, future markers are associated with assertiveness, which tends to be associated with main clauses (see also Bybee et al. 1991:19). Following their assumption, if future markers occur in subordinate clauses, this may mean that they have lost their assertive meaning, indicating that they may be progressing along the cline of grammaticalization (Bybee et al. 1991:28, 31, 1994:235, 244, 271, 274).

The third linguistic factor group is temporal distance. I coded for proximate events (those occurring within the same day of the utterance), as in (12), and distal events (those occurring after one day or more), as in (13). If there was no overt contextual indication in the ambient discourse as to when the future event will take place, then the event/state was coded as temporally indeterminate, as in (14).

(12) baʕed dʔi:ʔa b-aḥki:-l-ak ʕann-o. [POHA/24/12.11.2003/M/72/21:44]
 after minute FUT-tell.1s-to-you.m about-him
 ‘I will tell you about him in a minute.’

(13) b-etkharraj bi l-Fall. [09/L/M/21/0:25]
 FUT-graduate.1s in the-fall
 ‘I will graduate in the Fall (semester).’

- (14) *raḥ* *ešraḥ-l-ik* *kel* *ši.* [LPTC/Sitt El Kel/1974/255/M/21:34]
 FUT explain.1s-to-you.f every thing
 ‘I will explain everything to you.’

In Syrian Arabic, Cowell ([1964] 2005:322–23) states that *raḥ* is a marker of proximate events. The proclitic *ḥa-* is also said to be used mainly for proximate futures in some varieties of Arabic, such as Jerusalem Arabic (Rosenhouse 2011; see also Alkalesi 2006:141; Holes 2004:228; Woidich 2005b:331). On the other hand, some scholars posit that the *b-* future marker is used for distal futures in Levantine Arabic (Mitchell and El-Hassan 1994:25; see also Cowell [1964] 2005:324). For *bidd-*, there are no speculations in the literature on spoken Arabic about the effect of temporal distance on its selection, but preliminary observation of the data showed that it may be preferred with proximate futures.

The fourth linguistic factor group is temporal adverbial specification, which includes two factors: presence of a temporal adverbial (15) or its absence.

- (15) *raḥ* *ḥa:ki:-ha* *bukra.* [LPTC/Dawaleeb Al Hawa/1965/166/F/42:49]
 FUT talk.(to).her tomorrow
 ‘I will talk to her tomorrow.’

Cowell ([1964] 2005:323) argues that in contrast to *b-*, *raḥ* is used unambiguously to refer to the future, and this might also be the case for the future marker *ḥa-*, which is said to have developed from *raḥ* (see e.g., Jarad 2014; Leddy-Cecere 2020). On the other hand, *bidd-* and *b-* denote other non-future meanings, so the working hypothesis is that they may show a preference for contexts where there are temporal adverbials that disambiguate their future temporal reference.

The fifth linguistic factor group is grammatical person, coded here as a binary factor group differentiating first grammatical person from non-first grammatical person. The proclitic *ḥa-* is reportedly used to express the speaker’s own intention (El-Hassan 2007:265; Holes 2004:228; Mitchell and El-Hassan 1994:25), so, if this is the case, it can be expected to occur

with first-person subjects. Because *raḥ* and *ḥa-* are believed to have developed from the same lexical material (Jarad 2014; Leddy-Cecere 2020), *raḥ* may also show a preference for first-person subjects. Similarly, because *bidd-* is originally a volition verb, Mitchell and El-Hassan (1994:39) hold that in the Levant, it is often used with reference to speaker intent. It is therefore postulated that *bidd-* may be used with first-person subjects. Brustad (2000:244) also maintains that in Syrian Arabic, the *b-* future marker often occurs with first-person subjects because like *bidd-*, *b-* is reportedly used to express the intention of the speaker (Brustad 2000:244).¹³⁶

6.3 Results

6.3.1 Overall distribution of variants

I first consider overall variant distributions in the four datasets, presented in Table 6.2.

Table 6.2: Overall distribution of the exponents of futurity in the synchronic and diachronic datasets

	raḥ		ḥa-		bidd-		b-	
	%	N	%	N	%	N	%	N
PA synchronic dataset	11	62/571	26	148/571	24	137/571	39	224/571
LA synchronic dataset	18	81/447	21	94/447	22	98/447	39	174/447
POHA	13	83/634	4	26/634	28	176/634	55	349/634
LPTC	59	792/1342	2	23/1342	8	104/1342	31	423/1342

The table above shows that LA and PA share the same exponents of futurity, although there are notable differences between the two datasets in variant distributions. In terms of convergence of PA on LA, the working hypothesis is that the existence of (superficially) similar structures across these two varieties in terms of variant forms should render them hospitable to contact-induced change. However, sharing the same set of variants does not necessarily mean that LA and PA

¹³⁶ In addition to the linguistic factor groups discussed in this chapter, I coded for animacy (human versus non-human subjects) and verb type (stative versus motion). However, preliminary quantitative analysis showed that these factor groups made negligible contribution to variant choice in the four datasets, so they were subsequently excluded from the analysis.

share the same underlying grammar constraining variant selection. To pursue that possibility further, I examine the internal conditioning of variant choice in section 6.3.3, but before that, it is important to consider what the overall variant distributions in Table 6.2 tell us about the future temporal reference sector in LA and PA.

Table 6.2 shows that the *b-* future marker occurs more frequently than other variants in all the datasets except in the LPTC. *Bidd-* in the latter also occurs at a much lower rate than in other datasets. Furthermore, the LPTC patterns differently in showing a disproportionate preference for *rah*; however, it is not evident why this is the case. As we will see later, there is no clear extra-linguistic motivation why *rah* should be preferred in the LPTC.

The distributions of *bidd-* and the *b-* future marker are almost identical in LA and PA; however, there are notable differences between the two varieties with respect to the distributions of *rah* and *ha-*. *Rah* occurs more frequently in LA than in PA, while *ha-* is more frequent in PA than in LA. In the diachronic datasets, *ha-* occurs at notably lower rates in contrast with the synchronic datasets. This appears to be consistent with AbuAmsha's (2016) claim that *ha-* is a change in progress in PA (spoken outside Lebanon) (see also Leddy-Cecere 2018:82), and that older Palestinians seem to favour *bidd-* and *b-* (AbuAmsha 2016:9), as we see in the POHA. Since *ha-* is also virtually absent in the LPTC but increases in the LA synchronic dataset, this variant may also be a change in progress in contemporary LA. To explore this hypothesis further, I examine variant distributions in apparent time, shown in Figure 6.1.

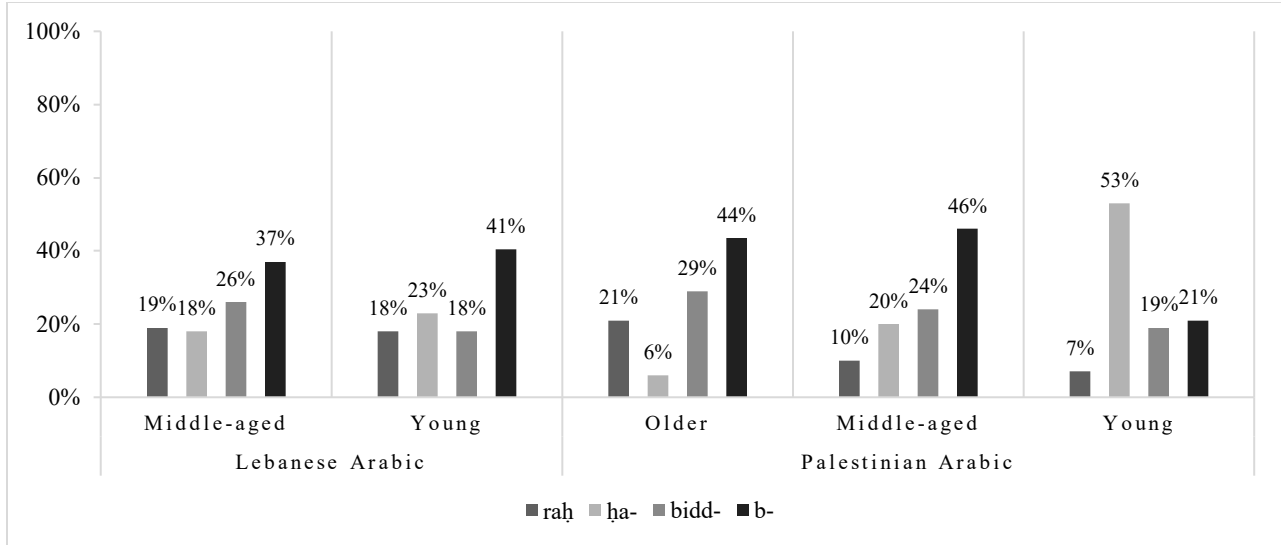


Figure 6.1: Overall variant distributions by age in the LA and PA synchronic datasets

In the figure above, we can see that there is an incremental rise in *ha-* in apparent time in PA, and it is younger Palestinians who lead in the use of this variant. Younger PA speakers use this variant at an even higher rate than younger (and middle-aged) Lebanese speakers. However, the increase in the rate of *ha-* in young PA speech has ramifications for the other variants, especially *rah* and the *b-* future marker, which occur at much lower rates in young PA speech in contrast to older PA speech, suggesting that the future temporal reference variable system in PA is undergoing longitudinal re-organization.

In LA, older speakers have been excluded from the analysis because they only produced 19 tokens of the target variable, precluding any detailed analysis of patterns in this age group. In the Lebanese dataset, we observe that *ha-* is used at similar rates across middle-aged and younger Lebanese speakers, and rates of other variants are also generally similar across these two age groups in LA. This is in contrast to PA, where there is a monotonic increase in the rate of *ha-* from older to younger speakers, suggesting that there is a change in progress in apparent time in

relation to *ha-* in PA (see Labov 1994:84, 2001:85, 171, 460).¹³⁷ A first conclusion to be drawn from Figure 6.1 is that variant rates distinguish between PA and LA, a preliminary finding that does not militate strongly in favour of the structural convergence of the comparison varieties.

6.3.2 Multivariate analysis of social factors

In this section, I subject the data to multivariate analysis to examine the social constraints on variant choice in future temporal reference.

Diachronic datasets

Table 6.3 shows the multivariate analysis of social factors in the POHA.

Table 6.3: Contribution of social factors to variant selection in the POHA

raḥ					ḥa- ¹³⁸		bidd-				b-			
Input prob.					0.084		Input prob.				0.419			
Total N					532 ¹³⁹		Total N				634			
Deviance					340.864		Deviance				671.802			
L.O.					N		Prop. App. Value				FW			
Sex/education/residence														
M/PS/outside	-.882	33	.061	.293	13.5	5/37	.042	37	.270	.510	.102	37	.541	.525
M/non-PS/outside	.133	89	.101	.533	4	4/108	-.428	108	.213	.395	.338	108	.667	.584
M/non-PS/inside	.700	135	.193	.668	4	7/165	-.120	165	.236	.470	-.145	165	.564	.464
F/non-PS/outside	.464	180	.206	.614	3	6/212	.358	212	.321	.589	-.458	212	.476	.387
F/non-PS/inside	-.415	95	.095	.398	4	4/112	.148	112	.321	.537	.163	112	.562	.541

The table above shows that the collapsed social factor groups of speaker sex, educational level and place of residence in the POHA are not selected as significant for any variant, though there

¹³⁷ Labov (1994:84) states that the classic pattern of linguistic change in progress in apparent time is a monotonic slope by age, e.g., “a steady increase in the proportion of use of a linguistic feature from one age cohort to the next” (Tagliamonte 2012:44).

¹³⁸ Only percentages and raw numbers are shown for *ha-* because its rate is too low in the POHA (N=26/634 or 4%), making multivariate analysis unreliable (see e.g., Tagliamonte 2006:84, 86–87).

¹³⁹ There are 102 tokens including distal futures that were excluded from the analysis (634 – 102 = 532) because of a K.O. value.

are slight differences in the factor weights (or log-odds) between speaker groups, especially with respect to *rah*. From the table above, we can see that female camp-dwellers, and especially educated males who live outside the camps disfavour this variant. However, the results from the educated males should be treated cautiously, as they only produced 37 tokens of the target variable. In terms of *bidd-* and *b-*, there are a few differences between speaker groups in the POHA. For example, female non-camp-dwellers weakly favour *bidd-*, while less educated males who live outside the camps disfavour it. The reverse direction of effects is observed for *b-*, where less educated male non-camp-dwellers weakly favour this variant, while females who live outside the camps disfavour it.

Turning to the LPTC, Table 6.4 shows the multivariate analysis of social factors in this diachronic dataset.

Table 6.4: Contribution of social factors to variant selection in the LPTC

(Note: shaded areas indicate statistically significant factor groups)

rah					ha ¹⁴⁰		bidd-				b-			
Input prob.					0.727		Input prob.				0.117			
Total N					1342		Total N				1342			
Deviance					1239.294		Deviance				1118.074			
L.O.					N		L.O.				N			
Prop.					%		Prop.				Prop.			
App.					N		App.				App.			
Value					FW		Value				Value			
FW					FW		FW				FW			
Sex														
Male					2		.055				.310			
Female					0.3		-.055				-.310			
Age														
Old					0		.537				-.456			
Middle-aged					3		-.073				.129			
Young					1		-.464				.327			
Social status														
Higher					1		-.393				.180			
Middle					2		-.045				-.148			
Lower					2		.438				-.032			

¹⁴⁰ Only percentages and raw numbers are shown for *ha-* because its rate is too low in the LPTC.

The table above shows that only speaker sex is significant in the selection of *rah* and *b-*. The former is favoured by females and disfavoured by males, while the direction of effects associated with speaker sex is reversed for *b-*. This sex-differentiated pattern in the LPTC is not reflected in the POHA, and as we will see later, speaker sex was also not selected as significant for any variant in the synchronic datasets. For other factor groups in the LPTC, factor weights are generally similar across speaker groups, but there are minor differences in variant selection. For example, *ha-* is mainly used by males, especially middle-aged males. However, we can see that it plays a very peripheral role in the LTPC, a finding that further bolsters the inference of change in the synchronic datasets. We also observe slight differences across speaker groups in the selection of *bidd-*, which is favoured by older and lower status characters. Nevertheless, the results in the LPTC generally lend support to the conviction that no exponent of futurity (in LA) is imbued with social prestige or stigma.

Synchronic datasets

Turning to the synchronic datasets, Table 6.5 shows the multivariate analyses of social factors in LA and PA.

Table 6.5: Contribution of social factors to variant selection in LA and PA synchronic datasets
(Note: shaded areas indicate statistically significant factor groups)¹⁴¹

Lebanese Arabic								Palestinian Arabic									
raḥ			ḥa-		bidd-		b-		raḥ			ḥa-		bidd-		b-	
Sex	Input prob.	0.112	Input prob.	0.142	Input prob.	0.219	Input prob.	0.356	Input prob.	0.064	Input prob.	0.131	Input prob.	0.059	Input prob.	0.388	
	Total N	447	Total N	447	Total N	447	Total N	447	Total N	571	Total N	571	Total N	571	Total N	571	
	Dev.	369.015	Dev.	354.342	Dev.	456.040	Dev.	535.461	Dev.	323.677	Dev.	463.702	Dev.	571.612	Dev.	636.796	
	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	
	Male	.481 .618	-.216 .446	.009 .502	-.285 .429	Male	-.286 .429	.163 .541	-.293 .427	.206 .551							
	Female	-.481 .382	.216 .554	-.009 .498	.285 .571	Female	.286 .571	-.163 .459	.293 .573	-.206 .449							
	Age/Education								Age/Education								
									Old/non-PS		.763 .682	-1.505 .182	.237 .559	.196 .549			
									Mid-aged/non-PS		.000 .500	.093 .523	-.159 .460	.330 .582			
	Mid-aged/PS		-.314 .422	-.040 .490	.209 .552	.044 .511	Mid-aged/PS		.068 .517	-.379 .406	.276 .569	.329 .582					
Young/PS		.314 .578	.040 .510	-.209 .448	-.044 .489	Young/PS		-.831 .303	1.791 .857	-.354 .412	-.855 .298						
								Residence									
								Outside		.200 .550	.343 .585	-.109 .473	-.091 .477				
								Inside		-.200 .450	-.343 .415	.109 .527	.091 .523				

¹⁴¹ The complete tables for the multivariate analyses of social factors in the LA and PA synchronic datasets (including Total N and proportion of application value) can be found in Appendix E.

Speaker sex is not selected as significant for any variant in the LA and PA synchronic datasets. Based on the direction of effects associated with non-significant factor groups, we can see that only in LA, *rah* is favoured by males, while it is disfavoured by females, which is the inverse of the direction of effects associated with speaker sex in PA (and in the LPTC). For the other variants in LA, speaker sex seems to play a minor role in variant choice, though we observe that females (weakly) favour *ha-* and *b-*, while speaker sex has a neutral effect on *bidd-*. On the other hand, in the PA synchronic dataset, it is males who weakly favour *ha-* and *b-*, while females favour *bidd-*. This shows that similar to *rah*, the direction of effects associated with speaker sex for the other co-variants is not replicated across LA and PA. Similar to speaker sex, place of residence in the PA dataset is not selected as significant, but there is a slight difference between speaker groups, as those outside the camps show a preference for *rah* and *ha-*, while camp residents weakly favour *bidd-* and the *b-* future marker.

In LA, age (collapsed with education) is also not selected as significant for any variant, as there are minimal differences between middle-aged and younger speakers, especially in terms of *ha-*, *bidd-* and the *b-* future marker. This reflects the distributional results in Figure 6.1 for LA. Only in terms of *rah*, factor weights slightly differ between younger and middle-aged Lebanese speakers.

Unlike in LA, there are clear differences between age groups in the PA synchronic dataset. In terms of *rah*, older Palestinians strongly favour this variant, while younger Palestinians strongly disfavour it. The direction of effects for *ha-* is reversed, where it is younger Palestinians who strongly favour this variant, in contrast to their older (and middle-aged) counterparts, lending further credence to the idea that this variant is a candidate for change in PA. Multivariate analysis confirms observations from Figure 6.1, as the fact that younger

Palestinians favour *ħa-* has repercussions for its co-variants, which are disfavoured by this age group, while they are generally favoured by older (and middle-aged) speakers. This is similar to AbuAmsha's (2016:9, 13) findings regarding the distribution of the exponents of futurity across age groups.

To sum up, the results regarding the social constraints on variant selection show that while speaker sex and place of residence for Palestinians were not selected as significant for any variant, speaker age/education in PA plays a key role in variant choice. The findings show that PA as spoken in Beirut seems to be undergoing ongoing change, where younger PA speakers are diverging from their elders with respect to *ħa-*, which may have just begun to enter the future temporal reference system for older PA speakers, as indicated by its lower rate (6%) in this age group, as we saw in Figure 6.1. The fact that *ħa-* is also an embryonic variant in the diachronic datasets also points in favour of a possible change in progress in PA (and possibly in LA too). However, as inferred from Figure 6.1, we can see that the future temporal reference variable system does not seem to be evolving in identical ways in LA and PA, especially in relation to *ħa-*, which is much more productive in young PA speech than in young LA speech. Preliminarily, this challenges a convergence analysis, but before we make any definitive claims about convergence, it is important to examine the linguistic conditioning of variant choice and compare it across relevant speaker cohorts.

6.3.3 Multivariate analysis of linguistic factors

In this section, I first conduct apparent-time comparisons across age groups in the LA and PA synchronic datasets to examine their participation in ongoing change. I also conduct real-time comparisons across the contact PA variety and the POHA. If contemporary PA features a constraint hierarchy that is different from that of LA, and if it was also found to differ from that

of the POHA (which approximates to the pre-contact PA variety), this means that we may be dealing with change in PA, but not one that is contact-induced; rather, it could be an internally-motivated change (see Poplack and Levey 2010:400–401).

LA and PA synchronic datasets

Apparent-time comparisons across age groups in LA and PA are presented in Table 6.6.¹⁴² It is important to note that for older Palestinians, only distributional analysis is presented because they produced only 85 tokens of future temporal reference.

¹⁴² Due to limited space, Table 6 does not include Total N, Deviance, log odds, and proportion of application value. It only includes (centred) factor weights, which RBRUL generates (see Chapter 4 for an explanation of centred FWs). The complete tables including apparent-time comparisons in the LA and PA synchronic datasets are found in Appendix F.

Table 6.6: Contribution of linguistic factors to variant selection across age groups in LA and PA synchronic datasets
(Note: shaded areas indicate statistically significant factor groups)

	Lebanese Speakers								Palestinian Speakers											
	Middle-aged				Younger				Older				Middle-aged				Younger			
	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-
	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Total %	Total %	Total %	Total %	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.	Input prob.
	0.006	0.260	0.291	0.422	0.140	0.139	0.158	0.252	21	6	29	44	0.046	0.107	0.065	0.466	0.051	0.752	0.119	0.245
	FW	FW	FW	FW	FW	FW	FW	FW	%	%	%	%	FW	FW	FW	FW	FW	FW	FW	FW
Polarity																				
Negative	.225	.785	<.001	.537	.513	.690	<.001	.443	64	14	0	22	.729	.648	.181	.366	<.001	.822	<.001	.369
Affirmative	.775	.215	>.999	.463	.487	.310	>.999	.557	13	4	35	48	.271	.352	.819	.634	>.999	.178	>.999	.631
Clause type																				
Main	.775	.505	.505	.343	.541	.614	.460	.421	29	9	33	29	.537	.651	.520	.373	.579	.618	.470	.358
Subordinate	.225	.495	.495	.657	.459	.386	.540	.579	14	2	26	58	.463	.349	.480	.627	.421	.382	.530	.642
Temporal distance																				
Proximate	.660	.569	.607	.402	.733	.712	.521	.163	8	0	50	42	.460	.647	.548	.432	.655	.201	.453	.702
Distal	.495	.534	.434	.557	.236	.339	.475	.772	8	0	50	42	.484	.350	.410	.611	.465	.623	.472	.536
Indeterminate	.344	.398	.458	.543	.540	.441	.504	.604	26	8	22	44	.556	.503	.542	.455	.378	.706	.574	.269
Temporal adverbial																				
Adverbial	.197	.230	.568	.679	.384	.270	.572	.626	6	0	43	51	.238	.277	.536	.631	.677	.411	.430	.597
No adverbial	.803	.770	.432	.321	.616	.730	.428	.374	32	10	20	38	.762	.723	.464	.369	.323	.589	.570	.403
Grammatical person																				
1 st person	.699	.563	.412	.454	.521	.400	.587	.480	19	3	37	41	.554	.422	.454	.566	.715	.434	.539	.507
Non-1 st	.301	.437	.588	.546	.479	.600	.413	.520	23	7.5	24.5	45	.446	.578	.546	.434	.285	.566	.461	.493

To assess whether there is convergent change in PA as spoken in Beirut, I begin by comparing patterns of variant conditioning across middle-aged LA and PA speakers, and then I move to younger speakers in both varieties. This will allow me to examine whether the grammars of LA and PA are becoming progressively more similar (i.e., converging) in apparent time. In the table above, I use the direction of effects constraining variant choice as the key heuristic for comparative purposes.

