The Thesis Committee for Emilie Pénélope Durand
Certifies that this is the approved version of the following thesis:

Word-Final *imaala* in Contemporary Levantine Arabic: a Case of Language Variation and Change

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor:

Kristen Brustad

Mahmoud Al-Batal

Word-Final *imaala* in Contemporary Levantine Arabic: a Case of Language Variation and Change

by

Emilie Pénélope Durand, B.A.

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2011

Abstract

Word-Final *imaala* in Contemporary Levantine Arabic: a Case of Language Variation and Change

Emilie Pénélope Durand, M.A.

The University of Texas at Austin, 2011

Supervisor: Kristen Brustad

The phenomenon of word-final *imaala*, or taa-marbuuTa raising, in the Levantine dialects of Arabic was well documented about 50 years ago by renowned Arabists who described the phenomenon as a purely phonological one. Today, after some major historical and sociological changes have taken place in Arab societies, this feature deserves to be revisited since this might shed some light on the processes of language change in those societies. The scope of this paper is to look into the issue of word-final *imaala* in contemporary Levantine Arabic (specifically after raa) through a wide lens, and to establish 1) whether there are patterns governing the production of taa marbuuTa after raa, and 2) whether the existing phonological rules account for all instances of word-final *imaala* as they appear in the speech of Levantine speakers nowadays. In order to do that, instances of all word tokens ending in -ra were extracted from 252 phone conversations recorded in 2004 and found in the LDC Levantine database. Those tokens were analyzed and the word-forms they represent were divided based on whether they exhibit any instances on word-final *imaala*. It

soon became clear that the existing sound rules cannot account for all current instances of taa-marbouTa raising. Two main factors were identified as having a possible effect on the production of taa marbuuTa after raa: word frequency and phonological word classes. Because of a lack of speaker-related information in the database coupled with some *imaala*-related discrepancies found in the transcriptions of the conversations, it was impossible to determine the exact social meaning(s) of word-final *imaala* in Levantine communities. However, this study shows that enough changes have taken place since the 1960's in terms of taa-marbuuTa raising, to consider it a case of language change in progress. This study also establishes some hypotheses which can be used as the base for a future sociolinguistic study whose scope will be to assign social meaning to word-final *imaala* in Levantine dialects.

Table of Contents

List of Tables

List of Figures

I. Introduction and Literature Review

In March 2010, while studying abroad in Syria, I was corrected by a dialect teacher who was born and raised in a traditional neighborhood of Damascus for using the word ʔaSiire. She explained rather passionately that while ʔaSiira was the right form, she considered ʔaSiire "wrong," "a made-up form," and "something a foreigner would say." A few moments later, one of her co-workers who is also a proud Damascene native, used the form ʔaSiire very naturally during a phone conversation that I happened to overhear. After this incident, I told the two what had happened. The coworker maintained that while there was nothing wrong with ʔaSiira, he felt that ʔaSiire was more neutral, while the dialect teacher stubbornly rejected ʔaSiire as a form that sounded wrong even in the mouth of a native Arabic speaker. Linguistic variation exists in the production of taa marbuuTa, and it seems to be in part sociolinguistic. Certainly, this story illustrates language change in progress.

As it is well known, the label "Arabic" constitutes an umbrella term which covers dozens of different dialects. Each variety of Arabic presents a unique set of linguistic components: phonological, morphological, syntactic, etc. In this paper, I will not be discussing the various possible classifications of dialects, but rather, for the sake of expediency, I will follow scholars who define Levantine (or Syrian) Arabic as the

group of dialects "spoken by the sedentary population of Greater Syria, that is, of present-day Syria, Lebanon, Jordan and the Arab population of Israel."[1]

One of the most distinctive features of Levantine Arabic is word-final *imaala*, a process by which the vowel corresponding to ة (taa marbuuTa) is raised from [a] to [æ], [ɛ], [e] or even [i] in some dialects. This phenomenon was studied by some of the most prominent Arabic scholars of the 20th century including Grotzfeld (1964, 1965), Cantineau (1960), Ferguson (1961), Cowell (1964), Barbot (1981) and more.