Across middle-aged LA and PA speech, there are some similarities in terms of the direction of effects constraining variant selection. For example, in both cohorts, *rah* is favoured in main clauses, especially in middle-aged LA speech, where the effect of clause type on *rah* is selected as statistically significant. These results are similar to Persson's (2008:33) findings, where *rah* occurred most frequently in main clauses in Gulf Arabic. In both middle-aged LA and PA speech, *rah* shows a statistically significant preference for contexts where there are no temporal adverbials, consistent with previous postulations that *rah* is used unambiguously to refer to the future (Cowell [1964] 2005:323). In terms of grammatical person, *rah* is favoured with first-person subjects, especially in middle-aged LA speech, where this effect is selected as significant.

Similar to *rah*, *ha-* in both middle-aged LA and PA speech does not necessarily require additional temporal adverbial specification to disambiguate its future temporal reference. Additionally, in both cohorts, *ha-* is strongly correlated with negative contexts. In contrast, *bidd-* across the comparison groups is mainly restricted to affirmative contexts, and unlike *rah* and *ha-*, *bidd-* has a weak propensity to occur in contexts where temporal adverbials are used. This is consistent with the hypothesis stated earlier, where *bidd-*, which has other non-future meanings, was expected to occur with adverbials which disambiguate its future temporal meaning.

Even though Mitchell and El-Hassan (1994:39) state that in the Levant, *bidd-* is often used with reference to the speaker's intention because of its original meaning of volition, results in the middle-aged LA and PA cohorts do not appear to support this view, as *bidd-* in both groups is weakly favoured with non-first-person subjects. Clause type seems to have a neutral effect on *bidd-* across the comparison groups. By contrast, the *b-* future marker shows a statistically significant correlation with subordinate clauses in both cohorts, suggesting that it has advanced along the pathway of grammaticalization, as mentioned previously (see Bybee et al. 1994:235). Both cohorts also display the same direction of effects associated with temporal distance, where *b-* is favoured with distal futures, in line with what Mitchell and El-Hassan (1994:25) conjectured for Levantine Arabic. In addition, as speculated, the presence of temporal adverbials in the ambient context has a statistically significant effect on the choice of *b-* in both varieties.

Despite the similarities between middle-aged LA and PA speakers, there are some key differences in the direction of effects conditioning variant choice. For example, while *rah* in middle-aged PA speech functions as a negative polarity item, in middle-aged LA speech, it is strongly favoured in affirmative contexts. Furthermore, *rah* in middle-aged LA speech is preferred with proximate futures, as Cowell ([1964] 2005:322–23) argued for Syrian Arabic. On the other hand, *rah* in middle-aged PA speech is weakly favoured with indeterminate futures, which implies that it is used in a more generalized context (as opposed to the more specialized context of proximate futures) (see e.g., Bybee 2003:605; Leddy-Cecere 2018:19–20; Schwenter and Torres Cacoullos 2008:30).¹⁴³

¹⁴³ Bybee et al. (1991:31) explain that as the future marker further proceeds along the cline of grammaticalization, it may be associated with less specialized contexts like those signifying simple futurity, i.e., without a specific temporal distance.

Turning to other variants, we also find several differences in the direction of effects across middle-aged LA and PA speech. In the former, *ha-* is weakly favoured with first-person subjects, so it may still be used to express speaker intention, similar to what El-Hassan (2007:265) alleged. On the other hand, *ha-* in middle-aged PA speech shows a preference for non-first-person subjects, indicating that it may be moving further along the cline of grammaticalization because it is not exclusively restricted to expressing speaker intention (see e.g., Bybee et al. 1991:26). Both groups also differ in terms of the direction of effects associated with temporal distance, though *ha-* in both cohorts is favoured with proximate futures, consistent with previous suppositions in the literature for Jerusalem Arabic (Rosenhouse 2011). Another important difference between the two comparison groups concerns the effect of polarity on *b-*. Middle-aged LA speakers weakly favour *b-* in negative contexts, while in middle-aged PA speech, the effect of affirmative contexts on *b-* is statistically significant. Moreover, *b-* in middle-aged LA speech is weakly favoured with non-first-person subjects, while the direction of effects is reversed in middle-aged PA speech, suggesting that *b-* in the latter may still be used to express speaker intent, as has been maintained in the literature (see e.g., Brustad 2000:244).

Table 6.7 summarizes the similarities and differences in the direction of effects constraining variant choice across middle-aged LA and PA speech. Despite some similarities across the two comparison groups, the patterns of variant conditioning are not entirely parallel, especially in terms of *rah*, *ha-* and *b-*. This further strengthens the case against convergence.

Table 6.7: Summary of the comparisons across middle-aged LA and middle-aged PA speech
(Note: shaded areas indicate statistically significant factor groups)

	Middle-aged LA speakers				Middle-aged PA speakers			
	rah	ha-	bidd-	b-	rah	ha-	bidd-	b-
Polarity	×	√	√	×	×	√	√	×
Clause type	√	—	—	√	√	√	—	√
Temporal distance	×	×	√	√	×	×	√	√
Temporal adverbials	√	√	√	√	√	√	√	√
Grammatical person	√	×	√	×	√	×	√	×

× = Different direction of effects √ = Same direction of effects — = Neutral effect¹⁴⁴

I now move to comparing between young LA and young PA speech. Focusing on the direction of effects constraining variant selection, we find several commonalities between young LA and young PA speakers. Both cohorts weakly favour *rah* in main clauses, similar to middle-aged speakers in both varieties. Additionally, young PA speakers strongly favour *rah* with first-person subjects, and the same direction of effects is displayed in young LA speech, though in the latter, the effect of grammatical person on *rah* appears to be neutralized. The two comparison groups also show similarities in the patterns of conditioning of *ha-*, which is strongly associated with negative polarity, and it shows a statistically significant preference for main clauses. Indications of its progress along the pathway of grammaticalization across the two comparison groups reside in the fact that *ha-* is favoured in the absence of temporal adverbials in the discourse context, and it shows a predilection for non-first-person subjects, similar to what we saw in middle-aged PA speech.

Similar to middle-aged speech in LA and PA, *bidd-* in young speech in both varieties is restricted to affirmative contexts. Furthermore, *bidd-* is weakly favoured in subordinate clauses

¹⁴⁴ A factor weight of around 0.5 is considered a neutral effect (see e.g., Tagliamonte 2012:141).

in young LA and PA speech. Both groups also weakly favour *bidd-* with first-person subjects. In contrast, grammatical person has a neutral effect on *b-* in young LA and PA speech, but as we saw in middle-aged speech in the two varieties, young LA and PA speakers also strongly favour *b-* in subordinate clauses. Additionally, *b-* shows a preference for affirmative contexts in young LA and PA speech, and as hypothesized, it is favoured in contexts where it is accompanied by temporal adverbials.

Even though there are some similarities across young LA and PA speech in the direction of effects conditioning variant choice, there are several important differences. For example, while the effect of polarity on *rah* is neutral in young LA speech, young PA speakers categorically favour *rah* in affirmative contexts. Furthermore, *rah* in young LA speech is strongly favoured without temporal adverbials, as we saw in middle-aged LA and PA speech. Nevertheless, *rah* in young PA speech shows a statistically significant correlation with contexts where there are temporal adverbials, though it should be stressed that this result is based on only 10 tokens of *rah* (N=10/141).

Consistent with Cowell's ([1964] 2005:322–23) assertion, both young LA and PA speakers favour *rah* with proximate futures, though they exhibit a different direction of effects associated with temporal distance. The two comparison groups also diverge in terms of the effect of temporal distance on *ha-*. Young LA speakers strongly favour *ha-* with proximate futures (though this is based on only 11 tokens of proximate states/events), while young PA speakers strongly favour *ha-* in the more general context of indeterminate futures, indicating possible progression along the grammaticalization pathway in young PA speech. The effect of temporal adverbials on *bidd-* is also different across the comparison groups, as young LA speakers weakly favour this variant with temporal adverbials, while the direction of effect is reversed in young PA

speech, indicating that *bidd-* in the latter may not necessarily need supporting material to disambiguate its future temporal reference. Another key difference between young LA and PA speech concerns the effect of temporal distance on *b-*. Young LA speakers strongly favour *b-* with distal futures. In young PA speech, by contrast, *b-* shows a statistically significant correlation with proximate futures.

Table 6.8 summarizes the results for young LA and PA speech.

Table 6.8: Summary of the comparisons across young LA and PA speech
(Note: shaded areas indicate statistically significant factor groups)

	Young LA speakers				Young PA speakers			
	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-
Polarity	—	√	√	√	×	√	√	√
Clause type	√	√	√	√	√	√	√	√
Temporal distance	×	×	×	×	×	×	×	×
Temporal adverbials	×	√	×	√	×	√	×	√
Grammatical person	—	√	√	—	√	√	√	—

× = Different direction of effects √ = Same direction of effects — = Neutral effect

Based on the comparisons of the patterns of variant conditioning across young LA and young PA speech, can we infer that young PA speakers are converging on young LA speakers in terms of future temporal reference? Poplack and Levey (2010:398) explain that a candidate for contact-induced change should parallel in a non-trivial way the behaviour of a counterpart feature in the presumed source of the change. Despite some parallels across young LA and PA speech, there are key differences in the underlying grammatical system constraining variant choice across the two groups, and the comparisons of the direction of effects constraining variant selection suggests that the same variants are not evolving identically in terms of all the factor groups in the

two cohorts. Building on the findings emerging from the comparisons between middle-aged LA and PA speakers, when we move down the age spectrum to young speakers, we can see that the grammars of LA and PA are not exhibiting greater similarity. This further suggests that we may not be dealing with the convergence of PA on LA in terms of future temporal reference. However, any inference of change (contact-induced or otherwise) should also rest on demonstrations that the behaviour of the candidate feature differs from its patterning in an antecedent stage of the variety in question (Poplack and Levey 2010:398). Accordingly, it is important to conduct apparent-time comparisons within each variety to examine whether young speakers behave like their elders or diverge from them.

With respect to young and middle-aged LA speakers, we find several similarities in the direction of effects constraining variant choice. In middle-aged LA speech, the effect of main clauses on *rah* is selected as statistically significant, and the same direction of effects is displayed in young LA speech. Additionally, both cohorts favour *rah* without temporal adverbials, and they share the same direction of effects associated with grammatical person, though only in middle-aged LA speech, the effect of first-person subjects on *rah* is statistically significant. Both cohorts also exhibit the same direction of effects conditioning *ha-* with respect to some factor groups. For example, the two age groups in LA favour *ha-* with negative polarity, and in contexts where there are no temporal adverbials. *Bidd-* displays an inverse direction of effects to that of *ha-* with respect to polarity and temporal adverbials, as both middle-aged and young LA speakers only use *bidd-* in affirmative contexts, and they weakly favour it with temporal adverbials. The *b-* future marker even shows a statistically significant correlation with temporally disambiguating adverbials, and it is the only variant that shows a significant association with subordinate clauses across the two comparison groups. Both age groups also

converge in terms of the effect of temporal distance on *b-*, as this variant is favoured with distal futures, followed by indeterminate futures.

Some differences between middle-aged and young LA speakers concern the effect of polarity on *rah*, which seems to be neutralized in young LA speech, while *rah* in middle-aged LA speech is strongly favoured in affirmative contexts. Although both groups favour *rah* with proximate futures, they do not share the same direction of effects associated with temporal distance. Likewise, with respect to *ha-*, the direction of effects associated with temporal distance is not parallel across the comparison groups, though in both middle-aged and young LA speech, *ha-* is also mainly used to express proximate futures. Middle-aged LA speakers also use *ha-* with first-person subjects, while the direction of effects is reversed in young LA speech. On the other hand, *bidd-* in the latter shows a statistically significant association with first-person subjects, while in middle-aged LA speech, it is weakly favoured with non-first-person subjects. Another difference between the two age groups in LA relates to the effect of polarity on the *b-* future marker. As we saw before, *b-* in middle-aged LA speech is weakly favoured in negative contexts, while in young LA speech, it is preferred in affirmative contexts.

Table 6.9 summarizes the results for middle-aged and young LA speakers.

Table 6.9: Summary of the comparisons across middle-aged and young LA speakers
(Note: shaded areas indicate statistically significant factor groups)

	Middle-aged LA speakers				Young LA speakers			
	<i>rah</i>	<i>ha-</i>	<i>bidd-</i>	<i>b-</i>	<i>rah</i>	<i>ha-</i>	<i>bidd-</i>	<i>b-</i>
Polarity	×	√	√	×	—	√	√	×
Clause type	√	—	—	√	√	√	×	√
Temporal distance	×	×	√	√	×	×	√	√
Temporal adverbials	√	√	√	√	√	√	√	√
Grammatical person	√	×	×	√	—	×	×	—

× = Different direction of effects √ = Same direction of effects — = Neutral effect

The comparison across middle-aged and young LA speech shows that despite some minor differences in patterns of variant conditioning, there are general structural affinities. This suggests that as far as these two age groups are concerned, the future temporal reference sector in LA seems to be relatively stable in apparent time with respect to patterns of variant conditioning.

In PA, there are a few structural affinities across older, middle-aged and younger speakers. In these three age groups, *rah* is favoured in main clauses, as we saw in young and middle-aged LA speech. Similarly, *ha-* is mainly restricted to main clauses in the three groups in PA, and it is strongly correlated with negative polarity. Furthermore, across the three age groups in PA, *ha-* is favoured without temporal adverbials, and it has a propensity to occur with non-first-person subjects, contra El-Hassan's (2007:265) postulation that *ha-* is used to express the speaker's own intention. The three cohorts in PA also converge in terms of the effect of polarity on *bidd-*, which is mainly restricted to affirmative contexts. The *b-* future marker displays the same direction of effects as *bidd-* with respect to polarity in the three age groups in PA. The *b-* future marker also mainly occurs in subordinate clauses in the three comparison groups in PA, as well as in contexts where its future temporal reference is disambiguated by a temporal adverbial, which is similar to what we saw in middle-aged and young LA speech.

Aside from the few similarities across older, middle-aged and younger PA speakers, there are striking differences in the direction of effects constraining variant selection, especially between younger Palestinians and their older counterparts. In terms of *rah*, this variant seems to function as a negative polarity item in older and middle-aged PA speech, while it only occurs in affirmative contexts in young PA speech. In addition, older and middle-aged PA speakers mainly use *rah* without temporal adverbials, similar to the two age groups in LA, but the direction of effects is reversed in young PA speech (though, as mentioned previously, this is based on only

10 tokens of *rah* produced by young PA speakers). On the other hand, young PA speakers strongly favour *rah* with first-person subjects, and the direction of effects is parallel in middle-aged PA speech, but older PA speakers are slightly more likely to use *rah* with non-first-person subjects. The three age groups also diverge in terms of the direction of effects associated with temporal distance, as *rah* is mainly restricted to proximate futures in young PA speech, while it has a propensity to occur with indeterminate futures in older and middle-aged PA speech. Older PA speakers also use *ha-* with indeterminate futures, while middle-aged PA speakers mainly use it for proximate futures. In young PA speech, temporal distance has a statistically significant effect on *ha-*, which is strongly favoured with indeterminate futures, followed by distal futures.

Turning to other variants, we observe that there are several differences in the direction of effects conditioning *bidd-* across the three age groups in PA. For example, in young PA speech, *bidd-* is weakly favoured in subordinate clauses, while the inverse direction of effects is exhibited in older and middle-aged PA speech. In addition, only in young PA speech, *bidd-* is weakly favoured in the absence of temporal adverbials, but this is not the case in older and middle-aged PA speech, where *bidd-* tends to occur with such adverbials. There are also a few differences across the three comparison groups in the patterns of conditioning of *b-*. A major difference concerns the direction of effects associated with temporal distance. In the distributional analysis for older PA speakers, we observe that *b-* occurs at similar rates across proximate, distal and indeterminate futures. However, middle-aged PA speakers mainly use *b-* for distal futures, while in young PA speech, *b-* is mainly restricted to proximate futures.

Table 6.10 provides a summary of the comparisons across the three age groups in PA.

Table 6.10: Summary of the comparisons across older, middle-aged and young speakers in PA

(Note: shaded areas indicate statistically significant factor groups)

	Older PA speakers				Middle-aged PA speakers				Young PA speakers			
	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-	rah	ħa-	bidd-	b-
Polarity	√	√	√	√	√	√	√	√	×	√	√	√
Clause type	√	√	×	√	√	√	—	√	√	√	×	√
Temporal distance	×	×	×	×	×	×	×	×	×	×	×	×
Temporal adverbials	√	√	√	√	√	√	√	√	×	√	×	√
Grammatical person	×	√	√	×	√	√	×	×	√	√	√	—

× = Different direction of effects √ = Same direction of effects — = Neutral effect

As the previous discussion and the table above show, there are key differences in the patterns of variant conditioning, where young Palestinians stand apart from their older and middle-aged counterparts, especially with respect to *rah* and *bidd-*. The fact that young PA speakers diverge from their elders, and they favour *ħa-* more than other age groups in LA and in PA may be attributed to an internal change in progress in PA as spoken in Beirut, rather than contact-induced change, consistent with what has been observed in non-contact PA spoken outside Lebanon (AbuAmsha 2016). In the next section, I compare patterns of variant conditioning in the speech of young Palestinians who are leading the change with those displayed by older Palestinians in the POHA who were born between 1897 and 1948 and whose speech, to all intents and purposes, approximates to the pre-contact PA variety. If young Palestinians also diverge from the POHA speakers, this is further evidence that the ongoing change in PA is the product of internal evolution.

POHA and Young PA speech

Table 6.11 shows the multivariate analyses of linguistic factors in the POHA and in young PA speech in the contemporary variety.

Table 6.11: Contribution of linguistic factors to variant selection in the POHA and in young PA speech (Note: shaded areas indicate statistically significant factor groups)¹⁴⁵

POHA								Young PA speech										
raḥ			ḥa-		bidd-		b-		raḥ		ḥa-		bidd-		b-			
Polarity Negative Affirmative Clause type Main Subordinate Temporal distance Proximate Distal Indeterminate Temporal adverbials Adverbial No adverbial Grammatical person 1 st -person Non-1 st -person	Input prob.	0.084	Total N	26/634	Input prob.	0.328	Total N	634	Input prob.	0.419	Total N	141	Input prob.	0.752	Total N	141	Input prob.	0.245
	Total N	532			Total N	634			Total N	634			Total N	141			Total N	141
	Dev.	340.864	%	4	Dev.	663.429	Dev.	671.802	Dev.	55.875	Dev.	154.752	Dev.	131.503	Dev.	119.824		
	L.O.	FW	%	N	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW	L.O.	FW		
	.991	.729	13	7/53	-8.851	<.001	-.283	.430	-9.304	<.001	1.532	.822	-8.796	<.001	-.538	.369		
	-.991	.271	3	19/581	8.851	>.999	.283	.570	9.304	>.999	-1.532	.178	8.796	>.999	.538	.631		
	.168	.542	6	8/128	.207	.552	-.413	.398	.318	.579	.483	.618	-.120	.470	-.584	.358		
	-.168	.458	4	18/506	-.207	.448	.413	.602	-.318	.421	-.483	.382	.120	.530	.584	.642		
	-1.166	.459	0	0/47	1.138	.757	-1.007	.268	.640	.655	-1.379	.201	-.187	.453	.856	.702		
	-9.664	<.001	0	0/102	-.581	.359	.665	.660	-.140	.465	.501	.623	-.112	.472	.146	.536		
.166	.541	5	26/485	-.557	.364	.342	.585	-.500	.378	.878	.706	.299	.574	-1.002	.269			
-1.158	.239	1	2/335	-.350	.413	1.050	.741	.741	.677	-.361	.411	-.280	.430	.392	.597			
1.158	.761	8	24/299	.350	.587	-1.050	.259	-.741	.323	.361	.589	.280	.570	-.392	.403			
-1.182	.455	6	15/265	-.518	.373	.412	.602	.922	.715	-.267	.434	.157	.539	.027	.507			
.182	.545	3	11/369	.518	.627	-.412	.398	-.922	.285	.267	.566	-.157	.461	-.027	.493			

¹⁴⁵ The complete table for the multivariate analysis of linguistic factors in the POHA (including Total N and proportion of application value) can be found in Appendix G.

Across the POHA and young PA speech, there are some structural similarities in the patterns of conditioning of variant choice, especially with respect to *ha-* and *b-*. As we saw before in the LA and PA synchronic datasets, *ha-* in the POHA and young PA speech is mainly reserved for negative contexts, and it shows a preference for main clauses, especially in young PA speech, where this effect is statistically significant. Both comparison groups also prefer to use *ha-* without temporal adverbials. In contrast to *ha-*, the *b-* future marker exhibits the inverse direction of effects associated with polarity, clause type and temporal adverbials. In the POHA and young PA speech, *b-* is favoured in affirmative contexts, and it is significantly correlated with subordinate clauses. As we saw before in LA, the effect of the presence of temporal adverbials on *b-* is selected as statistically significant in the POHA, and the same direction of effects is displayed in young PA speech. There are also a few similarities across the POHA and young PA speech in the patterns of conditioning of *bidd-*. Insofar as the diachronic data represented in the POHA show, it may be the case that *bidd-* originates in affirmative contexts, as it is also favoured in these contexts in young PA speech, similar to what we saw in other age groups in PA (and in LA). Both older speakers in the POHA and younger speakers in the PA synchronic dataset also share the same direction of effects associated with the factor group of temporal adverbials, as *bidd-* is preferred in the absence of such adverbials.

Apart from the few structural affinities discussed above, there are several differences in the direction of effects constraining variant choice when older speakers in the POHA are compared with younger Palestinians in the synchronic dataset. For example, in the POHA, *rah* is significantly correlated with negative polarity, similar to older and middle-aged PA speech, while in young PA speech, *rah* categorically occurs in affirmative contexts. Similar to what we saw in LA, as well as in older and middle-aged PA speech, *rah* in the POHA is strongly favoured

in contexts where temporal adverbials are absent, while young PA speakers exhibit the inverse direction of effects. As mentioned before, this is likely because young PA speakers produced too few tokens of *rah* (N=10/141), and 6 out of the 10 tokens of *rah* occurred with temporal adverbials. In terms of grammatical person, *rah* in the POHA is weakly favoured with non-first-person subjects, while in young PA speech, *rah* shows a statistically significant preference for first-person subjects (because 9 out of the 10 tokens of *rah* occurred with first-person subjects). On the other hand, *ha-* in young PA speech is weakly favoured with non-first-person subjects, while in the POHA, *ha-* is more likely to occur with first-person subjects. Both groups use *ha-* with indeterminate futures, but in young PA speech, it is also strongly favoured with distal futures, which is not consistent with previous claims in the literature (Rosenhouse 2011).