For the scope of this study, it is important to understand the conclusions these scholars came to when looking into this specific aspect of the Levantine phonology, and this is why the first part of this paper is dedicated to a literature review covering some of what has been written about word-final *imaala*. This report will be immediately followed by a discussion of the limitations of the works reviewed. The second part will introduce the present study in detail, including information about the database and the speakers. The third part will reveal what was found through careful analysis of the database with regards to word-final *imaala*, and the fourth part will introduce and discuss some problematic tokens of taa-marbuuTa raising as found in the database. The fifth part will briefly go over the most important conclusions reached through the study and offer suggestions for future research.

[1] Karl Stowasser and Moukhtar Ani, ed. *A Dictionary of Syrian Arabic* (Washington D.C.: Georgetown University Press, 1964), xiii.

In their Course in Levantine Arabic, McCarus, Qafisheh and Rammuny state that in order to make a feminine adjective from a masculine one in Levantine Arabic, most adjectives simply add the ending -e. Adjectives that end in one of the following consonants, however, take -ə (a non-raised production of taa marbuuTa) instead of -e: H, R, S, T, L, D or glottal stop (ʔ) (see Appendix for a list of all transcription codes). If the masculine adjective ends in -i, the feminine is formed by adding -yye.[2] In this overview of feminine adjective agreement, nothing is said of back consonants such as h, ʕ, kh or gh which are known to prevent taa-marbuuTa raising. It should also be noted that in their work, McCarus, Qafisheh and Rammuny treat emphatic and non-emphatic r as two separate phonemes.[3] Thackston, in his Vernacular Arabic of the Lebanon, explains that the feminine singular of all adjectives is formed by adding -e or -a to the masculine singular. If the masculine singular adjective ends in one of the back or velar consonants (h, H, ʕ, kh, gh, ʔ, q, T, D, S, Z), the ending is -a. In all other cases, the ending is -e. If an adjective ends in -iir, the feminine is -iire; otherwise -r is considered a velar consonant, and the feminine is marked with -ra.[4] Ferguson reports that in Damascus Arabic, what determines whether taa marbuuTa is raised or not is

[2] Ernest McCarus and Hamdi Qafisheh and Raji Rammuny, *A Course in Levantine Arabic* (Ann Harbor: the University of Michigan, 1978), 35.

[3] McCarus and Qafisheh and Rammuny, *Course in Levantine Arabic*, 16.

[4] W.M. Thackston, Jr, *The Vernacular Arabic of the Lebanon* (Massachusetts: Harvard University, 2003), 8.

the quality of the immediately preceding consonant. When it is a throat consonant (kh, gh, H, ʕ, h, ʔ) or a velarized one (T, D, S, DH), the ending is -a. After r, the ending is usually -a but sometimes (especially after -ii-) it is -e. After any other consonant, it is -e (with some exceptions in foreign words and irregular patterns).[5] Ferguson also argues that while emphatic and non-emphatic r can be found in contrastive distribution in a large part of the Syrian Arabic area (jaari 'running, flowing' vs. jaaRi 'my neighbor'), this is not the case in Damascus where r is in between the plain and velarized consonants in its effect on the sound a.[6] Cowell, in his famous Reference Grammar of Syrian Arabic, notes that most nouns come in pairs: one with the ending –e and one without it. Taa marbuuTa is normally produced as -a after velarized consonants (T, S, D, DH) and back consonants (kh, gh, q, H, ʕ, h, ʔ) and usually after r (but not usually after -iir-). Otherwise, it is pronounced -e. Some other exceptions are also found where -e occurs after r in Syrian Arabic (especially in the pattern fəʔla/e <fiʔla): إبرة, نمرة and sometimes after a velarized consonant: ʕaTSe 'sneeze' (but more usually ʕaTSa).[7]

Cantineau very carefully explains the *imaala* of the feminine ending as follows: it is an -a followed by an inaudible -h sound. Since -h is not pronounced, -a undergoes

[5] Charles A. Ferguson and Moukhtar Ani, *Damascus Arabic* (Washington D.C.: Center for Applied Linguistics, 1961), 115-116.

[6] Ferguson and Ani, *Damascus Arabic*, 174.

[7] Mark Cowell, *A Reference Grammar of Syrian Arabic* (Washington D.C.: Georgetown University Press, 1964), 138.

compensatory lengthening and becomes prone to *imaala* just like any long -aa can be. The feminine ending therefore becomes -e in dialects of Syria, Palestine and Iraq every time it is not preceded by an emphatic consonant.[8] He describes r in the frame of Classical Arabic as a phoneme which is never back, velar or uvular, but which can be velarized.[9] This velarization of r has to be triggered by neighboring consonants and has no contrastive function as it has been claimed by other scholars as seen above.[10] According to Barbot, there are two possible productions of r in Damascene Arabic: one emphatic and another non-emphatic, depending on the environment. r is emphatic when it is found in the vicinity of another emphatic sound, or in the contact of a/a:, u:/o: and u/ə (<u) (except if i: or i impedes it). r is non-emphatic when it is in contact with i: (or e<i, or e:<ay, or y), however if an emphatic consonant is found in the word, then this phenomenon is blocked and r is produced as an emphatic.[11] Word finally, if the consonant preceding taa marbuuTa is back or emphatic (which would include emphatic productions of r), the *imaala* cannot take place.[12] Finally, in his Syrian Arabic Grammar, Grotzfeld states that the production of taa marbuuTa as -a or -e depends on the preceding consonant. It is pronounced -a after: ʔ, H, kh, R, S, D, T, Z, ʕ, gh, h. It is

[8] Jean Cantineau, *Etudes de Linguistique Arabe* (Paris: Librairie C. Klincksieck, 1960), 100.

[9] Cantineau, *Etudes de Linguistique Arabe*, 172.

[10] Cantineau, *Etudes de Linguistique Arabe*, 182.

[11] M. Barbot, *Evolucion de l'Arabe Contemporain* (Paris: La Sorbonne, 1981), 396.

[12] Barbot, *Evolucion Arabe Contemporain*, 725.

pronounced -e after: b t r d r z ʃ f k l m n w y.[13] Grotzfeld also explains that r is produced as an emphatic in Damascene Arabic in the following phonological environments: 1) near an emphatic consonant (unless i or another non-velar consonant impedes the tafkhiim of r), 2) in contact with u or ə<u, 3) sometimes in contact with o:, 4) in contact with a or a: as long as i or one of the non-velarized consonants s, z, t, d is not there, otherwise the tafkhiim of r cannot happen.[14]

To summarize, Arabists and Semitists have described taa-marbuuTa raising with more or less rigor, but the more in-depth studies show that this phenomenon is phonologically conditioned in that it is determined by the quality of the immediately preceding consonant. Overall, scholars all seem to agree that *imaala* does not take place after emphatic and post-velar consonants, but does occur in all other phonological environments. This explains the contrast found in the production of the feminine ending in words such as baTTa 'female duck' (T is emphatic) and waaHa 'oasis' (H is post-velar) as opposed to salaame 'safety' and baSle 'an onion,' for example. One phoneme that seems to cause some degree of disagreement between scholars is /r/. There exist two possible productions of this voiced dental trill: emphatic and non-emphatic, and this affects the production of any following taa marbuuTa since

[13] Heinz Grotzfeld, *Syrisch-Arabische Grammatik* (Wiesbaden: Otto Harrassowitz, 1965), 45.

[14] Heinz Grotzfeld, *Laut- und Formenlehre des Damaszenisch-Arabischen* (Wiesbaden: Kommissionsverlag Franz Steiner GMBH, 1964), 18.

vowel-raising takes place after non-emphatic r but not after emphatic r.[15] This contributes to giving non-linguist native speakers, and foreigners trying to master a Levantine dialect, the false impression that taa marbuuTa-raising after r is a random process that affects some words and not others for no particular reason. Scholars themselves have taken slightly different stances on this issue: as mentioned above, McCarus, Qafisheh and Rammuny treat emphatic and non-emphatic r as contrastive phonemes,[16] Thackston's explanations is incomplete,[17] and Cowell's description remains somewhat vague, stating that taa marbuuTa *usually* takes the form -a after r though *usually* not after -iir-.[18] As for Cantineau, Grotzfeld and Barbot, they do not view emphatic and non-emphatic r as two distinct phonemes, but rather as allophones of the same non-emphatic phoneme.[19] Some of the claims made about *imaala* are very clean-cut, such as Cantineau's remarks according to which the feminine ending becomes -e in dialects of Syria, Palestine and Iraq *every time* it is not preceded by an emphatic consonant.[20] Table 1 below is a summary of each scholar's findings.