The two comparison groups also diverge in terms of the effect of some factor groups on *bidd-*. In the POHA, *bidd-* is preferred in main clauses, while in young PA speech, it is weakly favoured in subordinate clauses. As mentioned before, Bybee et al. (1994:234) suggest that the occurrence of future markers in subordinate clauses is an indication of an advanced developmental stage because future markers in an earlier stage of development are strongly associated with assertiveness, which tends to be associated with main clauses. Another difference between the POHA and young PA speech is that in the former, *bidd-* is mainly used to express proximate states/events, but in young PA speech, it is preferred in the more generalized context of indeterminate futures. On the other hand, *bidd-* in young PA speech is weakly favoured with first-person subjects, while in the POHA, it is strongly favoured with non-first person subjects, where it tends to mainly occur with indefinite subjects, e.g.:

- (16) yʔulu:-l-na **bidd-o** yʃi:r ɬareb.
 tell.3p-to-us FUT.3sm happen/become.3sm war
 Intended meaning: ‘They used to tell us there will be a war.’

[POHA/04/29.02.2004/M/79/04:14]

- (17) *ma həda yaʃref šu bidd-o yši:r.*
 NEG somebody know.3sm what FUT.3sm happen/become.3sm
 ‘Nobody knows what will happen.’ [POHA/64/29.05.1997/F/71/42:45]

Another important difference between the POHA and young PA speech is observed in the effect of temporal distance on the *b-* future marker. In the POHA, *b-* is strongly favoured with distal futures, similar to what we saw in middle-aged PA speech, while in young PA speech, it is mainly correlated with proximate futures.

Table 6.12 summarizes the comparisons between the POHA and young PA speech.

Table 6.12: Summary of the comparisons across the POHA and young PA speech

(Note: shaded areas indicate statistically significant factor groups)

	POHA				Young PA speakers			
	<i>raħ</i>	<i>ħa-</i>	<i>bidd-</i>	<i>b-</i>	<i>raħ</i>	<i>ħa-</i>	<i>bidd-</i>	<i>b-</i>
Polarity	×	√	√	√	×	√	√	√
Clause type	√	√	×	√	√	√	×	√
Temporal distance	×	×	×	×	×	×	×	×
Temporal adverbials	×	√	√	√	×	√	√	√
Grammatical person	×	×	×	√	×	×	×	—

× = Different direction of effects √ = Same direction of effects — = Neutral effect

Real-time comparisons between the POHA and young PA speech show that there are important differences in the patterns of variant conditioning, especially with respect to *raħ*, *bidd-*, and to a lesser extent, *ħa-*. Most importantly, *ħa-* is barely instantiated in the POHA, but it has risen vertiginously in younger (third-generation) speech, rapidly increasing within just a few generations. Accordingly, the fact that there are several ways in which young PA speakers diverge from older speakers in the POHA further suggests that PA as spoken in Beirut may be undergoing internal change in the future temporal reference sector.

6.4 Discussion

This chapter adopted an empirically accountable framework to address a major gap in the literature on the variable expression of futurity in spoken Arabic. Previous treatments of the future in spoken Arabic are largely based on scholars' intuitions about constraints on variant choice, rather than on systematic quantitative studies of actual speech that consider multiple social and linguistic factors simultaneously (see e.g., Jarad 2013, 2014; Mitchell and El-Hassan 1994). The results in the present chapter highlight the linguistic constraints on the selection of future variants in actual usage, where analysts' intuitions about variant patterning were not always confirmed, at least not in the varieties targeted in the present investigation. The results also reveal new linguistic constraints on variant selection not previously mentioned—to the best of my knowledge—in the literature on future temporal reference in spoken Arabic, particularly in LA and PA (e.g., sentence polarity, temporal adverbials).

The findings in this chapter have enabled me to address the research questions of the present study. In terms of social constraints on variant selection, social effects like speaker sex and place of residence (residence inside or outside camps among Palestinians) appear to play a relatively minor role in variant choice. As explained at the beginning of this chapter, a possible explanation for these findings is that we are dealing with a relatively infrequent morpho-syntactic variable in discourse which does not seem to attract normative commentary or overt social evaluation (see e.g., Cheshire 1999:61; Milroy and Gordon 2003:63, 170–71; Rickford and Wasow 1995:106; Winford 1996:188). However, results did reveal that speaker age/education in PA plays a key role in variant choice, especially as *ḥa-* is highly productive in the speech of young Palestinians, in contrast to other age groups in PA (as well as in LA). This was attributed

to a possible change in progress in PA, in line with previous conjectures in the literature (AbuAmsa 2016).

In order to check the source of the ongoing change in PA as spoken in Beirut, the underlying grammatical system constraining variant choice was compared across relevant benchmarks. In particular, the constraint hierarchy and direction of effects governing each linguistic factor group were used as a key diagnostic to assess convergence of PA on LA with respect to future temporal reference. Even though LA and PA share the same major exponents of futurity, this surface similarity masks some underlying structural differences between the two varieties, as the results turned up several lines of evidence which stack up against an interpretation of wholesale convergence of PA on LA in terms of the variable expression of futurity.

Apparent-time comparisons across LA and PA revealed several indications of divergence, where the patterns of conditioning of variant choice, especially *rah*, *bidd-* and *b-*, were not entirely parallel across middle-aged LA and PA speakers on the one hand, and across young LA and PA speakers on the other. As such, the same variants in LA and PA are not necessarily embedded in isomorphic grammars. Within LA, apparent-time comparisons showed that there were general parallels across middle-aged and younger speakers, with a few minor differences. In contrast, within PA, apparent-time comparisons across the three age groups showed that there were some differences, particularly between older and younger PA speakers, not only in terms of variant rates, but also in terms of the structure of variability, especially with regards to the constraint rankings associated with *rah* and *bidd-*. This suggests that PA is undergoing change in the future temporal reference variable system, especially in terms of the

dramatic rise of *ha-* in young PA speech, but, crucially, this change has unfolded internally and independently of contact with LA.

Real-time comparisons also seemed to confirm that we are dealing with an internal development in PA. Across the POHA and young PA speech in the contemporary contact variety, there were several differences in the patterns of variant conditioning. The fact that the hierarchy of constraints conditioning variant choice in the contact PA variety is different from the POHA, which approximates to pre-contact PA, as well as from LA, again shows that while there is ongoing change in the future temporal reference sector in Beirut PA, this change is not contact-induced. Since *ha-* increases in young PA speech, as well as in contemporary LA (in contrast to the LPTC), it is possible that these parallel developments with respect to this variant may be a product of drift (Sapir 1921:171–72; Trudgill 2004:132–33), i.e., parallel internally-motivated changes in several varieties (see e.g., Hinskens et al. 2005:16–17; Kerswill and Trudgill 2005:217–18; Labov 2001:416), though this change seems to be more advanced in PA, as inferred from the higher rates of *ha-* in young PA speech.

Situating the results in a wider perspective, the primary findings confirm previous claims that infrequently occurring (morpho-)syntactic variables in speech may be less accessible as a target for long-term dialect accommodation culminating in structural convergence (Cheshire et al. 2005:139). These findings also highlight the capacity of the variationist framework to determine whether there is change, and to identify the source of the change, i.e., to distinguish between contact-induced and internally-motivated change.

7 Conclusion

I am now in a position to review the major research objectives posed at the outset of this study and to discuss the contributions of the comparative methodological framework and analytical approach to elucidating these objectives.

This chapter is structured as follows. Section 7.1 summarizes the key motivations for the study, and section 7.2 reviews the major objectives of the study. Section 7.3 briefly explains the utility of the data sources in addressing the research objectives, and section 7.4 discusses the importance of the comparative analytical method for examining the major research issues at the heart of the investigation. Section 7.5 presents a brief recapitulation of the results of the study and explores their wider implications. Section 7.6 discusses the limitations of the study and suggests future lines of inquiry. The chapter concludes with an epilogue in section 7.7.

7.1 Key motivations for the study

The present study has sought to examine the outcomes of dialect contact in Beirut where there are several major social factors believed to be propitious to contact-induced change (see Thomason 2001; Thomason and Kaufman 1988), such as prolonged contact between Lebanese Arabic (LA) and Palestinian Arabic (PA), spoken by the local refugee population; the majority status of LA and the minority status of PA; and the (overt) prestige that is reportedly characteristic of Lebanese (Beirut) Arabic (Hennessey 2011:16; Versteegh 2001:153). A further linguistic factor believed to be favourable to convergent change between LA and PA is the typological congruence of these two varieties of Levantine Arabic (see e.g., Thomason 2001:97, 2010:40).

There are several gaps in the (socio)linguistic literature on spoken Arabic which this study aimed to address. We observed at the outset that there is currently a small body of sociolinguistic research targeting PA varieties, which have undergone drastic demographic and social change in recent decades (Cotter 2013; Horesh 2014, 2020; Horesh and Cotter 2016), and little attention has been paid to PA as spoken in Lebanon in particular (Fityan 1981; Hennessey 2011). We further observed that while studies on Arabic dialects in contact are a burgeoning area of inquiry, many are not articulated from a (comparative) variationist sociolinguistic perspective (Choueiri 2019:192; Owens 2013:10–11). Many studies also suffer from the absence of a socio-historical component incorporating diachronic data sources with which to better understand the process of language change in Arabic (Owens 2013:4, 9). We also noted that the focus of most Arabic sociolinguistic research is primarily on investigations of phonological variation (e.g., Al-Wer 2007b, 2011, 2020; Cotter 2013, 2020), with markedly fewer studies targeting (morpho-) syntactic variation (Brustad 2000; Choueiri 2019). Owens (2013:10–11) observes that while some phonological variables have received considerable attention in the sociolinguistic literature on spoken Arabic, other variables from different domains of the grammar, “of broad comparative potential”, await detailed treatment. Even in those studies focusing on (morpho-) syntactic variation in Arabic, systematic analyses of natural speech data are conspicuously rare (Jarad 2013, 2014; Lucas 2010; Mitchell and El-Hassan 1994; but see Owens, Dodsworth, and Kohn 2013; Owens, Dodsworth, and Rockwood 2009).

The original contributions of the present research reside in their capacity to speak directly to the aforementioned research gaps and to transcend the methodological shortcomings in the previous literature on spoken Arabic. A cornerstone of this research is its adoption of an empirically accountable approach to examine the underlying structure of variability in

spontaneous speech with respect to different linguistic sub-components, and its incorporation of a diachronic component to elucidate the process of linguistic change in PA as spoken in Beirut.

7.2 Objectives of the study

The overarching aims of the present research were to investigate the linguistic and social constraints on variant choice within three variable systems in PA and LA, and to examine the extent to which PA shows evidence, gauged from variant distributions and, crucially, from linguistic constraints on variant selection,¹⁴⁶ of becoming more structurally similar to LA with respect to different linguistic sub-components. The three variables targeted in the present study include a phonological variable involving the word-medial raising of /a:/ to [e:], known in the Arabic grammatical tradition as *imala*; and two morpho-syntactic variables, verbal negation and future temporal reference.

The working hypothesis enunciated at the outset of the thesis was that the phonological feature of raising of /a:/ to [e:] in word-medial position will be more favourable to contact-induced change than the two morpho-syntactic variables. This is because this feature is socially salient among the targeted Palestinian community. Trudgill (1986:20) explains that accommodation occurs by the modification of linguistic features that are salient in the variety to be accommodated to. Therefore, the fact that Palestinians are aware of word-medial *imala* as emblematic of LA should render it susceptible to convergent change. This working hypothesis also aligns with the traditional view propounded in the literature that (morpho-) syntactic variables may be less susceptible to contact-induced change than phonological variables, as

¹⁴⁶ As previously explained, the key line of evidence for the convergence of PA on LA comes from the direction of effects, or the order from more to less of factors within a factor group (Tagliamonte 2006:235, 2012:122), which shows the pattern of favouring and disfavouring effects on variant choice, or the constraint hierarchy.

previously explained in Chapter 6 (see also Cheshire et al. 2005; Sankoff 2001:658; but see Thomason 2001:71).¹⁴⁷

7.3 Utility of the data sources for addressing the research objectives

Poplack and Levey (2010:394) explain that an important prerequisite in establishing change is comparison over time (e.g., comparing a contemporary variety with an earlier stage of that variety), and this necessitates a socio-historical approach (see Auer et al. 2015; Romaine 1982). For that reason, the present study combined synchronic and diachronic data sources to pursue the inference of contact-induced change in PA as spoken in Beirut.

The synchronic data for this study come from extensive corpora of natural speech data from 39 Palestinian and 27 Lebanese speakers in Beirut, generating a total of 7,671 tokens representing the three targeted variables. The real-time component of the present investigation comes from two diachronic data sources which generated 15,382 tokens of the three targeted variables. The amount of data that has been brought to bear on the analyses expounded in Chapters 4–6 bolsters our confidence in the robustness of the results that were uncovered. By judiciously factoring two diachronic data sources into the analysis, the present study has demonstrated that even though it is claimed that the “diachronic dimension is not always accessible to investigation because most dialects of Arabic lack early written records” (Herin 2019:95), it is possible, with a little ingenuity, to locate diachronic speech surrogates of LA and PA, at least from the 20th century.

7.4 Contributions of the comparative variationist framework to addressing the research objectives

The present study is embedded in the framework of comparative variationist sociolinguistics

¹⁴⁷ As mentioned previously, Thomason (2001:71) argues that under the right social and linguistic conditions, “anything goes, including structural borrowing that results in major typological changes in the borrowing language.”

(Poplack and Tagliamonte 2001). The importance of this methodological framework lies in its capacity to tease apart idiosyncratic uses from systematic patterns among members of a speech community (Poplack 2021:49, 58). It is also capable of detecting linguistic change and discerning internal change from contact-induced change by relying on systematic comparisons of the fine-grained conditioning of variant selection across relevant (synchronic and diachronic) benchmarks and incorporating apparent- and real-time components, rather than merely relying on variant frequencies on which most previous inferences of change have been based. The results generated by this approach have highlighted the gap between scholars' impressionistic claims about patterns of variation and change and actual patterns observed in natural speech (especially with respect to future temporal reference in spoken Arabic). The results serve as a cautionary reminder that because the structure of variation lies so far below the level of consciousness, it cannot be intuited, rendering it invisible to any but systematic quantitative analysis of natural speech (Poplack and Torres Cacoullos 2015:272). This empirically accountable variationist methodology (combined with a comparative historical sociolinguistic component) could be profitably extended to future scientific investigations of patterns of linguistic variation and change in spoken Arabic.

7.5 Convergence and divergence in different grammatical components

What do the results reveal about the effects of dialect contact between LA and PA in Beirut? Quantitative analysis revealed that the clearest evidence of contact-induced change in PA as spoken in Beirut comes from the phonological variable of *imala* in word-medial position. As we saw in Chapter 4, the variable raising of word-medial /a:/ to [e:] was virtually absent in the speech of older Palestinians in the POHA, where only two speakers used *imala* in word-medial position to any appreciable extent, possibly influenced by the speech norms in their villages of

origin in Palestine, which were previously part of Lebanon. Older Palestinians in the synchronic dataset rarely use *imala* word-medially (N=2/566), again suggesting that the [e:] variant does not seem to be indigenous to PA, as spoken in Beirut (Hennessey 2011:59; Naïm 2005:277, 2011:925; Rosenhouse 2006:482–83). Among the second-generation (middle-aged) Palestinians in the synchronic dataset, word-medial *imala* only slightly increased (N=60/1065 or 6%), but it was mainly restricted to the speech of two female speakers, possibly due to their employment conditions, which presuppose regular exposure to LA speakers. This is despite the fact that one of the middle-aged female speakers lived in a refugee camp, suggesting that even though the dense social networks in refugee camps are expected to favour the maintenance of Palestinian vernacular norms (Hennessey 2011:1, 12, 19–20), individual conditions of social life may crucially determine accommodation towards the majority Lebanese speech patterns (see Villena-Ponsoda 2005:331).

In contrast to their older counterparts, younger (and educated) Palestinians in the synchronic dataset used *imala* in word-medial position much more frequently (N=282/654 or 43%), especially young females who live outside the camps, in line with the reported propensity of females to adopt supralocal linguistic features (Cheshire 2002:430; Milroy et al. 1994:352) and to lead change from above the level of conscious awareness. These findings suggest that we are not dealing with a case of transmission and incrementation of an internally-motivated change in PA from the second- to the third-generation Palestinians. Instead, because word-medial *imala* was found to be the default variant in the LA synchronic dataset and in the LPTC, and because younger PA speakers largely replicated the linguistic constraints on variant selection in word-medial *imala* found in young LA speech (where pharyngeals strongly disfavoured *imala* in the preceding and following phonological segments in both datasets), we concluded that we are

dealing with an externally-motivated change in PA due to the influence of LA. As such, it can be inferred that the higher rates of *imala* in young PA speech may be attributed to regular face-to-face interaction with Lebanese speakers, especially in post-secondary institutions, extrapolating from Al-Wer's (2002:51–52) argument that level of education is a proxy variable, which acts on behalf of amount and nature of contact with speakers of the target features.

But why did we find evidence of convergent change in terms of word-medial *imala* in the speech of the majority of the younger, third-generation Palestinians in this study? First, as detailed in Chapter 4, what likely facilitated this change in young PA speech is the fact that word-medial *imala* does not seem to be embedded in a complex structural system, as it does not require internal reconfiguration of the grammar. The fact that there is a similar feature in PA, word-final *imala* (in the feminine suffix of nouns and adjectives) (see e.g., Al-Wer 2020:561), with similar patterns of linguistic conditioning (as far as the effect of the preceding phonological segment is concerned) suggests that younger PA speakers may be simply extending this internal conditioning to word-medial *imala*. In addition, as discussed in Chapter 1, word-medial *imala* is overtly commented on in the targeted Palestinian speech community and associated by community members with LA. Accordingly, its conscious association with Lebanese speech patterns might explain why this feature is accessible as a target for long-term dialect accommodation for some younger Palestinians.

Notwithstanding the social salience of word-medial *imala* in the targeted Palestinian community and its productivity in the speech of the majority of the third-generation Palestinians, recall from Chapter 4 that there were three younger Palestinians who rarely or never used the [e:] variant word-medially. Two of these speakers live in refugee camps, where dense, multiplex social networks may have contributed to increased loyalty to conservative Palestinian speech

norms. Additionally, a few of the younger Palestinian outliers explained that they consciously maintain Palestinian speech patterns as a way of maintaining their Palestinian national identity, and this means avoiding word-medial *imala*.¹⁴⁸ In the case of these few younger speakers, then, the indexical meaning of the [e:] variant as characteristic of LA, and their positive attitudes towards their Palestinian identity and speech variety may have caused these speakers to resist the adoption of word-medial *imala*. This suggests that accommodation to Lebanese speech norms is subject to some degree of conscious control, at least in the case of a few Palestinian speakers. These findings are in line with Trudgill's (1986:125) claim that some linguistic features may be consciously avoided by speakers if they have extra-strong salience, that is, the features are overly strong markers of the dialect being accommodated to (Kerswill and Williams 2002:89), as in the stereotypical association of word-medial *imala* with LA here.

With respect to verbal negation, results revealed that in the speech of educated, younger (and middle-aged) Palestinians, there is a sharp decrease in the discontinuous negative construction *ma-š* and the post-verbal enclitic *-š*, which are reportedly stigmatized because they index rurality or provinciality (see e.g., Fleisch 1974:349, 353; Khairallah 2014:77; Wilmsen 2014:152). These two variants are also not characteristic of indigenous Beirut Arabic (Wilmsen 2014:150; see also Behnstedt and Woidich 2005:101; Fleisch 1974:349, 353). Results also showed that these two exponents of verbal negation were disfavoured by Palestinians who live outside the camps. By contrast, distributional analysis revealed that camp residents, especially first- and second-generation males, used *ma-š* and *-š* more frequently, again highlighting the role

¹⁴⁸ The explanations offered by the few Palestinian outliers may be expected in light of the circumstances in which Palestinians live in Lebanon. Had I interviewed more isolated speakers in the camps, these ideologies might have been even more pronounced. Still, it is important to acknowledge that younger Palestinians do not necessarily constitute a homogenous group and may evince different degrees of loyalty to Palestinian identity (see O'Connor 2018). This remains an important avenue to pursue in future research.

of the dense social networks of the camps in maintaining what is reportedly characteristic of rural PA speech norms (Awwad 1987:111; Mousa 2019:88).

In terms of the underlying linguistic constraints governing variant choice, there were general parallels between young LA and young PA speech. For example, in both cohorts, *ma* was generally favoured with most verb forms (past and present tense verbs, and pseudo-verbs), and it occurred more frequently with labial-initial verbs and in the absence of any intervening material between *ma* and the negated verb. On the other hand, *miš* was mainly restricted to negating future tense verbs or participles, and it occurred more frequently with non-labial-initial verbs and in the presence of intervening material.¹⁴⁹ Despite these parallels between younger LA and PA speakers, there were also striking similarities in the patterns of internal conditioning of variant choice across younger and older Palestinians in the synchronic dataset, as well as across younger Palestinians and the POHA speakers. These similarities in the underlying grammar between the contemporary contact variety and the POHA do not reveal compelling evidence in favour of structural change in PA as spoken in Beirut with respect to the linguistic factors constraining variant selection in verbal negation. However, as previously mentioned, there is convergent change in terms of variant rates in PA as spoken in Beirut. Results showed that the system of verbal negation in contemporary PA is becoming streamlined, such that *ma* is the unrivalled dominant variant, as in LA. This suggests that the system of verbal negation in PA is undergoing dialect levelling as a result of contact, with localised or distinctive minority forms being levelled away in favour of spatially more widely distributed variants, especially in the speech of educated

¹⁴⁹ Recall from Chapter 5 that young LA speakers did not use *ma-š* and *-š*, so only patterns of the linguistic conditioning of *ma* and *miš* were compared across young LA and young PA speech.