[15] Grotzfeld, *Laut- und Formenlehre*, 90.

[16] McCarus and Qafisheh and Rammuny, *Course in Levantine Arabic*, 16.

[17] Thackston, *Vernacular Arabic of Lebanon*, 8.

[18] Cowell, *Reference Grammar*, 138.

[19] Cantineau, *Etudes de Linguistique Arabe*, 182; Wolfdietrich Fischer and Otto Jastrow, *Handbuch der Arabischen Dialekte* (Wiesbaden: Otto Harrassowitz, 1980), 180; Barbot, *Evolucion Arabe Contemporain*, 396.

[20] Cantineau, *Etudes de Linguistique Arabe*, 100.

Scholar's Name	No *imaala* after:	*Imaala* takes place after:	Treatment of raa
McCarus, Qafisheh and Rammuny	H, R, S, T, L, D, ʔ	Everywhere else	No separate treatment of raa. They consider that emphatic raa (R) and non-emphatic raa (r) are two distinct phonemes. *Imaala* after r, but not after R.
Thackston	h, H, ʕ, kh, gh, ʔ, q, T, D, S, Z	Everywhere else	*Imaala* after -iir. In every other environment, no *imaala* (-ra).
Ferguson	kh, gh, H, ʕ, h, ʔ, T, D, S, DH	Everywhere else (with some exceptions in foreign words and irregular patterns)	• Usually -ra but sometimes (especially after ii), it is -re • one phoneme (not two) with different allophones
Cowell	T, S, D, DH, kh, gh, q, H, ʕ, h, ʔ	Everywhere else	Usually –ra but sometimes –re (especially after iir and in fiʕla pattern)
Cantineau	After emphatic consonants	Everywhere else	• Never back, velar or uvular, but can be velarized • Velarization is triggered by neighboring consonants and has no contrastive function
Barbot	If the consonant preceding taa marbuuTa is back or emphatic (which would include emphatic productions of raa)	If the consonant preceding taa marbuuTa is not back nor emphatic	• Raa is emphatic when: it is found in the vicinity of another emphatic sound, OR in the contact of a/a:, u:/o: and u/ə (<u) (except if i: or i impedes it) • Raa is not emphatic when it is in contact with i: (or e<i, or e:<ay, or y). However, if an emphatic consonant is found in the word, then this phenomenon is blocked and raa is produced as an emphatic
Grotzfeld	ʔ, H, kh, R, S, D, T, Z, ʕ, gh, h	b t r d r z s ʃ f k l m n w y	• Raa is emphatic (R): 1) near an emphatic consonant (unless i or another non-velar consonant impedes the tafkhiim of raa), 2) in contact with u or ə<u, 3) sometimes in contact with o:, 4) in contact with a or a: as long as i or one of the non-velarized consonants s, z, t, d is not there • Otherwise, raa is non-emphatic (r)

Table 1-Summary of Literature about Word-Final *Imaala*

Although the work of those renowned scholars, especially Cantineau, Grotzfeld and Barbot, constitutes a gigantic contribution to the field of Arabic, it is by no means absolute or irrevocable. The first limitation of such work is the fact that it is partial in nature. From a geographical perspective, each scholar had to focus on a specific part of the Levant, thus targeting a more defined variety of Arabic within the Levantine dialects. This presents an obvious advantage in that it limits variation and simplifies data organization and analysis, but it can also be seen as a disadvantage since it reduces the scope of data and impedes the erection of a more complete picture of features such as word-final *imaala*. As far as I am aware, no detailed broad-range comparative study of word-final *imaala* is available in English, French or German that offers descriptive rules of the phenomenon while taking into account a truly representative range of locations within the Levant. To mention only the scholars whose work was reviewed above, it may be worth pointing out that Cantineau's conclusions on word-final *imaala* were based on observations made in Syria, Palestine and Iraq (however, the exact communities remain unnamed), Cowell's work was based on city-dwelling Syrians (in particular natives of Damascus), Ferguson's and Barbot's on Damascus as well, Grotzfeld's on Damascus but also unnamed communities of Syria, Lebanon, Palestine and Transjordan, McCarus, Qafisheh and Rammuny's on the dialect spoken in Jerusalem, and finally Thackston's on the Lebanese mountain area and Beirut (leaving out the coast entirely). Thackston explains that in order to write his book, he had to leave out what he calls "unusual features" in order to be able to