Palestinians.¹⁵⁰ But why did we find evidence of this convergent change with respect to variant rates in verbal negation? One possibility is that certain exponents of verbal negation appear to be sensitive to prescriptive commentary or overt social evaluation, and speakers are aware of the social meanings that overtly attach to the choice of verbal negator, as indicated from explicit meta-commentary (see Chapter 5). As such, it is the social indexicality of *ma-š* and *-š* that renders them vulnerable to attrition.

In terms of future temporal reference, the comparative approach adopted in the present study did not turn up compelling evidence of convergence of PA on LA. Results indicated that there is a dramatic rise in the use of the proclitic future marker *ha-* among younger (and educated) Palestinians, while this variant is rarely used by older Palestinians (in the synchronic dataset and in the POHA). This indicated that there is a change in PA in this sector of the grammar, but it is arguably not a case of contact-induced change for several reasons. First, *ha-* was more productive in young PA speech than in young (and middle-aged) LA speech. Second, the rise in the proclitic future marker *ha-* has also been reported in a non-contact PA variety spoken outside Lebanon, in Gaza city, led by younger speakers as well (AbuAmsa 2016). This parallel development in a variety of PA spoken *outside* Lebanon is damaging to any claim that contact with Lebanese Arabic is the key motivation for the vertiginous ascendancy of proclitic *ha-* in PA spoken in Beirut. Third, even though there were a few similarities in the patterns of linguistic conditioning of variant selection across young LA and PA speech, there were striking differences which militated against a convergence analysis. Instead, it was inferred that PA as spoken in Beirut seems to be undergoing internal change in the future temporal reference variable system that is led by younger speakers, especially because younger Palestinians in the

¹⁵⁰ We could also apply a dialect levelling approach to word-medial *imala*, where the non-raised variant, associated with local, peripheral speech varieties in the Levant (see Hennessey 2011:35, 92 and Versteegh 2001:153), is giving way to the raised variant.

present study also diverged from their older counterparts (in the synchronic dataset and in the POHA) in terms of the patterns of linguistic conditioning of variant selection.

Even though there are social and linguistic factors that are favourable to the structural influence of LA on PA in Beirut, we only find evidence of convergent change in PA in terms of word-medial *imala* and verbal negation. But why did we find little compelling evidence of convergence in the variable expression of futurity? We can only speculate. One possibility is that the choice of future markers is not normally subject to correction or metalinguistic commentary, nor is this choice explicitly associated by speakers with one or another variety. In other words, future temporal reference seems to operate below the level of conscious awareness. For this reason, the future in these varieties seems to remain relatively inaccessible as a target for long-term dialect accommodation.

What are the key take-home messages of the linguistic results of this investigation? In terms of the susceptibility of linguistic components to convergent change, the findings of this study are at variance with what has been traditionally claimed in the literature, i.e., that (morpho-)syntactic variables may be less vulnerable to (contact-induced) change than phonological variables, which are more labile (see e.g., Cheshire et al. 2005; Hinskens et al. 2005; Sankoff 2001:658). In the present study, we do not find that change only takes place in the phonology, nor do we find a neat distinction between phonology and morpho-syntax because verbal negation appears to have more in common with *imala* than with the other morpho-syntactic variable, future temporal reference. The salience of word-medial *imala* to the Palestinian community in Beirut makes this feature a target for convergence, at least in the speech of some younger, educated Palestinians. We also saw convergent change in the verbal negation system in PA with respect to variant distributions. This is possibly because both *imala* and negation seem to operate

above the level of conscious awareness, though perhaps not to an identical extent. But, as mentioned above, this does not seem to be the case for future temporal reference. So to conclude, those variables that are socially indexical, and which people are aware of, and explicitly remark on, seem to be particularly propitious to convergence, at least in this study. These findings align with Trudgill's (1986:11) notion of salience,¹⁵¹ where he argues that in a cross-dialectal interaction, there is a general tendency for speakers to modify those features of their own varieties of which they are most aware.¹⁵²

The findings in this study also highlight the social embedding of linguistic change. For example, in terms of speaker sex, results showed that in *imala* and verbal negation, it was mainly (young) females who led the change towards the majority or unmarked variant. Place of residence for Palestinians also played a vital role in variant selection in these two variables, where residence outside the camps was generally more conducive to dialect accommodation towards Lebanese speech norms. In contrast, residence within the relatively socially and spatially

¹⁵¹ Though Trudgill (1986) only investigates phonology in his theory of salience.

Kerswill and Williams (2002:81) define salience as the “property of a linguistic item or feature that makes it in some way perceptually and cognitively prominent.” Siegel (2010:129) also states that salience “refers to the characteristic of being easily noticeable, prominent or conspicuous.” Similarly, Hickey (2000:57) explains that “salience is a reference to the degree to which speakers are aware of some linguistic feature.”

¹⁵² In this study, we refer to Trudgill's (1986) notion of salience in dialect contact, showing that some degree of salience may be necessary for a linguistic feature to become conspicuous, making it a possible candidate for contact-induced change. However, it is important to acknowledge that there is no guarantee that socially salient linguistic features will always be more susceptible to convergence than non-salient ones (see Erker 2017: 76-77). In other words, “salience is only a necessary, but not a sufficient, condition for a linguistic feature to be affected in accommodation” (Auer et al. 1998:167). As Boswijk and Coler (2020) explain, “salience plays a role in the propagation of some forms, but not others.” This observation raises questions relating to the precise nature of the relationship between social meaning, salience and language change. Relatedly, Hinskens (1996) and Kerswill and Williams (2002) reflect on Trudgill's (1986) idea that more salient dialectal variants will be accommodated to, but that variants that have ‘extra-strong salience’ will not be accommodated to (see Trudgill 1986:125), and suggest that this conception is circular because “salience is sometimes used as an explanation for accommodation [...] and sometimes to explain why accommodation does *not* take place” (Hinskens 1996: 11, emphasis in original; see Kerswill and Williams 2002:89). There is currently no objective measure of salience or how it interacts with other extra-linguistic (e.g., social psychological, sociodemographic) factors which may influence convergence (see Kerswill and Williams 2002; Levon and Fox 2014; Preston 2010, 2011).

isolated refugee camps in Beirut (Hennessey 2011:1, 12, 19–20) was generally more favourable to maintaining conservative Palestinian vernacular norms.

Most notably, the results in this study show that with respect to the three targeted variables, speaker age (collapsed with educational level) played a key role in variant selection, where the majority of younger (third-generation) speakers were found to lead linguistic change. This is in line with previous findings in the literature on spoken Arabic, where younger speakers are said to spearhead changes, while older speakers may retain conservative speech norms (e.g., Cotter 2013, 2020; Cotter and Horesh 2015; Fityan 1981; Hennessey 2011).

Beyond the Arabic-speaking world, these results are consistent with previous findings about inter-generational differences in language use in other dialect/language contact situations, where the first-generation immigrants continue to speak the dialect of their place of origin and do not significantly restructure their grammars, as we saw in the case of the first-generation (older) Palestinians in this study. The second generation generally speaks dialects which show some influence from their parents' (first-generation immigrants) regional dialects. On the other hand, the third generation tends to assimilate to the local speech norms of the host country (see e.g., Fishman 1966; Hoffman and Walker 2010; Trudgill 1986, 2004). This reflects the case of the majority of the third-generation, educated Palestinians in the present study, whose aspirations of upward social mobility (see Fityan 1981:102), as well as their presumed regular face-to-face interaction with Lebanese peers in post-secondary institutions, explain why they diverge from their elders and opt for the majority or unmarked forms characteristic of the Lebanese dialect of Beirut, at least in the case of word-medial *imala* (i.e., the [e:] variant) and verbal negation (i.e., the preverbal negator *ma*). Younger speakers also generally avoid forms that may be at variance

with aspirations to educatedness as well as orientations towards urban speech patterns (see Fityan 1981; Hennessey 2011; Mousa 2019:16).

7.6 Limitations of the study and future lines of inquiry

Even though the potential impacts of the present study are multi-faceted, it is important to acknowledge their limitations. One of the main limitations is the bias in the speaker sample towards educated individuals, especially among the younger speakers. These educated younger speakers are assumed to be in the vanguard of dialect contact on account of their mobility (Al-Wer 2002:42, 51). Because there are no younger speakers with lower levels of education in the population sample for this study, we cannot be sure of how those speakers behave. More conservative PA speech patterns may be expected among less educated younger speakers, especially camp-dwellers, because they presumably have limited opportunities to regularly interact with Lebanese speakers, unlike their more educated counterparts, especially those living outside the camps (see e.g., Hennessey 2011:111). Collecting data from less educated younger Palestinians who live inside and outside the camps and examining their speech behaviour remain an important future line of inquiry.

In terms of the linguistic variables targeted, even though the present study found that PA as spoken in Beirut may be undergoing contact-induced change, particularly with respect to word-medial *imala* and verbal negation, while the expression of futurity appears to be less amenable to convergent change, we cannot rule out that (contact-induced) change has not taken place in other areas of the speakers' grammar, unless we extend the investigation to additional (grammatical) variables. Accountable analyses of more variables in natural speech data may help us make more definitive claims about the effect of long-term contact with LA on patterns of linguistic variation and change in PA in different variable systems and sectors of the grammar.

In addition to examining additional grammatical variables, it is worth extending the analysis to other understudied areas and urban centres in Lebanon, as the results reported in this study are only applicable to the urban centre of Beirut and cannot necessarily be generalized to other Palestinian speech communities in Lebanon.¹⁵³ One urban centre that can be investigated is Sidon (also known as Saida) in Southern Lebanon, where there is reportedly a massive Palestinian population inside and outside refugee camps (UN-Habitat and UNICEF Lebanon 2019).¹⁵⁴

Another way we can build on the present study is to expand it on the qualitative level. As we saw in Chapter 4 on *imala*, there were some speakers whose linguistic behaviour was idiosyncratic. As such, it would be interesting to include a more fine-grained, qualitative analysis of subsets of the data (see e.g., Eckert 2000) in order to examine the potential multiple social indexations of word-medial *imala* and to investigate how socio-psychological factors, such as Palestinians' language attitudes and their perception of their identity (e.g., O'Connor 2018), as well as their in-group and out-group identifications may constrain their linguistic choices and contribute to (or hinder) the maintenance of PA speech norms. These extra-linguistic factors may highlight how Palestinians in Beirut manipulate linguistic variation to index multiple social meanings and to construct their social identities. A qualitative examination of the speech data would therefore generate a holistic view of the social embedding of variation and change in PA as spoken in Beirut.

¹⁵³ In other words, we cannot generalize the findings in this study from the context of Beirut to another context (whether inside or outside Lebanon, or whether involving Palestinians or not) because of the social variables associated with the contact situation in Beirut.

¹⁵⁴ UN-Habitat and UNICEF Lebanon (2019) report:
https://unhabitat.org/sites/default/files/2020/11/un-habitat-unicef_old_saida_neighbourhood_profile.pdf

Because this study showed that there is linguistic change in the three targeted variables in PA, and this change is being led by younger (third-generation) Palestinians, another suggestion for future research is the quantitative examination of the three variables targeted in the present study, as well as other variables from different grammatical sub-components, in the speech of Palestinian preadolescents (i.e., fourth-generation speakers) inside and outside refugee camps. This will allow us to investigate the acquisition of (rule-governed) variation in Arabic-speaking children, especially in dialect contact situations, as well as children's participation in patterns of variation and change observed in the Palestinian adult speech community in the present study. This future line of inquiry aligns with Roberts' (1994:177) argument that "the inclusion of children in variational research [is] important in formulating a complete picture of variation in the speech community."

Based on the discussion above, the present research can be extended to children aged between 5 and 12 years old, especially because the stage of vernacular reorganization begins around the age of 5 (Labov 2001). During this stage, children shift their grammars from parents' (or primary caregivers') input to a peer-oriented input (Labov 2001:415). In other words, children start speaking differently from their mothers or caretakers (Labov 2001:415) and they start getting influenced by (and accommodating to) the speech norms of their peer group (Kerswill 1996:196). Drawing on the comparative variationist framework, the results of the present study from the adult speech community can be used as a benchmark to contextualize children's speech patterns in relation to wider community norms. Such research would contribute to empirical characterizations of grammatical and phonological variation in the speech of preadolescents, as well as the acquisition of adult speech community norms by children in the Arabic-speaking world.

7.7 Epilogue

In conclusion, it is hoped that the results of the present study may contribute to a further understanding of the process of language change in spoken Arabic, especially in Palestinian Arabic in contact settings, and to foregrounding the importance of inspecting the fine-grained conditioning of variation in assessing language change in spoken Arabic, whether contact-induced or internally-motivated. The results of this study provide further empirical support for previous claims and findings that long-term contact between typologically similar languages or varieties may not necessarily lead to structural convergence in different sectors of the grammar, and that surface similarities between languages or dialects can hide structural differences, as we saw in the case of the variable expression of the future in LA and PA.

On a broader level, the present research has potential real-world impacts, as it makes new contributions to scientifically grounded descriptions of spoken Arabic. Accordingly, the results of this research are of direct interest not only to Arabic (socio)linguists, but are also of potential relevance to language professionals in the Arabic-speaking world and educators (or teachers of Arabic) who have a vested interest in access to socially realist accounts of spoken Arabic.

In the Arabic-speaking world, the only sociolinguistic notion that has been pertinent to applied linguistics (e.g., language teaching) is that of ‘diglossia’, which typically refers to the co-existence of two languages in some communities, one is the written standard used for high culture functions, and another is the spoken variety that is generally used in informal settings (Haeri 2003:159). In the court of public opinion, there is only one ‘legitimate’ variety of Arabic, popularly equated with the written standard, Classical Arabic, which is imbued with considerable social prestige. Spoken varieties of Arabic, by contrast, acquired natively by millions, are often construed as corrupt or degenerate versions of the written language, if not entirely dismissed as

being unworthy of academic study (see Haeri 2003). The present study confronts these ubiquitous prejudices by incorporating an empirically accountable sociolinguistic framework for the analysis of two Levantine Arabic varieties, showing that LA and PA are inherently variable, rule-governed and structured. Only more accountable analyses of the type pursued in the present investigation will help to breach the gap between highly idealised accounts of Arabic and the structured variability inherent in Arabic spoken on the ground in everyday social interactions.

References

- Abdel-Hafiz, Ahmed-Sokarno. 2005. "The Development of Future Markers in Arabic and the Nile Nubian Languages." *Journal of Arabic and Islamic Studies* 6:64–79.
- Abdel-Jawad, Hassan Rashid E. 1981. "Lexical and Phonological Variation in Spoken Arabic in Amman." PhD dissertation, University of Pennsylvania, Philadelphia, Pennsylvania.
- Abdel-Jawad, Hassan Rashid E. 1986. "The Emergence of an Urban Dialect in the Jordanian Urban Centers." *International Journal of the Sociology of Language* 61(1):53–63.
- Abdel-Jawad, Hassan Rashid E. 1987. "Cross-Dialectal Variation in Arabic: Competing Prestigious Forms." *Language in Society* 16(3):356–68.
- Abo Mokh, Noor, and Stuart Davis. 2020. "What Triggers *ʾImāla*: Focus on a Palestinian Variety with Phonological Analysis." Pp. 35–53 in *Perspectives on Arabic Linguistics XXXII: Papers Selected from the Annual Symposium on Arabic Linguistics, Tempe, Arizona, 2018*, edited by E. van Gelderen. Amsterdam: John Benjamins.
- Aboul-Fetouh, Hilmi. 1969. *A Morphological Study of Egyptian Colloquial Arabic*. The Hague: Mouton.
- AbuAmsa, Duaa. 2016. "The Future Marker in Palestinian Arabic: An Internal or Contact-Induced Change?" Pp. 1–15 in *Proceedings of the 2016 Annual Conference of the Canadian Linguistic Association*, edited by L. Hracs. University of Calgary.
- Abu-Haidar, Farida. 1979. *A Study of the Spoken Arabic of Baskinta*. Leiden: Brill.

- Abu-Haidar, Farida. 1991. *Christian Arabic of Baghdad*. Wiesbaden: Otto Harrassowitz.
- Abu-Haidar, Farida. 2005. "Baghdad Arabic." Pp. 222–31 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Al-Ani, Salman H. 1970. *Arabic Phonology: An Acoustical and Physiological Investigation*. Vol. 61. Berlin: Walter de Gruyter.
- Al-Bataineh, Hussein. 2019. "Emphasis Harmony in Arabic: A Critical Assessment of Feature-Geometric and Optimality-Theoretic Approaches." *Languages* 4(4):79–102.
- Al-Essa, Aziza. 2009. "When Najd Meets Hijaz: Dialect Contact in Jeddah." Pp. 203–22 in *Arabic Dialectology: In Honour of Clive Holes on the Occasion of His Sixtieth Birthday*, edited by E. Al-Wer and R. de Jong. Amsterdam: Brill.
- Al-Hashmi, Shadiya. 2016. "The Phonetics and Phonology of Arabic Loanwords in Turkish: Residual Effects of Gutturals." PhD dissertation, University of York, York, England.
- Al-Jallad, Ahmad. 2020. "Pre-Islamic Arabic." Pp. 37–55 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Alkalesi, Yasin M. 2006. *Modern Iraqi Arabic*. 2nd ed. Washington, D.C.: Georgetown University Press.
- Al-Khatib, Mahmoud. 1988. "Sociolinguistic Change in an Expanding Urban Center: A Case Study of Irbid City, Jordan." PhD dissertation, Durham University, Durham, England.

- Al-Momani, Islam M. 2011. "The Syntax of Sentential Negation in Jordanian Arabic." *Theory and Practice in Language Studies* 1(5):482–96.
- Al-Najjar, Balkees. 1991. "Grammaticalization of Lexical Markers in Kuwaiti Arabic." *Folia Linguistica* 25(3–4):665–76.
- Al-Nassir, Abdulmun'im Abdulamir. 1985. "Sibawayh the Phonologist: A Critical Study of the Phonetic and Phonological Theory of Sibawayh as Presented in His Treatise *AL KITAB*." PhD dissertation, University of York, York, England.
- Al-Nassir, Abdulmun'im Abdulamir. 1993. *Sibawayh the Phonologist*. London: Kegan Paul.
- Alqassas, Ahmad. 2012. "The Morpho-Syntax and Pragmatics of Levantine Arabic Negation: A Synchronic and Diachronic Analysis." PhD dissertation, Indiana University, Bloomington and Indianapolis, Indiana.
- Alqassas, Ahmad. 2019. *A Multi-Locus Analysis of Arabic Negation: Micro-Variation in Southern Levantine, Gulf and Standard Arabic*. Edinburgh: Edinburgh University Press.
- Alsarayreh, Atef. 2012. "The Licensing of Negative Sensitive Items in Jordanian Arabic." PhD dissertation, University of Kansas, Kansas.
- Al-Sayyed, Amany, and David Wilmsen. 2017. "Verbal Negation with *muš* in Maltese and Eastern Mediterranean Arabics." Pp. 151–72 in *Advances in Maltese Linguistics*, edited by B. Saade and M. Tosco. Berlin, Boston: De Gruyter Mouton.

- Alshboul, Sabri, Yousef Al Shaboul, and Sahail M. Asassfeh. 2010. "Grammaticalization Patterns: Evidence from Future Markers in Jordanian Arabic." *Journal of the Australasian Universities Language and Literature Association* 114:99–110.
- Altairi, Hamed, Jason Brown, Catherine Watson, and Bryan Gick. 2017. "Tongue Retraction in Arabic: An Ultrasound Study." Pp. 1–12 in *Proceedings of the 2016 Annual Meeting on Phonology*, edited by K. Jesney, C. O'Hara, C. Smith, and R. Walker. University of Southern California, Los Angeles, California: Linguistic Society of America.
- Altakhaineh, Abdel Rahman Mitib, and Aseel Zibin. 2018. "Verb + Verb Compound and Serial Verb Construction in Jordanian Arabic (JA) and English." *Lingua* 201:45–56.
- Al-Wer, Enam. 1991. "Phonological Variation in the Speech of Women from Three Urban Areas in Jordan." PhD dissertation, University of Essex, Colchester, England.
- Al-Wer, Enam. 1997. "Arabic Between Reality and Ideology." *International Journal of Applied Linguistics* 7(2):251–65.
- Al-Wer, Enam. 1999. "Why Do Different Variables Behave Differently? Data from Arabic." Pp. 38–57 in *Language and Society in the Middle East and North Africa: Studies in Variation and Identity*, edited by Y. Suleiman. Richmond, Surrey: Curzon Press.
- Al-Wer, Enam. 2002. "Jordanian and Palestinian Dialects in Contact: Vowel Raising in Amman." Pp. 63–79 in *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*, edited by M. C. Jones and E. Esch. Berlin: Mouton de Gruyter.

- Al-Wer, Enam. 2006. "Jordanian Arabic (Amman)." Pp. 505–17 in *Encyclopedia of Arabic Language and Linguistics*. Vol. II, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Al-Wer, Enam. 2007a. "Phonological Split." Pp. 605–7 in *Encyclopedia of Arabic Language and Linguistics*. Vol. III, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Al-Wer, Enam. 2007b. "The Formation of the Dialect of Amman: From Chaos to Order." Pp. 55–76 in *Arabic in the City: Issues in Dialect Contact and Language Variation*, edited by C. Miller, E. Al-Wer, D. Caubet, and J. C. E. Watson. London and New York: Routledge.
- Al-Wer, Enam. 2011. "Jordanian and Palestinian Dialects in Contact: Vowel Raising in Amman." Pp. 63–80 in *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*, edited by M. C. Jones and E. Esch. Berlin: De Gruyter Mouton.
- Al-Wer, Enam. 2020. "New Dialect Formation: The Amman Dialect." Pp. 551–66 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Al-Wer, Enam, and Khairia Al-Qahtani. 2016. "Lateral Fricative ḍād in Tihāmat Qaḥṭān: A Quantitative Sociolinguistic Investigation." Pp. 151–69 in *Perspectives on Arabic Linguistics XXVII: Papers from the Annual Symposium on Arabic Linguistics, Bloomington, Indiana, 2013*, edited by S. Davis and U. Soltan. Amsterdam: John Benjamins.

- Al-Wer, Enam, Uri Horesh, Deema Alammam, Hind Alaodini, Aziza Al-Essa, Areej Al-Hawamdeh, Khairia Al-Qahtani, and Abeer Ab Hussain. 2020. "Probing Linguistic Change in Arabic Vernaculars: A Sociohistorical Perspective." *Language in Society* 51(1):1–22.
- Al-Wer, Enam, and Rudolf de Jong. 2018. "Dialects of Arabic." Pp. 523–34 in *The Handbook of Dialectology*, edited by C. Boberg, J. Nerbonne, and D. Watt. Oxford: Wiley-Blackwell.
- Amara, Muhammad. 2005. "Language, Migration, and Urbanization: The Case of Bethlehem." *Linguistics* 43(5):883–901.
- Amiri, Zeinab. 2016. "Boundaries and Political Agency of Palestinian Refugee Camps in Lebanon." MA thesis, Iowa State University, Ames, Iowa.
- Anera. 2021. "Where We Work: Lebanon." *Anera*. Retrieved January 9, 2022 (<https://www.anera.org/where-we-work/lebanon/>).
- Aoun, Joseph, Elabbas Benmamoun, and Lina Choueiri. 2010. *The Syntax of Arabic*. Cambridge: Cambridge University Press.
- Armstrong, Nigel. 1998. "The Sociolinguistic Gender Pattern in French: A Comparison of Two Linguistic Levels." *Journal of French Language Studies* 8(2):139–58.
- Arnold, Werner. 2005. "Antiochia Arabic." Pp. 111–19 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Auer, Anita, Catharina Peersman, Simon Pickl, Gijsbert Rutten, and Rik Vosters. 2015. "Historical Sociolinguistics: The Field and its Future." *Journal of Historical Sociolinguistics* 1(1):1–12.
- Auer, Peter, Birgit Barden, and Beate Grosskopf. 1998. "Subjective and Objective Parameters in Determining 'Salience' in Long-Term Dialect Accommodation." *Journal of Sociolinguistics* 2(2):163–87.
- Awwad, Mohammad Amin. 1987. "Free and Bound Pronouns as Verbs in Rural Palestinian Colloquial Arabic." *Zeitschrift Für Arabische Linguistik* 16:108–18.
- Ayres-Bennett, Wendy. 2000. "Voices from the Past: Sources of Spoken Seventeenth-Century French." *Romanische Forschungen* 112(3):328–48.
- Ayres-Bennett, Wendy. 2004. *Sociolinguistic Variation in Seventeenth-Century France: Methodology and Case Studies*. Cambridge: Cambridge University Press.
- Backus, Ad. 2005. "Code-Switching and Language Change: One Thing Leads to Another?" *International Journal of Bilingualism* 9(3–4):307–40.
- Baldi, Philip. 1990. "Introduction: The Comparative Method." Pp. 1–13 in *Linguistic Change and Reconstruction Methodology*, edited by P. Baldi. Berlin: Mouton de Gruyter.
- Barkat, Melissa. 2008. "Vowel Raising." Pp. 678–82 in *Encyclopedia of Arabic Language and Linguistics*. Vol. Vol IV, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Bauer, Leonhard. 1926. *Das Palästinische Arabisch: Die Dialekte des Städters und des Fellachen. Grammatik, Übungen und Chrestomathie Dargestellt*. Leipzig: Hinrichs.
- Baxter, Gareth J., Richard A. Blythe, William Croft, and Alan J. McKane. 2009. "Modeling Language Change: An Evaluation of Trudgill's Theory of the Emergence of New Zealand English." *Language Variation and Change* 21(2):257–96.
- Bayley, Robert. 2013. "The Quantitative Paradigm." Pp. 85–107 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Bazzi, Youssef. 2014. "Ziad Rahbani: His Life as a Performance." *The New Arab*, October 11.
- Behnstedt, Peter. 1985. *Die Nordjemenitischen Dialekte*. Vol. 1. Wiesbaden-Dotzheim: Reichert.
- Behnstedt, Peter. 1997. *Sprachatlas von Syrien*. Vol. 1. Wiesbaden: Otto Harrassowitz.
- Behnstedt, Peter, and Manfred Woidich. 2005. *Arabische Dialektgeographie: Eine Einführung*. Leiden: Brill.
- Behnstedt, Peter, and Manfred Woidich. 2013. "Arabic Dialectology." Pp. 300–325 in *The Oxford Handbook of Arabic Linguistics*, edited by J. Owens. Oxford: Oxford University Press.
- Benmamoun, Elabbas. 1992. "Inflectional and Functional Morphology: Problems of Projection, Representation and Derivation." PhD dissertation, University of Southern California, Los Angeles, California.

Benmamoun, Elabbas. 2000. *The Feature Structure of Functional Categories: A Comparative Study of Arabic Dialects*. Oxford: Oxford University Press.

Benmamoun, Elabbas, Mahmoud Abunasser, Rania Al-Sabbagh, Abdelaadim Bidaoui, and Dana Shalash. 2014. "Variations on the Same Theme: Sentential Negation and the Negative Copula in Arabic." Pp. 121–38 in *Perspectives on Arabic Linguistics XXIV–XXV: Papers from the Annual Symposia on Arabic Linguistics. Texas, 2010 and Arizona, 2011*, edited by S. Farwaneh and H. Ouali. Amsterdam: John Benjamins.

Bergs, Alexander. 2012. "The Uniformitarian Principle and the Risk of Anachronisms in Language and Social History." Pp. 80–98 in *The Handbook of Historical Sociolinguistics*, edited by J. M. Hernández-Campoy and J. C. Conde-Silvestre. Oxford: Wiley-Blackwell.

Blanc, Haim. 1953. *Studies in North Palestinian Arabic: Linguistic Inquiries Among the Druzes of Western Galilee and Mt. Carmel*. Jerusalem: Israel Oriental Society.

Blau, Josua. 1960. *Syntax des Palastinensischen Bauerndialekts von Bir-Zet*. Walldorf: Verlag für Orientkunde.

Bocco, Riccardo. 2009. "Unrwa and the Palestinian Refugees: A History Within History." *Refugee Survey Quarterly* 28(2–3):229–52.

- Boersma, Paul, and David Weenink. 2016. "Praat: Doing Phonetics by Computer." Retrieved March 1, 2019 (<https://www.fon.hum.uva.nl/praat/>).
- Borg, Alexander. 2005. "Cypriot Maronite Arabic." Pp. 536–43 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Boswijk, Vincent, and Matt Coler. 2020. "What is Saliency?" *Open Linguistics* 6(1):713–22.
- Bourdieu, Pierre. 1991. *Language and Symbolic Power*. Cambridge, MA: Harvard University Press.
- Britain, David. 2003. "Exploring the Importance of the Outlier in Sociolinguistic Dialectology." Pp. 191–208 in *Social Dialectology: In Honour of Peter Trudgill*, edited by D. Britain and J. Cheshire. Amsterdam: John Benjamins.
- Britain, David. 2004. "Space and Spatial Diffusion." Pp. 603–37 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers, P. Trudgill, and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Britain, David. 2010a. "Language and Space: The Variationist Approach." Pp. 142–62 in *Language and space: An international handbook of linguistic variation. Vol. 1: Theories and methods*, edited by P. Auer and J. E. Schmidt. Berlin: de Gruyter.
- Britain, David. 2010b. "Supralocal Regional Dialect Levelling." Pp. 193–204 in *Language and Identities*, edited by C. Llamas and D. Watt. Edinburgh: Edinburgh University Press.

- Britain, David. 2018. "Dialect Contact and New Dialect Formation." Pp. 143–58 in *The Handbook of Dialectology*, edited by C. Boberg, J. Nerbonne, and D. Watt. Oxford: Wiley-Blackwell.
- Britain, David, and Peter Trudgill. 2005. "New Dialect Formation and Contact- Induced Reallocation: Three Case Studies from the English Fens." *International Journal of English Studies* 5(1):183–209.
- Brustad, Kristen. 2000. *The Syntax of Spoken Arabic: A Comparative Study of Moroccan, Egyptian, Syrian, and Kuwaiti Dialects*. Washington, D.C.: Georgetown University Press.
- Bullock, Barbara E., and Almeida Jacqueline Toribio. 2004. "Introduction: Convergence as an Emergent Property in Bilingual Speech." *Bilingualism: Language and Cognition* 7(2):91–94.
- Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2002. "Cognitive Processes in Grammaticalization." Pp. 145–67 in *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*. Vol. II, edited by M. Tomasello. New Jersey: Lawrence Erlbaum Associates Inc.
- Bybee, Joan. 2003. "Mechanisms of Change in Grammaticization: The Role of Frequency." Pp. 602–23 in *The Handbook of Historical Linguistics*, edited by B. D. Joseph and R. D. Janda. Oxford: Wiley-Blackwell.

- Bybee, Joan. 2011. "Usage-Based Theory and Grammaticalization." Pp. 69–78 in *The Oxford Handbook of Grammaticalization*, edited by B. Heine and H. Narrog. Oxford: Oxford University Press.
- Bybee, Joan L., and Östen Dahl. 1989. "The Creation of Tense and Aspect Systems in the Languages of the World." *Studies in Language* 13(1):51–103.
- Bybee, Joan L., William Pagliuca, and Revere D. Perkins. 1991. "Back to the Future." Pp. 17–58 in *Approaches to Grammaticalization: Volume Ii. Types of Grammatical Markers*, edited by E. C. Traugott and B. Heine. Amsterdam: John Benjamins.
- Bybee, Joan, and William Pagliuca. 1985. "Cross-Linguistic Comparison and the Development of Grammatical Meaning." Pp. 59–83 in *Historical Semantics - Historical Word-Formation*, edited by J. Fisiak. Berlin, New York: De Gruyter Mouton.
- Bybee, Joan, and William Pagliuca. 1987. "The Evolution of Future Meaning." Pp. 109–22 in *Papers from the 7th International Conference on Historical Linguistics*, edited by A. Giacalone-Ramat, O. Carruba, and G. Bernini. Amsterdam: John Benjamins.
- Bybee, Joan, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: The University of Chicago Press.
- Campbell, Lyle. 1998. *Historical Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.
- Cantineau, Jean. 1946. *Les Parlers Arabes du Hōrān*. Paris: Librairie C. Klincksieck.

- Cantineau, Jean. 1960. *Études de Linguistique Arabe*. Paris: C. Klincksieck.
- Card, Elizabeth Anne. 1983. "A Phonetic and Phonological Study of Arabic Emphasis." PhD dissertation, Cornell University, Ithaca, NY.
- Chalala, Elie. 2018. "Jalal Khoury (1933-2017): Brechtian Realist Forged by 1967 War, and the Birth of Modern Lebanese Theater." *Al Jadid Magazine*.
- Chambers, J. K. 2003. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Wiley-Blackwell.
- Chambers, J. K., and Peter Trudgill. 2004. *Dialectology*. 2nd ed. Cambridge: Cambridge University Press.
- Cheshire, Jenny. 1999. "Taming the Vernacular: Some Repercussions for the Study of Syntactic Variation and Spoken Grammar." *Cuadernos de Filología Inglesa* 8:59–80.
- Cheshire, Jenny. 2002. "Sex and Gender in Variationist Research." Pp. 423–43 in *The handbook of language variation and change*, edited by J. K. Chambers, P. Trudgill, and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Cheshire, Jenny, Paul Kerswill, and Ann Williams. 2005. "Phonology, Grammar, and Discourse in Dialect Convergence." Pp. 135–67 in *Dialect Change: Convergence and Divergence in European Languages*, edited by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press.
- Childs, Claire. 2019. "Interviewer Effects on the Phonetic Reduction of Negative Tags, *innit?*" *Journal of Pragmatics* 142:31–46.

- Choueiri, Lina. 2019. "Syntactic Variation." Pp. 185–200 in *The Routledge Handbook of Arabic Sociolinguistics*, edited by E. Al-Wer and U. Horesh. London and New York: Routledge.
- Cohen, David. 1975. *Le Parler Arabe des Juifs de Tunis: Étude Linguistique*. Berlin: Mouton de Gruyter.
- Cotter, William. 2013. "Dialect Contact and Change in Gaza City." MA thesis, University of Essex, Colchester, England.
- Cotter, William. 2016. "(q) as a Sociolinguistic Variable in the Arabic of Gaza City." Pp. 229–46 in *Perspectives on Arabic Linguistics XXVIII: Papers from the Annual Symposium on Arabic Linguistics, Gainesville, Florida, 2014*, edited by Y. A. Haddad and E. Potsdam. Amsterdam: John Benjamins.
- Cotter, William. 2020. "Dialect Contact and Phonological Change." Pp. 567–81 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Cotter, William, and Uri Horesh. 2015. "Social Integration and Dialect Divergence in Coastal Palestine." *Journal of Sociolinguistics* 19(4):460–83.
- Cotter, William, and Rudolf de Jong. 2019. "Regional Variation." Pp. 47–62 in *The Routledge Handbook of Arabic Sociolinguistics*, edited by E. Al-Wer and U. Horesh. London and New York: Routledge.
- Coupland, Nikolas. 1980. "Style-Shifting in a Cardiff Work-Setting." *Language in Society* 9(1):1–12.

- Coveney, Aidan. 1996. *Variability in Spoken French*. Exeter: Elm Bank.
- Cowell, Mark W. [1964] 2005. *A Reference Grammar of Syrian Arabic: Based on the Dialect of Damascus*. Washington, D.C.: Georgetown University Press.
- Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. London: Longman.
- Cukor-Avila, Patricia, and Guy Bailey. 2001. "The Effects of the Race of the Interviewer on Sociolinguistic Fieldwork." *Journal of Sociolinguistics* 5(2):254–70.
- Cukor-Avila, Patricia, and Guy Bailey. 2013. "Real Time and Apparent Time." Pp. 239–62 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Culpeper, Jonathan, and Merja Kytö. 2010. *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
- D'Alverny, André. 1970. *Petite Introduction au Parler Libanais*. 2nd ed. Beyrouth: Dar el-Machreq.
- D'Arcy, Alexandra, and Sali Tagliamonte. 2015. "Not Always Variable: Probing the Vernacular Grammar." *Language Variation and Change* 27(3):255–85.
- Dat, Mihai. 2005. "Deixis." Pp. 565–69 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Davis, Stuart. 1995. "Emphasis Spread in Arabic and Grounded Phonology." *Linguistic Inquiry* 26(3):465–98.
- Denis, Derek, Matt Hunt Gardner, Marisa Brook, and Sali Tagliamonte. 2019. "Peaks and Arrowheads of Vernacular Reorganization." *Language Variation and Change* 31(1):43–67.
- Digesto, Salvatore. 2019. "A Variationist Analysis of Subjunctive Variability Across Space and Time: From Contemporary Italian Back to Latin." PhD dissertation, University of Ottawa, Ottawa, Ontario.
- Doraï, Mohamed Kamel. 2010. "Palestinian Refugee Camps in Lebanon: Migration, Mobility and the Urbanization Process." Pp. 67–80 in *Palestinian Refugees: Identity, Space and Place in the Levant*, edited by A. Knudsen and S. Hanafi. London and New York: Routledge.
- Doss, Madiha. 2008. "Evolving Uses in Cairene Egyptian Arabic Negation Forms." *Estudios de Dialectología Norteafricana y Andalusí* 12:83–91.
- Driver, G. R. 1925. *A Grammar of the Colloquial Arabic of Syria and Palestine*. London: Probsthain & Co.
- Eckardt, Regine. 2011. "Grammaticalization and Semantic Change." Pp. 389–400 in *The Oxford Handbook of Grammaticalization*, edited by B. Heine and H. Narrog. Oxford: Oxford University Press.
- Eckert, Penelope. 2000. *Linguistic Variation as Social Practice*. Oxford: Wiley-Blackwell.

- Eid, Mushira. 1983. "The Copula Functions of Pronouns." *Lingua* 59(2–3):197–207.
- Eisele, John C. 1990. "Time Reference, Tense, and Formal Aspect in Cairene Arabic." Pp. 173–212 in *f*. Vol. 1, edited by M. Eid. Amsterdam: John Benjamins.
- Eisele, John C. 2005. "Aspect." Pp. 195–201 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Eksell, Kerstin. 2006. "Analytic Genitive." Pp. 82–85 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- El Masri, Yafa. 2020. "72 Years of Homemaking in Waiting Zones: Lebanon's "Permanently Temporary" Palestinian Refugee Camps." *Frontiers in Sociology* 5:1–13.
- El-Hassan, Shahir. 2007. "Mood (Arabic Dialects)." Pp. 262–69 in *Encyclopedia of Arabic Language and Linguistics*. Vol. 3, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden, The Netherlands: Brill.
- Elihay, J. 2004. *The Olive Tree Dictionary: A Transliterated Dictionary of Conversational Eastern Arabic (Palestinian)*. Jerusalem: Minerva Instruction & Consultation.
- Erker, Daniel. 2017. "Contact, Co-Variation and Sociolinguistic Salience: What Mister Rogers Knows about Language Change." *University of Pennsylvania Working Papers in Linguistics* 23(2):68–78.

- Esseesy, Mohssen. 2021. "Grammaticalization in Arabic." Pp. 277–308 in *The Cambridge Handbook of Arabic Linguistics*, edited by K. C. Ryding and D. Wilmsen. Cambridge: Cambridge University Press.
- Farwaneh, Samira. 2021. "Diglossia, Variation, and Structural Complexity." Pp. 155–80 in *The Cambridge Handbook of Arabic Linguistics*, edited by K. C. Ryding and D. Wilmsen. Cambridge: Cambridge University Press.
- Fasold, Ralph W. 1972. *Tense Marking in Black English: A Linguistic and Social Analysis*. Washington, D.C.: Center for Applied Linguistics.
- Fasold, Ralph W. 1984. *The Sociolinguistics of Society*. Oxford: Wiley-Blackwell.
- Fasold, Ralph W. 1990. *The Sociolinguistics of Language*. Oxford: Wiley-Blackwell.
- Fassi Fehri, Abdelkader. 1993. *Issues in the Structure of Arabic Clauses and Words*. Dordrecht: Kluwer.
- Feghali, Michel. 1928. *Syntaxe des Parlers Arabes Actuels du Liban*. Paris: Imprimerie Nationale.
- Fishman, Joshua A. 1966. "Language Maintenance and Language Shift: The American Immigrant Case within a General Theoretical Perspective." *Sociologus* 16(1):19–39.
- Fityan, Ahmed. 1981. "Subdialects in Contact: Palestinian Arabic in Baalbeck Al-Jalil Camp." MA thesis, American University of Beirut, Beirut, Lebanon.
- Fleisch, Henri. 1974. *Etudes d'Arabe Dialectal*. Beyrouth: Dar el-Machreq.

- Fought, Carmen. 2004. "Ethnicity." Pp. 444–72 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers, P. Trudgill, and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Ghandour, Hind. 2017. "Naturalised Palestinians in Lebanon: Experiences of Belonging, Identity and Citizenship." PhD dissertation, Swinburne University of Technology, Melbourne, Australia.
- Giles, Howard, and Peter Powesland. 1997. "Accommodation Theory." Pp. 232–39 in *Sociolinguistics: A Reader*, edited by N. Coupland and A. Jaworski. Basingstoke: Macmillan.
- Givón, Talmy. 1973. "The Time-Axis Phenomenon." *Language* 49(4):890–925.
- Government of Lebanon, and United Nations. 2021. *Lebanon Crisis Response Plan 2017 - 2021 (2021 Update)*. Lebanon: Government of Lebanon and the United Nations.
- Guy, Gregory. 1988. "Advanced VARBRUL Analysis." Pp. 124–36 in *Linguistic Change and Contact: NWAV-XVI (Texas Linguistic Forum 30)*, edited by K. Ferrara, B. Brown, K. Walters, and J. Baugh. Austin: Department of Linguistics, University of Texas.
- Habib, Rania. 2008. "New Model for Analyzing Sociolinguistic Variation: The Interaction of Social and Linguistic Constraints." PhD dissertation, University of Florida, Gainesville, Florida.
- Habib, Rania. 2010. "Towards Determining Social Class in Arabic-Speaking Communities and Implications for Linguistic Variation." *Sociolinguistic Studies* 4(1):175–200.

- Habib, Rania. 2012. “‘*Imala* and Rounding in a Rural Syrian Variety: Morphophonological and Lexical Conditioning.” *The Canadian Journal of Linguistics/La revue Canadienne de linguistique* 57(1):51–75.
- Haeri, Niloofar. 1987. “Male/Female Differences in Speech: An Alternative Interpretation.” Pp. 173–83 in *Variation in Language: NWAV-XV at Stanford: Proceedings of the Fifteenth Annual Conference on New Ways of Analyzing Variation*, edited by K. M. Denning, S. Inkelas, J. R. Rickford, and F. McNair-Knox. Stanford: Stanford University Press.
- Haeri, Niloofar. 1991. “Sociolinguistic Variation in Cairene Arabic: Palatalization and the ‘Qaf’ in the Speech of Men and Women.” PhD dissertation, University of Pennsylvania, Philadelphia, Pennsylvania.
- Haeri, Niloofar. 1996. *The Sociolinguistic Market of Cairo: Gender, Class and Education*. London and New York: Kegan Paul.
- Haeri, Niloofar. 2003. *Sacred Language, Ordinary People: Dilemmas of Culture and Politics in Egypt*. New York: Palgrave MacMillan.
- Haeri, Niloofar, and William Cotter. 2019. “Form and Ideology Revisited.” Pp. 243–58 in *The Routledge Handbook of Arabic Sociolinguistics*, edited by E. Al-Wer and U. Horesh. London and New York: Routledge.
- Håland, Eva Marie. 2011. “*miš yinfa* ‘ – A Change in Progress? A Study of Extended Usage of the Negation Marker *miš* in Cairene Arabic.” MA thesis, University of Oslo, Oslo, Norway.

- Hanafi, Sari. 2009. "Palestinian Refugee Camps in the Palestinian Territory: Territory of Exception and Locus of Resistance." Pp. 495–519 in *The Power of Inclusive Exclusion: Anatomy of Israeli Rule in the Occupied Palestinian Territories*, edited by A. Ophir, M. Givoni, and S. Hanafi. New York: Zone books.
- Harrell, Richard. 1962. *A Short Reference Grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.
- Haspelmath, Martin. 2016. "The Serial Verb Construction: Comparative Concept and Cross-Linguistic Generalizations." *Language and Linguistics* 17(3):291–319.
- Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford: Oxford University Press.
- Heine, Bernd. 2003. "Grammaticalization." Pp. 573–601 in *The Handbook of Historical Linguistics*, edited by B. D. Joseph and R. D. Janda. Oxford: Wiley-Blackwell.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Heine, Bernd, and Tania Kuteva. 2002a. "On the Evolution of Grammatical Forms." Pp. 376–97 in *The Transition to Language*, edited by A. Wray. Oxford: Oxford University Press.
- Heine, Bernd, and Tania Kuteva. 2002b. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd, and Tania Kuteva. 2007. *The Genesis of Grammar: A Reconstruction*. Oxford: Oxford University Press.