generalize on the way Lebanese Arabic behaves.[21] It is not impossible that one of those features may have had to do with instances of taa-marbuuTa raising. It is also very likely that other scholars whose work covered more than one linguistic community had to do what Thackston did and select which aspects of the language they would account for, based on majority rule across the varieties under study. Therefore it seems like while the dialect of Damascus is well documented, one can hardly say that word-final *imaala* studies have been representative of the entire Levant area.

The second limitation that the available research presents, rests in the fact that the majority of these studies are no longer up-to-date. The book by Cantineau used as a reference in this study dates back to 1960, Cowell's book to 1964, Ferguson's to 1961, Grotzfeld's 1964 and 1965, McCarus, Qafisheh and Rammuny's to 1974, and Barbot's to 1981. The only exception is Thackston's 2003 textbook on the Vernacular Arabic of the Lebanon. Considering recent Middle Eastern history and the extent and impact of population shifts due to more or less constant waves of migration, immigration and urbanization in the Arab world over the past hundred years or so, it is only natural to expect the kinds of Arabic spoken today in Damascus, Baghdad, Amman or Jerusalem to differ rather substantially from the varieties that were spoken in those same urban centers in the 1960's, 1970's and even 1980's. The ʔaSiira/ʔaSiire story mentioned at the very beginning of this paper illustrates this point well. As mentioned

[21] Thackston, *Vernacular Arabic of Lebanon*, viii.

above, Grotzfeld's study of Damascene Arabic reveals that r is produced as an emphatic near an emphatic consonant unless i or another non-velar consonant impedes the tafkhiim of r.[22] Cantineau also states that -a is pronounced -e every time it is preceded by a non-emphatic sound.[23] Based on these rules, which constitute the general consensus, one would predict that the feminine of the adjective ʔaSiir 'short' is ʔaSiire. However, this form is by no means accepted by all Damascene speakers. The existing rules are not sufficient to account for such variation and changes because the studies of which they sprang out are 50 years old. For that reason alone, they need to be revisited. Doing this would be especially beneficial since there are no real longitudinal or diachronic studies in Arabic.

Another example of the limitations of these rules appears in Cowell's Reference Grammar of Syrian Arabic in which he reports that words of the fəʕla form, such as nəmre 'number' and ʔəbre 'needle' allow for vowel-raising, without making any distinction between the different words that belong to that form.[24] In my experience however, a distinction should be made since it is much rarer to hear *imaala* on khəbra 'experience' or fəkra 'idea' than on 'number' and 'needle.' Cowell's rule is inoperable

[22] Grotzfeld, *Syrisch-Arabische Grammatik*, 7.

[23] Cantineau, *Etudes de Linguistique Arabe*, 182.

[24] Cowell, *Reference Grammar*, 138.

here because variation is found and this variation is not explained by velarized/non-velarized rules (fəkra vs. ʔəbre), nor is it explained by the quality of the vowel found in the first syllable (the same vowels is found in both words). Hence the best hypothesis is that the variables are social. In his famous case study about the social value of diphthongs on Martha's Vineyard, William Labov explained:

> (…) an account of structural pressures can hardly tell the whole story. Not all changes are highly structured, and no change takes place in a social vacuum. Even the most systematic chain shift occurs with a specificity of time and place that demands an explanation.[25]

Some of the social factors which may be influencing this phenomenon are: the speaker's place of origin, gender, age, social class, religious affiliation, etc. Unfortunately, the scope of this project does not allow for an investigation of sociolinguistic variables because no available corpus takes into account enough of those variables to allow for any relevant conclusions. What is possible and what I will do here is to look at the issue of vowel-raising after r through a wide lens and to answer the following questions: in actual contemporary speech across the Levantine dialects, which words that end in raa+taa marbuuTa typically allow for vowel-raising and which do not? Of course, some variation can be expected, but how much exactly

[25] William Labov, "The Social Motivation of a Sound Change," *Word* 19 (1963): 274.

and on which words? To what extent do those instances of taa-marbuuTa raising agree with older studies? If some tokens do not fit the current rules, then what can they tell us about possible changes and new trends that may be taking place within some of the Levantine linguistic communities?