- Hellmuth, Sam. 2013. "Phonology." Pp. 45–70 in *The Oxford Handbook of Arabic Linguistics*, edited by J. Owens. Oxford: Oxford University Press.
- Henkin, Roni. 2007. "Negev Arabic." Pp. 360–69 in *Encyclopedia of Arabic Language and Linguistics*. Vol. III, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Hennessey, Andrew. 2011. "The Linguistic Integration of the Palestinian Refugees in Beirut: A Model for Analysis." MA thesis, American University of Beirut, Beirut, Lebanon.
- Herin, Bruno. 2019. "Traditional Dialects." Pp. 93–105 in *The Routledge handbook of Arabic sociolinguistics*, edited by E. Al-Wer and U. Horesh. London and New York: Routledge.
- Herzallah, Rukayyah. 1990. "Aspects of Palestinian Arabic Phonology: A Non-Linear Approach." PhD dissertation, Cornell University, Ithaca, NY.
- Hickey, Raymond. 2000. "Salience, Stigma and Standard." Pp. 57–72 in *The Development of Standard English, 1300–1800: Theories, Descriptions, Conflicts*, edited by L. Wright. Cambridge: Cambridge University Press.
- Hickey, Raymond. 2010. "Language Contact: Reconsideration and Reassessment." Pp. 1–28 in *The Handbook of Language Contact*, edited by R. Hickey. Oxford: Wiley-Blackwell.
- Hinskens, Frans. 1996. *Dialect Levelling in Limburg: Structural and Sociolinguistic Aspects*. Tübingen: Niemeyer.

- Hinskens, Frans, Peter Auer, and Paul Kerswill. 2005. "The Study of Dialect Convergence and Divergence: Conceptual and Methodological Considerations." Pp. 1–48 in *Dialect Change: Convergence and Divergence in European Languages*, edited by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press.
- Hoenigswald, Henry. 1960. *Language Change and Linguistic Reconstruction*. Chicago: University of Chicago Press.
- Hoffman, Michol F., and James A. Walker. 2010. "Ethnolects and the City: Ethnic Orientation and Linguistic Variation in Toronto English." *Language Variation and Change* 22(1):37–67.
- Holes, Clive. 1983. "Patterns of Communal Variation in Bahrain." *Language in Society* 12(4):443–57.
- Holes, Clive. 1986. "Communicative Function and Pronominal Variation in Bahraini Arabic." *Anthropological Linguistics* 28(1):10–30.
- Holes, Clive. 1990. *Gulf Arabic*. London and New York: Routledge.
- Holes, Clive. 2004. *Modern Arabic: Structures, Functions, and Varieties*. Washington, D.C.: Georgetown University Press.
- Holes, Clive. 2005. "Bahraini Arabic." Pp. 241–55 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Holes, Clive, ed. 2018. *Arabic Historical Dialectology: Linguistic and Sociolinguistic Approaches*. Oxford: Oxford University Press.
- Hopper, Paul J. 1991. "On Some Principles of Grammaticization." Pp. 17–36 in *Approaches to Grammaticalization: Volume I. Theoretical and Methodological Issues*, edited by E. C. Traugott and B. Heine. Amsterdam: John Benjamins.
- Hopper, Paul J., and Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.
- Horesh, Uri. 2014. "Phonological Outcomes of Language Contact in the Palestinian Arabic Dialect of Jaffa." PhD dissertation, University of Essex, Colchester, England.
- Horesh, Uri. 2020. "Palestinian Dialects and Identities Shifting Across Physical and Virtual Borders." *Multilingua* 40(5):647–73.
- Horesh, Uri, and William Cotter. 2015. "Sociolinguistics of Palestinian Arabic." in *Encyclopedia of Arabic Language and Linguistics Online*, edited by L. Edzard and R. de Jong. Leiden: Brill.
- Horesh, Uri, and William Cotter. 2016. "Current Research on Linguistic Variation in the Arabic-Speaking World." *Language and Linguistics Compass* 10(8):370–81.

- Hoyt, Frederick MacNeill. 2006. "Negative Concord and Restructuring in Palestinian Arabic: A Comparison of TAG and CCG Analyses." Pp. 49–56 in *Proceedings of the 8th International Workshop on Tree Adjoining Grammar and Related Formalisms*, edited by T. Becker and L. Kallmeyer. Sydney, Australia: Association for Computational Linguistics.
- Hoyt, Frederick MacNeill. 2010. "Negative Concord In Levantine Arabic." PhD thesis, The University of Texas at Austin, Austin, Texas.
- Hussein, Lutfi. 1990. *Serial Verbs in Colloquial Arabic. Working Paper*. 39. Ohio: The Ohio State University, Department of Linguistics.
- Ibrahim, Muhammad H. 1986. "Standard and Prestige Language: A Problem in Arabic Sociolinguistics." *Anthropological Linguistics* 28(1):115–26.
- Ingham, Bruce. 1994. *Najdi Arabic: Central Arabian*. Amsterdam: Benjamins.
- Institute for Palestine Studies. 1988. "Appendix D: Maps: Arab Villages Emptied and Jewish Settlements Established in Palestine, 1948-49." *Journal of Palestine Studies* 18(1):38–50.
- International Labour Organization (ILO). 2012. *Palestinian Employment in Lebanon - Facts and Challenges: Labour Force Survey Among Palestinian Refugees Living in Camps and Gatherings in Lebanon*. Geneva: International Labour Organization (ILO).
- Jarad, Najib Ismail. 2013. "The Evolution of the *b*-Future Marker in Syrian Arabic." *Lingua Posnaniensis* 55(1):69–85.

- Jarad, Najib Ismail. 2014. "The Grammaticalization of the Motion Verb 'raḥ' as a Prospective Aspect Marker in Syrian Arabic." *Al-'Arabiyya* 47:101–18.
- Jarad, Najib Ismail. 2017. "Grammaticalization in Emirati Arabic." *Arabica* 64(5–6):742–60.
- Jastrow, Otto O. 2006. "Iraq." Pp. 414–24 in *Encyclopedia of Arabic Language and Linguistics*. Vol. II, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Johnson, Daniel Ezra. 2009. "Getting Off the Goldvarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis." *Language and Linguistics Compass* 3(1):359–83.
- Johnson, Daniel Ezra. 2010. "Rbrul Manual." *Daniel Ezra Johnson*. Retrieved June 1, 2021 (http://www.danielezrajohnson.com/Rbrul_manual.html).
- Johnstone, T. M. 1967. *Eastern Arabian Dialect Studies*. London: Oxford University Press.
- Kaufman, Asher. 2006. "Between Palestine and Lebanon: Seven Shi'i Villages as a Case Study of Boundaries, Identities, and Conflict." *Middle East Journal* 60(4):685–706.
- Kelly, Niamh E. 2017. "A Phonetic Case Study of a Bidialectal Speaker of Lebanese and Palestinian Arabic." Presented at the Structural and Developmental Aspect of Bidialectalism, The Arctic University of Norway, Tromsø, Norway.
- Kerswill, Paul. 1996. "Children, Adolescents, and Language Change." *Language Variation and Change* 8(2):177–202.
- Kerswill, Paul. 2002. "Models of Linguistic Change and Diffusion: New Evidence from Dialect Levelling in British English." *Reading Working Papers in Linguistics* 6:187–216.

- Kerswill, Paul. 2003. "Dialect Levelling and Geographical Diffusion in British English." Pp. 223–43 in *Social Dialectology: In Honour of Peter Trudgill*, edited by D. Britain and J. Cheshire. Amsterdam: John Benjamins.
- Kerswill, Paul, and Peter Trudgill. 2005. "The Birth of New Dialects." Pp. 196–220 in *Dialect Change: Convergence and Divergence in European Languages*, edited by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press.
- Kerswill, Paul, and Ann Williams. 2002. "'Salience' as an Explanatory Factor in Language Change: Evidence from Dialect Levelling in Urban England." Pp. 81–110 in *Language Change: The Interplay of Internal, External and Extra-Linguistic Factors*, edited by M. Jones and E. Esch. Berlin: Mouton de Gruyter.
- Khairallah, Natalie. 2014. "Negation in the Lebanese Dialect of Zeitoun, Keserwan: An Examination of Claims, Concepts, and Usage." MA thesis, American University of Beirut, Beirut, Lebanon.
- Khalafallah, Abdelghany A. 1969. *A Descriptive Grammar of Sa'i:Di Egyptian Colloquial Arabic*. The Hague: Mouton.
- Khoury, Yvette. 2009. "Mansour Rahbani: Dramatist and Writer of the Classic Songs of a Lebanese Golden Age." *The Guardian*, April 17.
- Khuri, Fuad I. 1960. "Education as a Function of Social Stratification in Cedarstown." MA thesis, American University of Beirut, Beirut, Lebanon.

- Khuri, Fuad I. 1969. "The Changing Class Structure in Lebanon." *Middle East Journal* 23(1):29–44.
- Kulikov, Vladimir, Fatemeh Mohsenzadeh, and Rawand M. Syam. 2020. "Effect of Emphasis Spread on Coronal Stop Articulation in Qatari Arabic." Pp. 16–28 in *Proceedings of the Linguistic Society of America*. Vol. 5, edited by P. Farrell. New Orleans: Linguistic Society of America.
- Kyrtziz, Amy, and Susan Ervin-Tripp. 1999. "The Development of Discourse Markers in Peer Interaction." *Journal of Pragmatics* 31:1321–38.
- Labov, William. 1963. "The Social Motivation of a Sound Change." *Word* 19(3):273–309.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
- Labov, William. 1969. "Contraction, Deletion, and Inherent Variability of the English Copula." *Language* 45(4):715–62.
- Labov, William. 1972a. *Language in the Inner City*. Philadelphia, Pennsylvania: University of Pennsylvania Press.
- Labov, William. 1972b. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 1972c. "Some Principles of Linguistic Methodology." *Language in Society* 1(1):97–120.
- Labov, William. 1973. *Sample Questionnaire Used by the Project on Linguistic Change and Variation*. Philadelphia, Pennsylvania: University of Pennsylvania.

- Labov, William. 1978. "Where Does the Linguistic Variable Stop? A Response to Beatriz Lavandera." Pp. 4–21 in *Working Papers in Sociolinguistics*. Vol. 44. Austin (TX): Southwest Educational Development Laboratory.
- Labov, William. 1982. "Building on Empirical Foundations." Pp. 17–92 in *Perspectives on Historical Linguistics: Papers from a Conference Held at the Meeting of the Language Theory Division, Modern Language Assn., San Francisco, 27–30 December 1979*, edited by W. P. Lehmann and Y. Malkiel. Amsterdam: John Benjamins.
- Labov, William. 1984. "Field Methods of the Project on Linguistic Change and Variation." Pp. 28–54 in *Language in Use: Readings in Sociolinguistics*, edited by J. Baugh and J. Sherzer. Englewood Cliffs, NJ: Prentice-Hall.
- Labov, William. 1994. *Principles of Linguistic Change. Vol. 1: Internal Factors*. Oxford: Wiley-Blackwell.
- Labov, William. 2001. *Principles of Linguistic Change. Vol. 2: Social Factors*. Oxford: Wiley-Blackwell.
- Labov, William. 2004. "Quantitative Reasoning in Linguistics." Pp. 6–22 in *Sociolinguistics/Soziolinguistik: An International Handbook of the Science of Language and Society*. Vol. 1, edited by U. Ammon, N. Dittmar, K. J. Mattheier, and P. Trudgill. Berlin: De Gruyter Mouton.
- Labov, William. 2006. *The Social Stratification of English in New York City*. 2nd ed. Cambridge: Cambridge University Press.

- Labov, William. 2007. "Transmission and Diffusion." *Language* 83(2):344–87.
- Labov, William. 2010. *Principles of Linguistic Change. Vol. 3: Cognitive and Cultural Factors*. Oxford: Wiley-Blackwell.
- Labov, William. 2015. "The Discovery of the Unexpected." *Asia-Pacific Language Variation* 1(1):7–22.
- Lamiroy, Béatrice, and Walter De Mulder. 2011. "Degrees of Grammaticalization Across Languages." Pp. 302–17 in *The Oxford Handbook of Grammaticalization*, edited by B. Heine and H. Narrog. Oxford: Oxford University Press.
- Lane, Edward William. 1984. *Arabic-English Lexicon*. Vol. 2. Cambridge: Islamic Texts Society.
- Laraba, Ahmed. 1981. "A Linguistic Description of the Algerian Arabic Dialect of Constantine." PhD dissertation, University of Manchester, Manchester, England.
- Leddy-Cecere, Thomas. 2020. "Contact-Induced Grammaticalization Between Arabic Dialects." Pp. 603–23 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Leddy-Cecere, Thomas Alexander. 2018. "Contact-Induced Grammaticalization as an Impetus for Arabic Dialect Development." PhD dissertation, The University of Texas at Austin, Austin, Texas.
- Lentin, Jérôme. 2018. "The Levant." Pp. 170–205 in *Arabic Historical Dialectology: Linguistic and Sociolinguistic Approaches*, edited by C. Holes. Oxford: Oxford University Press.

- Levin, Aryeh. 1978. "The 'imāla of 'alif fā'il in Old Arabic." *Israel Oriental Studies* 8:174–203.
- Levin, Aryeh. 1992. "The Authenticity of Sibawayhi's Description of the 'imāla." *Jerusalem Studies in Arabic and Islam* 15:74–93.
- Levin, Aryeh. 1998. *Arabic Linguistic Thought and Dialectology*. Jerusalem: Institute of Asian and African Studies.
- Levin, Aryeh. 2006. "Imāla." Pp. 311–15 in *Encyclopedia of Arabic Language and Linguistics*. Vol. II, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Levon, Erez, and Sue Fox. 2014. "Social Salience and the Sociolinguistic Monitor: A Case Study of ING and TH-Fronting in Britain." *Journal of English Linguistics* 42(3):185–217.
- Lindholm Schulz, Helena. 2003. *The Palestinian Diaspora: Formation of Identities and Politics of Homeland*. London and New York: Routledge.
- Lucas, Christopher. 2007. "Jespersen's Cycle in Arabic and Berber." *Transactions of the Philological Society* 105(3):398–431.
- Lucas, Christopher. 2010. "Negative -š in Palestinian (and Cairene) Arabic: Present and Possible Past." *Brill's Journal of Afroasiatic Languages and Linguistics* 2(1):165–201.
- Lucas, Christopher. 2015. "On Wilmsen on the Development of Postverbal Negation in Dialectal Arabic." *SOAS Working Papers in Linguistics* 17:77–95.
- Lucas, Christopher. 2018. "On Wilmsen on the Development of Postverbal Negation in Dialectal Arabic." *Zeitschrift Für Arabische Linguistik* 67:44–70.

- Lucas, Christopher. 2020. "Contact and the Expression of Negation." Pp. 643–67 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Lucas, Christopher, and Stefano Manfredi. 2020. "Introduction." Pp. 1–33 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Mansour, Jacob. 2005. "Baghdad Arabic Jewish." Pp. 231–41 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.
- McCarthy, John J. 1994. "The Phonetics and Phonology of Semitic Pharyngeals." Pp. 191–233 in *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology Iii*, edited by P. Keating. Cambridge: Cambridge University Press.
- McCarthy, John J. 1997. "Process-Specific Constraints in Optimality Theory." *Linguistic Inquiry* 28(2):231–51.
- Meillet, Antoine. 1967. *The Comparative Method in Historical Linguistics*. Paris: Librairie Honoré Champion.
- Miller, Catherine. 2007. "Arabic Urban Vernaculars: Development and Change." Pp. 1–31 in *Arabic in the City: Issues in Dialect Contact and Language Variation*, edited by C. Miller, E. Al-Wer, D. Caubet, and J. C. E. Watson. London and New York: Routledge.

- Miller, Catherine, Enam Al-Wer, Dominique Caubet, and Janet C. E. Watson, eds. 2007. *Arabic in the City: Issues in Dialect Contact and Language Variation*. London and New York: Routledge.
- Milroy, James. 1992. *Linguistic Variation and Change*. Oxford: Wiley-Blackwell.
- Milroy, James, and Lesley Milroy. 2012. *Authority in Language: Investigating Standard English*. London and New York: Routledge.
- Milroy, James, Lesley Milroy, and Sue Hartley. 1994. "Local and Supra-Local Change in British English: The Case of Glottalisation." *English World-Wide* 15(1):1–32.
- Milroy, Lesley. 1980. *Language and Social Networks*. University Park, Pennsylvania: University Park Press.
- Milroy, Lesley. 2007. "Off the Shelf or Under the Counter? On the Social Dynamics of Sound Changes." Pp. 143–71 in *Studies in the History of the English Language Iii. Managing Chaos: Strategies for Identifying Change in English*, edited by C. M. Cain and G. Russom. Berlin: Mouton de Gruyter.
- Milroy, Lesley, and Matthew Gordon. 2003. *Sociolinguistics: Method and Interpretation*. Oxford: Wiley-Blackwell.
- Milroy, Lesley, and Carmen Llamas. 2013. "Social Networks." Pp. 409–27 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers and N. Schilling-Estes. Oxford: Wiley-Blackwell.

- Mitchell, T. F., and Shahir El-Hassan. 1994. *Modality, Mood and Aspect in Spoken Arabic with Special Reference to Egypt and the Levant*. London and New York: Kegan Paul International.
- Mitchell, Terence Frederick. 1956. *An Introduction to Egyptian Colloquial Arabic*. Oxford: Oxford University Press.
- Moore, Colette. 2011. *Quoting Speech in Early English*. Cambridge: Cambridge University Press.
- Morris, Benny. 2003. *The Birth of the Palestinian Refugee Problem Revisited*. 2nd ed. Cambridge: Cambridge University Press.
- Mousa, Neimeh. 2019. "Grammatical Aspects of Rural Palestinian Arabic." PhD dissertation, Arizona State University, Tempe, Arizona.
- Moutaouakil, Ahmed. 1993. *al-wathiifa wa l-binya*. Casablanca: 'ocaadh.
- Muysken, Pieter, and Tonjes Veenstra. 1995. "Serial Verbs." Pp. 289–301 in *Pidgins and Creoles: An Introduction*, edited by J. Arends, P. Muysken, and N. Smith. Amsterdam: John Benjamins.
- Naïm, Samia. 2005. "Beirut Arabic." Pp. 174–286 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Naïm, Samia. 2011. "Dialects of the Levant." Pp. 920–35 in *The Semitic Languages: An International Handbook*, edited by S. Weninger. Berlin: Mouton de Gruyter.

- Obler, Loraine. 1975. "Reflexes of the Classical Arabic šay'un 'thing' in the Modern Dialects: A Study in Patterns of Language Change." PhD dissertation, University of Michigan, Ann Arbor, Michigan.
- O'Connor, Dermot Paul. 2018. "Geographies of Youth Identity, Memory and Belonging in the Palestinian Refugee Camps of Lebanon." PhD dissertation, Ulster University, Coleraine, Northern Ireland.
- Ouhalla, Jamal. 1993. "Negation, Focus and Tense: The Arabic Maa and Laa." *Revisita Di Linguistica* 5:275–300.
- Owens, Jonathan. 1984. *A Short Reference Grammar of Eastern Libyan Arabic*. Wiesbaden, Germany: Otto Harrassowitz Verlag.
- Owens, Jonathan. 1993. "'Imāla' in Eastern Libyan Arabic." *Zeitschrift Für Arabische Linguistik* 25:251–59.
- Owens, Jonathan. 2006. *A Linguistic History of Arabic*. Oxford: Oxford University Press.
- Owens, Jonathan. 2013. "A House of Sound Structure, of Marvelous Form and Proportion: An Introduction." Pp. 1–22 in *The Oxford Handbook of Arabic Linguistics*, edited by J. Owens. Oxford: Oxford University Press.
- Owens, Jonathan. 2018. "Dialects (Speech Communities), the Apparent Past, and Grammaticalization: Towards an Understanding of the History of Arabic." Pp. 206–56 in *Arabic Historical Dialectology: Linguistic and Sociolinguistic Approaches*, edited by C. Holes. Oxford: Oxford University Press.

- Owens, Jonathan, Robin Dodsworth, and Mary Kohn. 2013. "Subject Expression and Discourse Embeddedness in Emirati Arabic." *Language Variation and Change* 25(3):255–85.
- Owens, Jonathan, Robin Dodsworth, and Trent Rockwood. 2009. "Subject-Verb Order in Spoken Arabic: Morpholexical and Event-Based Factors." *Language Variation and Change* 21(1):39–67.
- Palva, Heikki. 2004. "Negations in the Dialect of Es-Salt, Jordan." Pp. 221–36 in *Approaches to Arabic Dialects: A Collection of Articles Presented to Manfred Woidich on the Occasion of His Sixtieth Birthday*, edited by M. Haak, R. de Jong, and K. Versteegh. Leiden: Brill.
- Palva, Heikki. 2005. "Dialects: Classification." Pp. 604–13 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Parkinson, Dilworth B. 2003. "Future Variability: A Corpus Study of Arabic Future Particles." Pp. 191–211 in *Perspectives on Arabic Linguistics XV: Papers from the Fifteenth Annual Symposium on Arabic Linguistics, Salt Lake City 2001*, edited by D. B. Parkinson and S. Farwanah. Amsterdam: John Benjamins.
- Pereira, Christophe. 2009. "Tripoli Arabic." Pp. 548–56 in *Encyclopedia of Arabic Language and Linguistics*. Vol. IV, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Persson, Maria. 2008. "The Role of the *b*-Prefix in Gulf Arabic Dialects as a Marker of Future, Intent and/or Irrealis." *Journal of Arabic and Islamic Studies* 8:26–52.

- Piamenta, Moshe. 1966. *Studies in the Syntax of Palestinian Arabic*. Jerusalem: Brill.
- Poplack, Shana. 1982. "Bilingualism and the Vernacular." Pp. 1–24 in *Issues in International Bilingual Education, the Role of the Vernacular*, edited by A. Valdman and C. R. Foster. New York: Plenum Press.
- Poplack, Shana. 2021. "A Variationist Perspective on Language Contact." Pp. 46–62 in *Routledge Handbook of Language Contact*, edited by E. Adamou and Y. Matras. London and New York: Routledge.
- Poplack, Shana, and Nathalie Dion. 2009. "Prescription Vs. Praxis: The Evolution of Future Temporal Reference in French." *Language* 85(3):557–87.
- Poplack, Shana, and Stephen Levey. 2010. "Contact-Induced Grammatical Change." Pp. 391–419 in *Language and Space – an International Handbook of Linguistic Variation: Volume 1 – Theories and Methods*, edited by P. Auer and J. E. Schmidt. Berlin: Mouton de Gruyter.
- Poplack, Shana, and Elisabete Malvar. 2007. "Elucidating the Transition Period in Linguistic Change: The Expression of the Future in Brazilian Portuguese." *Probus* 19(1):121–69.
- Poplack, Shana, and Marjory Meechan. 1998. "Introduction: How Languages Fit Together in Codemixing." *International Journal of Bilingualism* 2(2):127–38.
- Poplack, Shana, and Anne St-Amand. 2007. "A Real-Time Window on Nineteenth-Century Vernacular French: The Récits du Français Québécois d'autrefois." *Language in Society* 36(5):707–34.

- Poplack, Shana, and Sali Tagliamonte. 1999. "The Grammaticization of *going to* in (African American) English." *Language Variation and Change* 11(3):315–42.
- Poplack, Shana, and Sali Tagliamonte. 2001. *African American English in the Diaspora*. Oxford: Wiley-Blackwell.
- Poplack, Shana, and Rena Torres Cacoullos. 2015. "Linguistic Emergence on the Ground: A Variationist Paradigm." Pp. 267–91 in *The Handbook of Language Emergence*, edited by B. MacWhinney and W. O'Grady. Oxford: Wiley-Blackwell.
- Poplack, Shana, and Danielle Turpin. 1999. "Does the *Futur* Have a Future in (Canadian) French?" *Probus* 11(1):133–64.
- Poplack, Shana, Lauren Zentz, and Nathalie Dion. 2012. "Phrase-Final Prepositions in Quebec French: An Empirical Study of Contact, Code-Switching and Resistance to Convergence." *Bilingualism: Language and Cognition* 15(2):203–25.
- Preston, Dennis. 2010. "Variation in Language Regard." Pp. 7–27 in *Empirische Evidenzen und theoretische Passungen sprachlicher Variation*, edited by P. Gilles, J. Scharloth, and E. Zeigler. Frankfurt: Peter Lang.
- Preston, Dennis. 2011. "The Power of Language Regard: Discrimination, Classification, Comprehension and Production." *Dialectologia*. Special Issue II. 9–33.
- Procházka, Stephan. 2005. "Cilician Arabic." Pp. 388–97 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Procházka, Stephan. 2020. "Arabic in Iraq, Syria, and Southern Turkey." Pp. 83–114 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Procházka, Stephan. 2021. "Arabic Dialectology." Pp. 214–43 in *The Cambridge Handbook of Arabic Linguistics*, edited by K. C. Ryding and D. Wilmsen. Cambridge: Cambridge University Press.
- van Putten, Marijn. 2020. "Classical and Modern Standard Arabic." Pp. 57–82 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.
- Reinhardt, Carl. 1894. *Ein Arabischer Dialekt Gesprochen in Omān und Zanzibar: Nach Praktischen Gesichtspunkten für das Seminar für Orientalische Sprachen in Berlin*. Berlin and Stuttgart: W. Spemann.
- Retsö, Jan. 2011. "Classical Arabic." Pp. 782–810 in *The Semitic Languages: An International Handbook*, edited by S. Weninger. Berlin: Mouton de Gruyter.
- Richter, Tonio Sebastian. 2005. "Coptic." Pp. 495–501 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Rickford, John R., and Faye McNair-Knox. 1994. "Addressee- and Topic-Influenced Style Shift: A Quantitative Sociolinguistic Study." Pp. 235–76 in *Perspectives on Register: Situating Register Variation Within Sociolinguistics*, edited by D. Biber and E. Finegan. Oxford: Oxford University Press.

- Rickford, John R., and Thomas A. Wasow. 1995. "Syntactic Variation and Change in Progress: Loss of the Verbal Coda in Topic-Restricting *as far as* Constructions." *Language* 71(1):102–31.
- Roberts, Julie. 1994. *Acquisition of Variable Rules: (-t,d) Deletion and (ing) Production in Preschool Children*. IRCS-96-09. Philadelphia: University of Pennsylvania: Institute for Research in Cognitive Science (IRCS).
- Rogers, Henry. 2000. *The Sounds of Language: An Introduction to Phonetics*. London and New York: Routledge.
- Romaine, Suzanne. 1982. *Socio-Historical Linguistics: Its Status and Methodology*. Cambridge: Cambridge University Press.
- Romaine, Suzanne. 1988. "Historical Sociolinguistics: Problems and Methodology." Pp. 1452–69 in *Sociolinguistics: An International Handbook of the Sciences of Language and Society*. Vol. 2, edited by U. Ammon, N. Dittmar, and K. J. Mattheier. Berlin, New York: W. de Gruyter.
- Romaine, Suzanne. 2003. "Variation in Language and Gender." Pp. 98–118 in *The Handbook of Language and Gender*, edited by J. Holmes and M. Meyerhoff. Oxford: Wiley-Blackwell.
- Rosenhouse, Judith. 2006. "Jerusalem Arabic." Pp. 481–93 in *Encyclopedia of Arabic Language and Linguistics*. Vol. II, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Rosenhouse, Judith. 2011. "Jerusalem Arabic." in *Encyclopedia of Arabic Language and Linguistics*, edited by L. Edzard and R. de Jong. Leiden: Brill.
- Russell, Joan. 1982. "Networks and Sociolinguistic Variation in an African Urban Setting." Pp. 125–40 in *Sociolinguistic Variation in Speech Communities*, edited by S. Romaine. London: Edward Arnold.
- Ryding, Karin C. 2005. *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.
- Salem, Elise. 1999. "Imagining Lebanon Through Rahbani Musicals." *Al Jadid Magazine*.
- Salmon, Vivian. [1965] 1987. "Sentence Structures in Colloquial Shakespearian English." Pp. 265–300 in (*First published in*) *Transactions of the Philological Society, Vol. 64*, edited by Salmon and Burness.
- Sankoff, David. 1988. "Sociolinguistics and Syntactic Variation." Pp. 140–61 in *Linguistics: The Cambridge Survey*, edited by F. J. Newmeyer. Cambridge: Cambridge University Press.
- Sankoff, Gillian. 2001. "Linguistic Outcomes of Language Contact." Pp. 638–68 in *The Handbook of Language Variation and Change*, edited by P. Trudgill, J. K. Chambers, and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Sankoff, Gillian. 2005. "Cross-Sectional and Longitudinal Studies in Sociolinguistics." Pp. 1003–13 in *Sociolinguistics: An International Handbook of the Science of Language and Society, Vol. 1*, edited by U. Ammon, N. Dittmar, K. J. Mattheier, and P. Trudgill. Berlin: Mouton de Gruyter.

- Sankoff, Gillian, and Hélène Blondeau. 2007. "Language Change Across the Lifespan: /r/ in Montreal French." *Language* 83(3):560–88.
- Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace & Co.
- Sayigh, Rosemary. 2005. "A House Is Not a Home: Permanent Impermanence of Habitat for Palestinian Expellees in Lebanon." *Holy Land Studies* 4(1):17–39.
- Schilling-Estes, Natalie. 2007. "Sociolinguistic Fieldwork." Pp. 165–89 in *Sociolinguistic Variation: Theory, Methods and Application*, edited by R. Bayley and C. Lucas. Cambridge: Cambridge University Press.
- Schilling-Estes, Natalie. 2013. *Sociolinguistic Fieldwork*. Cambridge: Cambridge University Press.
- Schmidt, Hans, and Paul Kahle. 1918. *Volkserzählungen aus Palästina V. I*. Göttingen: Vandenhoeck and Ruprecht.
- Schneider, Edgar W. 2013. "Investigating Historical Variation and Change in Written Documents: New Perspectives." Pp. 57–81 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Schwenter, Scott A., and Rena Torres Cacoullos. 2008. "Defaults and Indeterminacy in Temporal Grammaticalization: The 'Perfect' Road to Perfective." *Language Variation and Change* 20(1):1–39.

- Seeger, Ulrich. 1996. "Der Arabische Dialekt von Il-Xalīl (Hebron)." Magister dissertation, Universität Heidelberg, Heidelberg, Germany.
- Shahin, Kimary. 2002. *Postvelar Harmony*. Amsterdam: John Benjamins.
- Shahin, Kimary. 2011. "Pharyngeals." Pp. 604–27 in *The Blackwell Companion to Phonology*, edited by M. Van Oostendorp, C. J. Ewen, E. Hume, and K. Rice. Oxford: Wiley-Blackwell.
- Shahin, Kimary N. 2000. *Rural Palestinian Arabic (Abu Shusha Dialect)*. Munich: Lincom Europa.
- Shahin, Kimary N. 2007. "Palestinian Arabic." Pp. 526–38 in *Encyclopedia of Arabic Language and Linguistics*. Vol. III, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Shlonsky, Ur. 1997. *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. Oxford: Oxford University Press.
- Siegel, Jeff. 2010. *Second Dialect Acquisition*. Cambridge: Cambridge University Press.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Clarendon.
- Simeone-Senelle, Marie-Claude. 2006. "Horn of Africa." Pp. 268–75 in *Encyclopedia of Arabic Language and Linguistics*. Vol. II, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

- Soltan, Usama. 2007. "On Formal Feature Licensing in Minimalism: Aspects of Standard Arabic Morphosyntax." PhD dissertation, University of Maryland, College Park, Maryland.
- Soltan, Usama. 2014. "Splitting Neg: The Morphosyntax of Sentential Negation in Cairene Egyptian Arabic Revisited." Pp. 91–119 in *Perspectives on Arabic Linguistics XXVI: Papers from the Annual Symposium on Arabic Linguistics. New York, 2012*, edited by R. Khamis-Dakwar and K. Froud. Amsterdam: John Benjamins.
- Soltan, Usama. 2021. "On Arabic Morphosyntax and its Implications for the Theory of Generative Grammar." Pp. 371–404 in *The Cambridge Handbook of Arabic Linguistics*, edited by K. C. Ryding and D. Wilmsen. Cambridge: Cambridge University Press.
- Souag, Lameen. 2005. "Notes on the Algerian Arabic Dialect of Dellys." *Estudios de Dialectología Norteafricana y Andalusí* 9:151–80.
- Stewart, Devin J. 1998. "Clitic Reduction in the Formation of Modal Prefixes in the Post-Classical Arabic Dialects and Classical Arabic *Sa-/Sawfa*." *Arabica* 45(1):104–28.
- Suleiman, Jaber. 2010. "Trapped Refugees: The Case of Palestinians in Lebanon." *Refugee Studies Centre Working Paper Series* 64:7–35.
- Tagliamonte, Sali. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Tagliamonte, Sali. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.

- Tagliamonte, Sali. 2013. "Comparative Sociolinguistics." Pp. 128–56 in *The Handbook of Language Variation and Change*, edited by J. K. Chambers and N. Schilling-Estes. Oxford: Wiley-Blackwell.
- Tagliamonte, Sali, and Alexandra D'Arcy. 2009. "Peaks Beyond Phonology: Adolescence, Incrementation, and Language Change." *Language* 85(1):58–108.
- Tagliamonte, Sali, Mercedes Durham, and Jennifer Smith. 2014. "Grammaticalization at an Early Stage: Future *be going to* in Conservative British Dialects." *English Language & Linguistics* 18(1):75–108.
- Thackston, Wheeler McIntosh. 1996. *The Vernacular Arabic of the Lebanon*. Cambridge: Department of Near Eastern Languages and Civilizations, Harvard University.
- Thomason, Sarah. 2001. *Language Contact*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah. 2010. "Contact Explanations in Linguistics." Pp. 1–28 in *The Handbook of Language Contact*, edited by R. Hickey. Oxford: Wiley-Blackwell.
- Thomason, Sarah, and Terrence Kaufman. 1988. *Language Contact, Creolization and Genetic Linguistics*. Los Angeles: University of California Press.
- Torres Cacoullos, Rena, and James A. Walker. 2009. "The Present of the English Future: Grammatical Variation and Collocations in Discourse." *Language* 85(2):321–54.
- Traugott, Elizabeth Closs. 2010. "Grammaticalization." Pp. 271–85 in *The Bloomsbury Companion to Historical Linguistics*, edited by S. Luraghi and V. Bubenik. London: Bloomsbury Academic.

- Traugott, Elizabeth Closs. 2011. "Grammaticalization and Mechanisms of Change." Pp. 19–30 in *The Oxford Handbook of Grammaticalization*, edited by B. Heine and H. Narrog. Oxford: Oxford University Press.
- Traugott, Elizabeth Closs, and Ekkehard König. 1991. "The Semantics-Pragmatics of Grammaticalization Revisited." Pp. 189–218 in *Approaches to Grammaticalization: Volume I. Theoretical and Methodological Issues*, edited by E. C. Traugott and B. Heine. Amsterdam: John Benjamins.
- Trudgill, Peter. 1972. "Sex, Covert Prestige, and Linguistic Change in Urban British English." *Language in Society* 1(2):179–95.
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, Peter. 1983. *On Dialect: Social and Geographical Perspectives*. Oxford: Wiley-Blackwell.
- Trudgill, Peter. 1986. *Dialects in Contact*. Oxford: Wiley-Blackwell.
- Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Oxford: Oxford University Press.
- Trudgill, Peter. 2008. "Colonial Dialect Contact in the History of European Languages: On the Irrelevance of Identity to New-Dialect Formation." *Language in Society* 37(2):241–80.
- Tuten, Donald N. 2008. "Identity Formation and Accommodation: Sequential and Simultaneous Relations." *Language in Society* 37(2):259–62.

UN-Habitat, and UNICEF Lebanon. 2019. *Old Saida Neighbourhood Profile 2019*. Beirut: UN-Habitat Lebanon.

United Kingdom: Home Office. 2018. *Country Policy and Information Note - Lebanon: Palestinians*. London: United Kingdom: Home Office.

UNRWA. n.d. “Where We Work.” *UNRWA*. Retrieved January 9, 2022 (<https://www.unrwa.org/where-we-work/lebanon>).

Versteegh, Kees. 2001. *The Arabic Language*. Edinburgh: Edinburgh University Press.

Versteegh, Kees. 2009. “Serial Verbs.” Pp. 195–99 in *Encyclopedia of Arabic Language and Linguistics*. Vol. IV, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

Vicente, Ángeles. 2019. “Dialect Contact and Urban Dialects.” Pp. 106–16 in *The Routledge Handbook of Arabic Sociolinguistics*, edited by E. Al-Wer and U. Horesh. London and New York: Routledge.

Vicente, Ángeles. 2020. “Andalusi Arabic.” Pp. 225–44 in *Arabic and Contact-Induced Change*, edited by C. Lucas and S. Manfredi. Berlin: Language Science Press.

Villena-Ponsoda, Juan Andrés. 2005. “How Similar are People who Speak Alike? An Interpretive Way of Using Social Networks in Social Dialectology Research.” Pp. 303–34 in *Dialect Change: Convergence and Divergence in European Languages*, edited by P. Auer, F. Hinskens, and P. Kerswill. Cambridge: Cambridge University Press.

- Vollers, Karl. 1890. *Lehrbuch der Aegypto-Arabischen Umgangssprache, mit Übungen und einem Glossar*. Karlsruhe, Germany: Selbstverlag des Verf.
- Watson, Janet C. E. 1993. *A Syntax of San 'ani Arabic*. Wiesbaden: Otto Harrassowitz.
- Watson, Janet C. E. 1999. "The Directionality of Emphasis Spread in Arabic." *Linguistic Inquiry* 30(2):289–300.
- Watson, Janet C. E. 2005. "Adverbs." Pp. 21–25 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. "Empirical Foundations for a Theory of Language Change." Pp. 95–195 in *Directions for Historical Linguistics: A Symposium*, edited by W. P. Lehmann and Y. Malkiel. Austin, Texas: University of Texas Press.
- Wilmsen, David. 2006. "Egypt." Pp. 1–12 in *Encyclopedia of Arabic Language and Linguistics*. Vol. II, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Wilmsen, David. 2013. "The Interrogative Origin of the Arabic Negator /-š/: Evidence from Copular Interrogation in Andalusī Arabic, Maltese, and Modern Spoken Egyptian and Moroccan Arabic." *Zeitschrift für Arabische Linguistik* 58:5–31.
- Wilmsen, David. 2014. *Arabic Indefinites, Interrogatives and Negators: A Linguistic History of Western Dialects*. Oxford: Oxford University Press.

- Wilmsen, David. 2020. "Croft's Cycle in Arabic: The Negative Existential Cycle in a Single Language." *Linguistics* 58(2):493–535.
- Wilmsen, David. 2021. "Maltese: A Peripheral Dialect in the Historical Dialectology of Arabic." Pp. 244–74 in *The Cambridge Handbook of Arabic Linguistics*, edited by K. C. Ryding and D. Wilmsen. Cambridge: Cambridge University Press.
- Winford, Donald. 1996. "The Problem of Syntactic Variation." Pp. 177–92 in *Sociolinguistic Variation: Data, Theory, and Analysis. Selected Papers from NWAV 23, Stanford University*, edited by J. Arnold, R. Blake, B. Davidson, S. Schwenter, and J. Solomon. Stanford, California: Center for the Study of Language and Information.
- Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Oxford: Wiley-Blackwell.
- Woidich, Manfred. 1968. "Negation und Negative Sätze im Ägyptisch-Arabischen." PhD dissertation, Ludwig Maximilian University of Munich, Munich, Germany.
- Woidich, Manfred. 2002. "Verbalphrasen mit Asyndetischem Perfekt im Ägyptisch-Arabischen." *Estudios de Dialectología Norteafricana y Andalusí* 6:121–92.
- Woidich, Manfred. 2005a. "B'ēri Arabic." Pp. 299–308 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Woidich, Manfred. 2005b. "Cairo Arabic." Pp. 323–33 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.

Wolfram, Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, DC: Center for Applied Linguistics.

Wolfram, Walt, and Ralph W. Fasold. 1974. *The Study of Social Dialects in American English*. Englewood Cliffs, NJ: Prentice-Hall.

Wolfram, Walt, and Natalie Schilling-Estes. 2016. *American English: Dialects and Variation*. 3rd ed. Oxford: Wiley-Blackwell.

Wolfson, Nessa. 1976. "Speech Events and Natural Speech: Some Implications for Sociolinguistic Methodology." *Language in Society* 5(2):189–209.

Younes, Munther. 1982. "Problems in the Segmental Phonology of Palestinian Arabic." PhD dissertation, The University of Texas at Austin, Austin, TX.

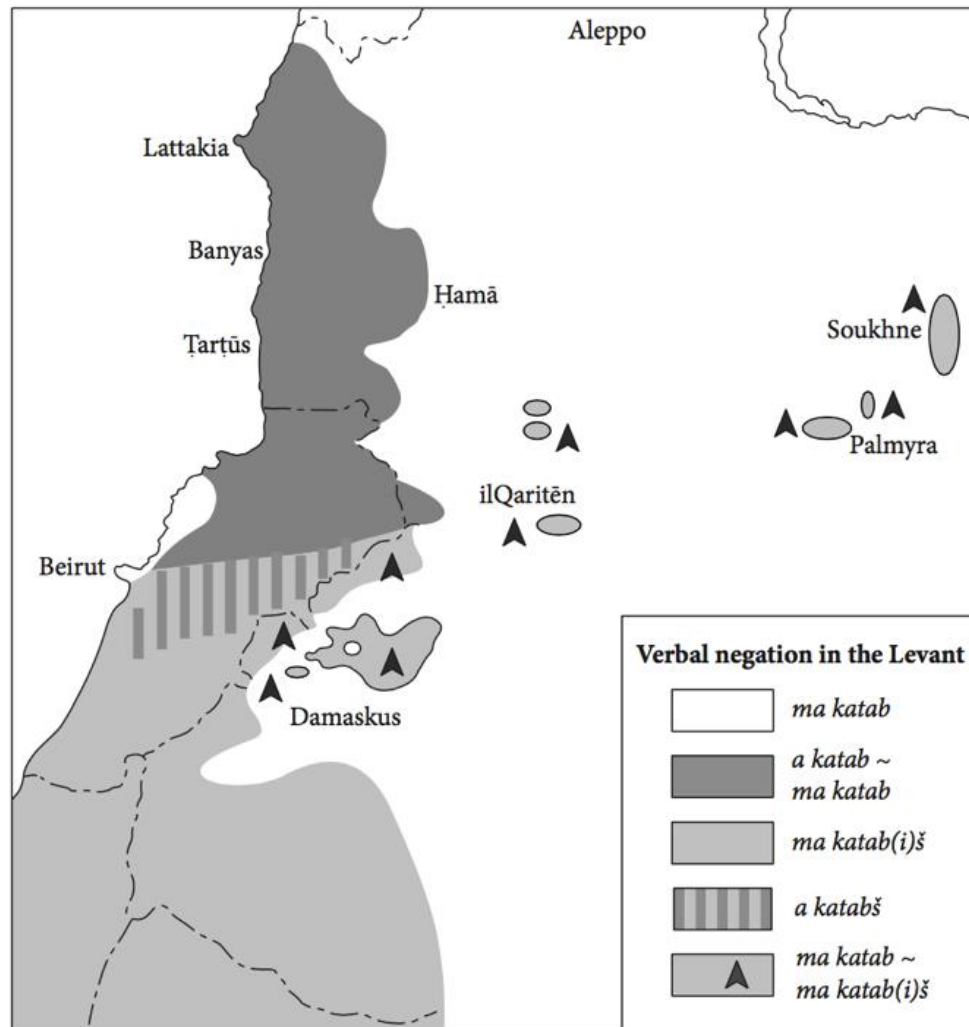
Young, Richard, and Robert Bayley. 1996. "Varbrul Analysis for Second Language Acquisition Research." Pp. 253–306 in *Second Language Acquisition and Linguistic Variation*, edited by R. Bayley and D. R. Preston. Amsterdam: John Benjamins.

Zaki, Mai, David Wilmsen, and Dana Abdulrahim. 2021. "The Utility of Arabic Corpus Linguistics." Pp. 473–503 in *The Cambridge Handbook of Arabic Linguistics*, edited by K. C. Ryding and D. Wilmsen. Cambridge: Cambridge University Press.

Zawaydeh, Bushra. 1997. "An Acoustic Analysis of Uvularization Spread in Ammani-Jordanian Arabic." *Studies in the Linguistic Sciences* 27(1):185–200.

- Zawaydeh, Bushra. 1998. "Gradient Uvularization Spread in Ammani-Jordanian Arabic." Pp. 117–42 in *Perspectives on Arabic Linguistics XI: Papers from the Annual Symposium on Arabic Linguistics*, edited by E. Benmamoun, M. Eid, and N. Haeri. Amsterdam: John Benjamins.
- Zemánek, Petr. 2005. "Assimilation." Pp. 204–6 in *Encyclopedia of Arabic Language and Linguistics*. Vol. I, edited by K. Versteegh, M. Eid, A. Elgibali, M. Woidich, and A. Zaborski. Leiden: Brill.
- Zsiga, Elizabeth C. 2013. *The Sounds of Language: An Introduction to Phonetics and Phonology*. Oxford: Wiley-Blackwell.