II. The Corpus

In order to answer these questions, I decided to take a quantitative look at instances of raa+taa marbuuTa as found in the only database available to me: the LDC Levantine database, one of the very few existing databases of spoken Arabic. The LDC presents 982 native Levantine Arabic speakers taking part in relatively spontaneous telephone conversations in Colloquial Levantine Arabic. A total of 985 conversations of five to six minutes each are provided. These conversations were recorded and transcribed in 2004 by Appen Pty Ltd, Sydney, Australia.

In terms of speaker demographics, Appen only provides three pieces of information about each participant: gender, country of origin and age. From the standpoint of dialectology and sociolinguistics, this type of speaker-related information is lacking at best. It is so limited that it greatly restricts the options of researchers using the LDC, and the depth and impact of their work. At the very least, the exact place of origin of each speaker should be known—where they grew up (which area, which village, which neighborhood even). Other crucial data would include: other places the speaker has lived, for how long and when, where they live now, what kind of schools they attend(ed), other social spaces they spend time in, their religious affiliation and degree of commitment to their faith, their level of education, their occupation, information about their social networks and connections, sexual orientation, and more.

Unfortunately, none of this information is available in the LDC, which points to a question that transcends this study and applies to our field as a whole: how serious are we in our efforts to propel Arabic language and linguistics into the sphere of accredited scientific scholarship? This field is in dire need of more data, more experiments, more studies, but what kind? Within which parameters? How do we intend on pushing the field forward? What is the best framework for each study? Clearly, the LDC team did not dwell too long on the potential their database could have as a research tool. Rather, they just got the work done which is much better than nothing, but a little extra effort would have led to significantly better results if only the methodological foundations of this field were properly laid out. Too often in Arabic linguistics, it is not only the answers that are lacking, but the questions themselves, and this fact is widely illustrated in the LDC Levantine database.

This lack of demographic information does not constitute the only weak point exemplified by the LDC. Another of its flaws has to do with the transcribers who helped on the project. After the conversations were recorded, a team of people listened to each interview and transcribed it in both Arabic and in transliteration. Since the database itself gives absolutely no information about the transcribers, I contacted Appen Pty Ltd directly and inquired about them. On March 30, 2011, I received the following email response from Phil Hall, Senior Vice President of Business Development for Appen Pty Ltd:

> Most of the team would have come from the regions of collection, but there may also have been transcribers from other regions (we had a lot of Iraqi Arabic running at the time, so the best performing members of that team may have contributed also). The team supervisors we had at the time were also native speakers, and came from Palestine, Jordan and Iraq.

This of course tells us nothing about the specific transcriber of individual conversations that may need more attention. In addition, why would non-native Levantine speakers (i.e. Iraqis) be involved in this most valuable process? If precision is the goal, then why not leave this task to those who are natives of this dialect and its regional variations? The strange choice of having Iraqis take up the role of transcribers coupled with the lack of information concerning the exact conversations each transcriber worked on, turned out to be very problematic in this study, as will be explained in more detail.

The counting of instances of raa+taa marbuuTa was done on the basis of the transcriptions provided in the LDC. An F-search allows users to easily locate each relevant token, and then by looking at the transliteration of each found token, one can see whether taa marbuuTa was raised or not in each particular instance. For the purpose of this study, only free tokens were considered. What is meant by free tokens is the following: Arabic, along with a few other Afro-Asiatic languages, has a special construct form (*idaafa*) which combines two lexical items (usually two nouns, but sometimes one adjective and one noun) in which the second item is in a genitive

relation to the first. Possession can be expressed with this construction, as in kitaab al-ustaaz 'the teacher's book,' but also genitive expressions other than possession, such as madiinit al-qaahira 'the city of Cairo' and Tayyib al-qalb 'good of heart.' If the first term of this construct form ends in taa marbuuTa, then *imaala* always takes place and taa marbuuTa is pronounced -it (or -et).[26] Free tokens are those words that are not the first term of the construct form and thus allow for variation on word-final *imaala*.

Because token counting is such a time-consuming task, I focused on the first 252 conversations found in the LDC, which meant the speech of 504 speakers. Demographics were only available for 483 of those speakers, and I did not analyze the speech of any speaker whose demographics was not available. Table 2 presents all the data available about those 483 speakers.

[26] Grotzfeld, *Syrisch-Arabische Grammatik*, 45.

Gender	Age	Country of Origin				
		Lebanon	Syria	Jordan	Palestine	All Countries
MALES	15-19	23	1	8	2	34
	20-29	56	40	55	2	153
	30-39	12	19	20	1	52
	40-49	6	1	7	1	15
	50-59	1	1	0	1	3
	60-69	0	0	0	0	0
	All Ages	98	62	90	7	257
FEMALES	15-19	20	6	19	0	45
	20-29	26	34	28	0	88
	30-39	20	18	24	1	63
	40-49	8	4	10	1	23
	50-59	3	1	1	1	6
	60-69	1	0	0	0	1
	All Ages	78	63	82	3	226
ALL GENDERS	15-19	43	7	27	2	79
	20-29	82	74	83	2	241
	30-39	32	37	44	2	115
	40-49	14	5	17	2	38
	50-59	4	2	1	2	9
	60-69	1	0	0	0	1
	All Ages	176	125	172	10	483

Table 2-Demographics of Speakers

By considering each of the three demographic variants (gender, country of origin and age) independently of the other two, one obtains Figures 1 to 3, which are added here as a quick visual reference meant to help readers understand basic speaker-related data at a glance.

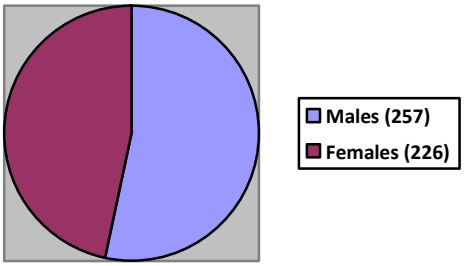

Figure 1-Gender of Speakers

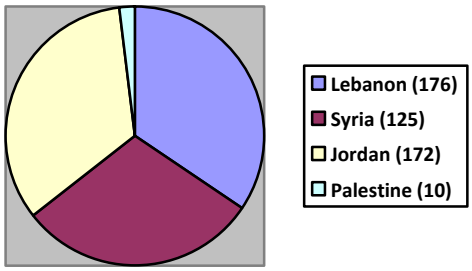

Figure 2-Speakers' Country of Origin

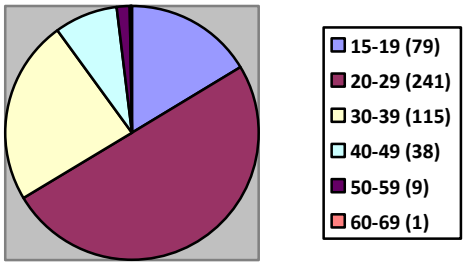

Figure 3-Speakers' Age

Males and females are represented in a fairly balanced fashion in this group of speakers (see Figure 1). However, as illustrated in Figure 2, this same balance is not reached with regards to country of origin, since Palestine is very poorly represented.

This happened by chance but undoubtedly will have some repercussions on the results.

Finally, age is also very unbalanced among speakers (see Figure 3): half of the speakers are in their twenties, followed by almost a quarter of speakers in their thirties, about 15% of speakers between fifteen and nineteen, and only about 10% of speakers between fourty and sixty (the only speaker in the '60-69' age group was sixty years old exactly). This means that it will be difficult to distinguish the role of age in word-final *imaala*; since most speech is coming from people under forty, it follows logically that most instances of taa-marbuuTa raising will also be produced by people under forty. Combining the demographics reveals that the largest groups of speakers are the following: males in their twenties from Lebanon, Syria and Jordan, then females in their twenties from the same three countries, then males in their thirties from the same countries, and finally females in their thirties from the same countries again.