Appendix A: Distribution of negative constructions in northern Levantine dialects



Source: Behnstedt and Woidich (2005:101, via Wilmsen 2014: 153)

Appendix B: Distributional analysis of the contribution of linguistic factors to the selection of word-medial raised [e:]

Distributional analysis of the contribution of the preceding and following phonological contexts to the selection of word-medial raised [e:] in the speech of younger Lebanese and younger (third-generation) Palestinians

Younger Lebanese			Younger Palestinians		
	%	Total N		%	Total N
	59	522/885		43	282/654
	%	N		%	N
Preceding phonological context			Preceding phonological context		
Palatal	95	82/86	Labial	68	66/97
Dental-alveolar	88	219/248	Dental-alveolar	67	108/162
Velar	80	109/137	Palatal	62.5	50/80
Labial	78	95/122	Velar	42	37/88
Guttural	8	12/148	Guttural	13	18/142
/r/	7	4/61	/r/	6.5	2/31
Emphatic	2	1/48	Uvular	5	1/19
Uvular	0	0/35	Emphatic	0	0/35
Following phonological context			Following phonological context		
Uvular	93	14/15	Uvular	90.5	19/21
Dental-alveolar	71	324/456	Velar	52	15/29
Velar	67	8/12	Dental-alveolar	48	161/334
Labial	58	63/109	Palatal	43	16/37
Guttural	52	66/128	Labial	43	30/70
Palatal	47	34/73	Guttural	37	33/90
/r/	22	13/58	/r/	15	8/54
Emphatic	0	0/34	Emphatic	0	0/19

Distributional analysis of the contribution of the preceding and following phonological contexts to the selection of word-medial raised [e:] in the LPTC

Lebanese Popular Theatre Corpus (LPTC)		
	%	Total N
	64	2390/3741
	%	N
Preceding phonological segment		
Palatal	83	400/483
Dental-Alveolar	80	835/1041
Labial	77	427/554
Velar	78	300/384
Guttural	51	328/643
Uvular	49	68/138
Secondary pharyngeal [r]	12	31/264
Emphatic	0.5	1/234
Following phonological segment		
Uvular	94	66/70
Velar	86.5	147/170
Dental-Alveolar	73.5	1201/1634
Labial	68	309/457
Guttural	65	328/503
Palatal	59	263/446
Secondary pharyngeal [r]	25	76/299
Emphatic	0	0/162

Appendix C: Overall distribution of the variants of verbal negation according to linguistic factors across different age groups in LA and PA

Overall distribution of variants according to linguistic factors across different age groups in the LA synchronic dataset

LA Synchronic Dataset																
Older								Middle-aged				Younger				
ma			ma-š		-š		miš		ma		miš		ma		miš	
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N
Total N	67	88/131	22	29/131	6	8/131	4	5/131	94	311/332	6	21/332	95	883/931	5	48/931
Verb																
Imperfect (Present)	75	30/40	20	8/40	2.5	1/40	2.5	1/40	98	172/175	2	3/175	98	511/522	2	11/522
Perfect (Past)	32	12/37	54	20/37	14	5/37	0	0/37	100	54/54	0	0/54	96	175/182	4	7/182
Future or Participle	60	6/10	0	0/10	0	0/10	40	4/10	56	22/39	44	17/39	58	42/72	42	30/72
Pseudo-verb	91	40/44	2	1/44	5	2/44	0	0/44	98	63/64	2	1/64	100	155/155	0	0/155
Initial consonant																
Labial	85	57/67	7.5	5/67	4.5	3/67	1.5	1/67	97	198/205	3	7/205	97.5	575/590	2.5	15/590
Non-labial	48	31/64	38	24/64	8	5/64	6	4/64	89	113/127	11	14/127	90	308/341	10	33/341
Intervening material																
None	83	81/98	5	5/98	8	8/98	3	3/98	94.5	273/289	5.5	16/289	96	804/836	4	32/836
Present	21	7/33	73	24/33	0	0/33	6	2/33	88	38/43	12	5/43	83	79/95	17	16/95

Overall distribution of variants according to linguistic factors across different age groups in the PA synchronic dataset

PA Synchronic Dataset																						
Older								Middle-aged								Younger						
ma		ma-š		-š		miš		ma		ma-š		-š		miš		ma N		-š		miš		
	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N		
Total N	74	259/ 351	9	32/ 351	12	42/ 351	5	18/ 351	84	631/ 754	4.5	34/ 754	5	40/ 754	6.5	49/ 754	95	440/4 65	2	9/465	3	15/ 465
Verb																						
Imperfect (Present)	77	103/ 134	8	10/ 134	13	18/ 134	2	3/ 134	86	323/ 374	5	17/ 374	6	24/ 374	3	10/ 374	97	231/ 238	2	5/ 238	1	2/238
Perfect (Past)	78	65/83	22	18/83	0	0/83	0	0/83	87	111/ 127	9	11/ 127	0	0/ 127	4	5/ 127	100	105/ 105	0	0/ 105	0	0/105
Future or Participle	40	10/25	0	0/25	0	0/25	60	15/ 25	49	33/67	0	0/67	0	0/67	51	34/67	70	30/43	0	0/43	30	13/43
Pseudo-verb	74	81/ 109	4	4/ 109	22	24/ 109	0	0/ 109	88	164/ 186	3	6/ 186	9	16/ 186	0	0/ 186	94	74/79	5	4/79	0	0/79
Initial consonant																						
Labial	76	161/ 213	5	10/ 213	17	37/ 213	2	5/ 213	85	395/ 466	4	17/ 466	8	40/ 466	3	14/ 466	96	260/ 272	3	9/ 272	1	3/272
Non-labial	71	98/ 138	16	22/ 138	4	5/ 138	9	13/ 138	82	236/ 288	6	17/ 288	0	0/ 288	12	35/ 288	93	180/ 193	0	0/ 193	6	12/193
Intervening material																						
None	85	235/ 276	4	10/ 276	7	19/ 276	4	12/ 276	89	593/ 663	2	11/ 663	4	23/ 663	5	36/ 663	97	395/ 408	1	6/ 408	2	7/408
Present	32	24/75	29	22/75	31	23/75	8	6/75	42	38/91	25	23/91	19	17/91	14	13/91	79	45/57	5	3/57	14	8/57

Appendix D: Contribution of linguistic factors to variant selection in verbal negation in the POHA

Contribution of linguistic factors to variant selection in verbal negation in the POHA

(Notes: shaded areas indicate statistically significant factor groups; K.O.= invariant context)

Palestinian Oral History Archive (POHA)																				
ma					ma-š				-š				miš							
Verb	Input prob.		0.486		Input prob.		0.144		Input prob.		0.0338		Input prob.		0.00486					
	Total N		4014		Total N		3807 ¹⁵⁵		Total N		3807 ¹⁵⁶		Total N		4014					
	Deviance		3618.616		Deviance		2833.929		Deviance		2002.852		Deviance		371.693					
	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW				
	Perfect (Past)				1.265	1449	.636	.780	-.466	1449	.342	.386	-.794	1449	.018	.311	-1.959	1449	.003	.124
	Imperfect (Present)				.503	1332	.649	.623	-.140	1332	.154	.465	.547	1332	.193	.633	-1.924	1332	.005	.127
	Pseudo-verb				.286	1026	.592	.571	.606	1026	.189	.647	.247	1026	.216	.561	-2.121	1026	.003	.107
	Future or participle				-2.054	207	.280	.114	K.O.	207	.000	K.O.	K.O.	207	.000	K.O.	6.004	207	.720	.998
	Initial consonant of the verb																			
	Non-labial				.163	2206	.611	.541	1.043	2060	.348	.739	-1.736	2060	.018	.150	-.060	2206	.048	.485
Labial				-.163	1808	.611	.459	-1.043	1747	.102	.261	1.736	1747	.268	.850	.060	1808	.032	.515	
Intervening material																				
None				1.442	2963	.732	.809	-1.091	2767	.112	.251	-.912	2767	.119	.287	.309	2963	.052	.577	
Present				-1.442	1051	.267	.191	1.091	1040	.563	.749	.912	1040	.169	.713	-.309	1051	.008	.423	

¹⁵⁵ There were 207 tokens (4014 – 207 = 3807) including future tense verbs/participles that were excluded from the analysis because of a K.O. value.

¹⁵⁶ There were 207 tokens (4014 – 207 = 3807) including future tense verbs/participles that were excluded from the analysis because of a K.O. value.

Appendix E: Contribution of social factors to variant selection in future temporal reference in LA and PA

Contribution of social factors to variant selection in the LA synchronic dataset

LA synchronic dataset																
rah					ha				bidd-				b-			
Input prob. 0.112					Input prob. 0.142				Input prob. 0.219				Input prob. 0.356			
Total N 447					Total N 447				Total N 447				Total N 447			
Deviance 369.015					Deviance 354.342				Deviance 456.040				Deviance 535.461			
L.O. N Prop. App. Value FW					L.O. N Prop. App. Value FW				L.O. N Prop. App. Value FW				L.O. N Prop. App. Value FW			
Sex																
Male					.481 88 .250 .618				-.216 88 .193 .446				.009 88 .216 .502			
Female					-.481 359 .164 .382				.216 359 .214 .554				-.009 359 .220 .498			
Age/Education																
Middle-aged / post-secondary					-.314 188 .186 .422				-.040 188 .181 .490				.209 188 .266 .552			
Young / post-secondary					.314 259 .178 .578				.040 259 .232 .510				-.209 259 .185 .448			

Contribution of social factors to variant selection in the PA synchronic dataset
(Note: shaded areas indicate statistically significant factor groups)

PA synchronic dataset																				
rah					ḥa				bidd-				b-							
Sex	Input prob.		0.064		Input prob.		0.131		Input prob.		0.059		Input prob.		0.388					
	Total N		571		Total N		571		Total N		571		Total N		571					
	Deviance		323.677		Deviance		463.702		Deviance		571.612		Deviance		636.796					
	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW				
	Male	-.286	177	.073	.429	.163	177	.249	.541	-.293	177	.186	.427	.206	177	.492	.551			
	Female	.286	394	.124	.571	-.163	394	.264	.459	.293	394	.264	.573	-.206	394	.348	.449			
Age/Education																				
Old/ Primary or secondary		.763	85	.212	.682	-1.505		85	.059	.182	.237		85	.294	.559	.196		85	.435	.549
Middle-aged / primary or secondary		.000	203	.094	.500	.093		203	.217	.523	-.159		203	.217	.460	.330		203	.473	.582
Middle-aged / post-secondary		.068	142	.106	.517	-.379		142	.169	.406	.276		142	.289	.569	.329		142	.437	.582
Young / post-secondary		-.831	141	.071	.303	1.791		141	.532	.857	-.354		141	.191	.412	-.855		141	.206	.298
Residence																				
Outside camps		.200	455	.116	.550	.343		455	.290	.585	-.109		455	.235	.473	-.091		455	.358	.477
Inside camps		-.200	116	.078	.450	-.343		116	.138	.415	.109		116	.259	.527	.091		116	.526	.523

Appendix F: Contribution of linguistic factors to variant selection in future temporal reference across age groups in LA and PA

Contribution of linguistic factors to variant selection in middle-aged LA speech

(Note: shaded areas indicate statistically significant factor groups)

Middle-aged Lebanese speakers																			
rah					ħa				bidd-				b-						
Input prob.		0.006			Input prob.		0.260		Input prob.		0.291		Input prob.		0.422				
Total N		188			Total N		188		Total N		188		Total N		188				
Dev.		118.392			Dev.		130.463		Dev.		211.917		Dev.		218.227				
L.O.		N	Prop.	FW	L.O.		N	Prop.	FW	L.O.		N	Prop.	FW	L.O.		N	Prop.	FW
			App.	Value				App.	Value				App.	Value				App.	Value
Polarity																			
Negative	-1.238	23	.044	.225	1.295	23	.609	.785	-8.692	23	.000	<.001	.149	23	.348	.537			
Affirmative	1.238	165	.206	.775	-1.295	165	.121	.215	8.692	165	.303	>.999	-.149	165	.370	.463			
Clause type																			
Main	1.239	95	.305	.775	.021	95	.179	.505	.022	95	.253	.505	-.648	95	.263	.343			
Subordinate	-1.239	93	.065	.225	-.021	93	.183	.495	-.022	93	.280	.495	.648	93	.473	.657			
Temporal distance																			
Proximate	.664	47	.298	.660	.276	47	.128	.569	.436	47	.298	.607	-.399	47	.277	.402			
Distal	-.020	19	.053	.495	.136	19	.053	.534	-.267	19	.263	.434	.227	19	.632	.557			
Indeterminate	-.644	122	.164	.344	-.412	122	.221	.398	-.169	122	.254	.458	.172	122	.361	.543			
Temporal adverbials																			
Adverbial	-1.402	72	.153	.197	-1.207	72	.028	.230	.274	72	.333	.568	.747	72	.486	.679			
No adverbial	1.402	116	.207	.803	1.207	116	.276	.770	-.274	116	.224	.432	-.747	116	.293	.321			
Grammatical person																			
1 st -person	.843	91	.308	.699	.255	91	.187	.563	-.354	91	.209	.412	-.185	91	.297	.454			
Non-1 st	-.843	97	.072	.301	-.255	97	.175	.437	.354	97	.320	.588	.185	97	.433	.546			

Contribution of linguistic factors to variant selection in young LA speech
(Note: shaded areas indicate statistically significant factor groups)

Younger Lebanese speakers																
rah					ħa				bidd-				b-			
Input prob.		0.140			Input prob.		0.139		Input prob.		0.158		Input prob.		0.252	
Total N		259			Total N		259		Total N		259		Total N		259	
Dev.		181.241			Dev.		215.042		Dev.		236.419		Dev.		310.339	
Prop.					Prop.				Prop.				Prop.			
L.O.	N	App.	FW		L.O.	N	App.	FW	L.O.	N	App.	FW	L.O.	N	App.	FW
		Value					Value				Value				Value	
Polarity																
Negative	.053	40	.225	.513	.800	40	.500	.690	-9.001	40	.000	<.001	-.229	40	.275	.443
Affirmative	-.053	219	.169	.487	-.800	219	.183	.310	9.001	219	.219	>.999	.229	219	.429	.557
Clause type																
Main	.165	123	.203	.541	.464	123	.325	.614	-.161	123	.154	.460	-.318	123	.317	.421
Subordinate	-.165	136	.154	.459	-.464	136	.147	.386	.161	136	.213	.540	.318	136	.485	.579
Temporal distance																
Proximate	1.012	11	.273	.733	.904	11	.455	.712	.085	11	.182	.521	-1.638	11	.091	.163
Distal	-1.173	40	.125	.236	-.666	40	.075	.339	-.099	40	.200	.475	1.217	40	.600	.772
Indeterminate	.161	208	.183	.540	-.238	208	.250	.441	.014	208	.183	.504	.421	208	.385	.604
Temporal adverbials																
Adverbial	-.473	68	.118	.384	-.997	68	.059	.270	.292	68	.250	.572	.514	68	.574	.626
No adverbial	.473	191	.199	.616	.997	191	.293	.730	-.292	191	.162	.428	-.514	191	.346	.374
Grammatical person																
1 st -person	.083	101	.188	.521	-.406	101	.158	.400	.351	101	.248	.587	-.078	101	.406	.480
Non-1 st	-.083	158	.171	.479	.406	158	.278	.600	-.351	158	.146	.413	.078	158	.405	.520

Overall distribution of variants according to linguistic factors in older PA speech

Older Palestinian speakers								
	raḥ		ḥa		bidd-		b-	
	%	N	%	N	%	N	%	N
Total N	21	18/85	6	5/85	29	25/85	44	37/85
Polarity								
Negative	64	9/14	14	2/14	0	0/14	22	3/14
Affirmative	13	9/71	4	3/71	35	25/71	48	34/71
Clause Type								
Main	29	12/42	9	4/42	33	14/42	29	12/42
Subordinate	14	6/43	2	1/43	26	11/43	58	25/43
Temporal Distance								
Proximate	8	1/12	0	0/12	50	6/12	42	5/12
Distal	8	1/12	0	0/12	50	6/12	42	5/12
Indeterminate	26	16/61	8	5/61	22	13/61	44	27/61
Temporal Adverbials								
Adverbial	6	2/35	0	0/35	43	15/35	51	18/35
No adverbial	32	16/50	10	5/50	20	10/50	38	19/50
Grammatical Person								
First-person	19	6/32	3	1/32	37	12/32	41	13/32
Non-first-person	23	12/53	7.5	4/53	24.5	13/53	45	24/53

Contribution of linguistic factors to variant selection in middle-aged PA speech
(Note: shaded areas indicate statistically significant factor groups)

Middle-aged Palestinian speakers																	
rah					ħa				bidd-				b-				
Polarity Negative Affirmative Clause Type Main Subordinate Temporal Distance Proximate Distal Indeterminate Temporal Adverbials Adverbial No adverbial Grammatical Person 1 st -person Non-1 st	Input prob.		0.046		Input prob.		0.107		Input prob.		0.065		Input prob.		0.466		
	Total N		345		Total N		345		Total N		345		Total N		345		
	Dev.		170.102		Dev.		258.658		Dev.		348.742		Dev.		402.728		
	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	
	.992	52	.327	.729	.610	52	.481	.648	-1.507	52	.019	.181	-.550	52	.173	.366	
	-.992	293	.058	.271	-.610	293	.147	.352	1.507	293	.287	.819	.550	293	.509	.634	
	.149	217	.115	.537	.625	217	.249	.651	.080	217	.253	.520	-.519	217	.382	.373	
	-.149	128	.070	.463	-.625	128	.109	.349	-.080	128	.234	.480	.519	128	.586	.627	
	-.161	61	.033	.460	.608	61	.148	.647	.192	61	.262	.548	-.272	61	.557	.432	
	-.063	45	.022	.484	-.621	45	.044	.350	-.362	45	.244	.410	.453	45	.689	.611	
.224	239	.130	.556	.013	239	.238	.503	.170	239	.243	.542	-.181	239	.389	.455		
-1.164	140	.014	.238	-.959	140	.043	.277	.143	140	.300	.536	.536	140	.643	.631		
1.164	205	.156	.762	.959	205	.302	.723	-.143	205	.210	.464	-.536	205	.332	.369		
.218	145	.103	.554	-.314	145	.145	.422	-.185	145	.221	.454	.266	145	.531	.566		
-.218	200	.095	.446	.314	200	.235	.578	.185	200	.265	.546	-.266	200	.405	.434		

Contribution of linguistic factors to variant selection in young PA speech
(Note: shaded areas indicate statistically significant factor groups)

Younger Palestinian speakers																			
rah					ħa					bidd-					b-				
		Input prob.		0.051			Input prob.		0.752			Input prob.		0.119			Input prob.		0.245
		Total N		141			Total N		141			Total N		141			Total N		141
		Dev.		55.875			Dev.		154.752			Dev.		131.503			Dev.		119.824
		L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW	L.O.	N	Prop. App. Value	FW		
Polarity																			
Negative		-9.304	20	.000	<.001	1.532	20	.950	.822	-8.796	20	.000	<.001	-.538	20	.050	.369		
Affirmative		9.304	121	.083	>.999	-1.532	121	.463	.178	8.796	121	.223	>.999	.538	121	.231	.631		
Clause Type																			
Main		.318	82	.085	.579	.483	82	.610	.618	-.120	82	.171	.470	-.584	82	.134	.358		
Subordinate		-.318	59	.051	.421	-.483	59	.424	.382	.120	59	.220	.530	.584	59	.305	.642		
Temporal Distance																			
Proximate		.640	6	.167	.655	-1.379	6	.333	.201	-.187	6	.167	.453	.856	6	.333	.702		
Distal		-.140	37	.135	.465	.501	37	.324	.623	-.112	37	.135	.472	.146	37	.405	.536		
Indeterminate		-.500	98	.041	.378	.878	98	.622	.706	.299	98	.214	.574	-1.002	98	.122	.269		
Temporal Adverbials																			
Adverbial		.741	32	.188	.677	-.361	32	.281	.411	-.280	32	.125	.430	.392	32	.406	.597		
No adverbial		-.741	109	.037	.323	.361	109	.606	.589	.280	109	.211	.570	-.392	109	.147	.403		
Grammatical Person																			
1 st -person		.922	69	.130	.715	-.267	69	.435	.434	.157	69	.188	.539	.027	69	.246	.507		
Non-1 st		-.922	72	.014	.285	.267	72	.625	.566	-.157	72	.194	.461	-.027	72	.167	.493		

Appendix G: Contribution of linguistic factors to variant selection in future temporal reference in the POHA

Contribution of linguistic factors to variant selection in the POHA (Note: shaded areas indicate statistically significant factor groups)

rah					ha		bidd-				b-			
Input prob.					Total N		Input prob.				Input prob.			
Total N					26/634		Total N				Total N			
Deviance					4		Deviance				Deviance			
Prop. App. Value					%		Prop. App. Value				Prop. App. Value			
L.O. N					%		L.O. N				L.O. N			
FW					N		FW				FW			
Polarity														
Negative	.991	50	.560	.729	13	7/53	-8.851	53	.000	<.001	-.283	53	.340	.430
Affirmative	-.991	482	.114	.271	3	19/581	8.851	581	.303	>.999	.283	581	.570	.570
Clause Type														
Main	.168	114	.140	.542	6	8/128	.207	128	.375	.552	-.413	128	.438	.398
Subordinate	-.168	418	.160	.458	4	18/506	-.207	506	.253	.448	.413	506	.579	.602
Temporal Distance														
Proximate	-.166	47	.021	.459	0	0/47	1.138	47	.468	.757	-1.007	47	.511	.268
Distal	-9.664	102	.000	<.001	0	0/102	-.581	102	.167	.359	.665	102	.833	.660
Indeterminate	.166	485	.169	.541	5	26/485	-.557	485	.282	.364	.342	485	.495	.585
Temporal Adverbials														
Adverbial	-1.158	234	.030	.239	1	2/335	-.350	335	.227	.413	1.050	335	.746	.741
No adverbial	1.158	298	.255	.761	8	24/299	.350	299	.334	.587	-1.050	299	.331	.259
Grammatical Person														
1 st -person	-.182	203	.143	.455	6	15/265	-.518	265	.170	.373	.412	265	.664	.602
Non-1 st	.182	329	.164	.545	3	11/369	.518	369	.355	.627	-.412	369	.469	.398

¹⁵⁷ There are 102 tokens including distal futures that were excluded from the analysis (634 – 102 = 532) because of a K.O. value (social and linguistic factors are run together in Rbrul).