In spite of all of these problems, I still chose to use the LDC corpus simply because it is the only database we have for spoken Levantine Arabic. Looking at the data found in the LDC gives us an overview of what is happening with word-final *imaala*, and allows us to frame some more specific questions which in turns may be used as the base for future studies.

III. *Imaala* in the Corpus

In this section, we will be looking at all free tokens that end in -ra in the data, whether or not they exhibit word-final *imaala*. The goal is to try and identify patterns in the production of taa marbuuTa in those words. Two main possible factors will be considered: the first is word frequency, tightly connected to the concept of lexical diffusion, and the other is the idea of natural word classes. After looking at the possible connections between word frequency and *imaala* within all instances of -ra (Table 3), and then only within those instances of –ra that exhibit word-final *imaala* (Table 4), four word classes will be identified based on the most common phonological environments exhibited by the raised tokens (Table 5). Subsequent tables in this section (6 through 12) will take both word frequency and word classes into account in an attempt to discover patterns of taa-marbuuTa raising.

The concept of lexical diffusion was expressed and promoted by sociolinguists such as Labov, who, following a long line of historical linguists whose main interest lied in describing *how* language changes over time, started investigating the possible reasons *why* language changes. Labov explains that the majority of variations in language occur only once in the mouth of one speaker and are immediately extinguished. However, in some cases, a variation can be imitated more or less widely, and spread to a degree where it creates a contrast with the older form. Later, one

form usually wins, the other form disappears and regularity is restored.[27] Phillips and other linguists have pointed out that lexical frequency can affect the diffusion of particular sound changes; some variation in modern American English can be attributed to changes affecting the most frequent words, and some to changes affecting the least frequent words. Phillips explains that:

> The connection between word frequency and lexical diffusion emphasizes the degree of attention or lexical analysis required of the speaker. That is, if the features of individual segments are blurred (as in assimilations) or ignored (as in certain stress shifts or segment reductions), the most frequent words are affected first. If speakers have to access specific information about a word, however, in order to implement a change, then the least frequent words are affected first. [28]

Although there is no evidence that these same principles can be applied to Arabic, it remains nonetheless true that word frequency is worth looking into when dealing with language variation such as word-final *imaala*.

Table 3 presents all the free tokens of raa+taa marbuuTa (both raised and unraised) found in the speech of the 483 speakers mentioned above, and organized by word frequency (from more frequent to less frequent) on the base of those 252

[27] Labov, "Social Motivation of Sound Change," 273-274.

[28] Betty S. Phillips, "Fast Words, Slow Words," *American Speech* 75, no. 4 (2000): 414-415.

conversations. The focus here should be the correlation (or lack thereof) between word frequency (indicated by the number of tokens for each word-form) and the percentage of *imaala* within the same word-form. Note that the transcription codes used in the LDC differ from the more conventional ones. For a complete list of correspondences between Arabic letters and the LDC transcription codes, refer to the Appendix. Also note that words are listed exactly as they appear in the LDC transcriptions. For the sake of consistency, all words are listed in their unraised form (ending in -ap).

Word-forms (with a few minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa-marbuuTa raising in decimals (i.e. 1 = 100%)
bukrap بكرة	321	321	0	0.00
mar~ap مرة	192	192	0	0.00
Als~ay~aArap السيارة	187	178	9	0.05
fikrap فكرة	55	54	1	0.02
Ea$arap عشرة	47	35	12	0.26
fatrap فترة	42	41	1	0.02
zgiyrap صغيرة	41	5	36	0.88
zyaArap زيارة	38	37	1	0.03
kbiyrap كبيرة	32	6	26	0.81
sahrap سهرة	32	29	3	0.09
dawrap دورة	20	20	0	0.00
Almarap المرة	19	19	0	0.00

Table 3-All instances of raa+taa marbuuTa organized by word frequency

Table 3 (continued)

bandawrap بندورة	19	19	0	0.00
layrap ليرة (and liyrap ليرة)	18	18	0	0.00
Suwrap صورة	17	17	0	0.00
SaAyrap صايرة	16	11	5	0.31
msaAfrap مسافرة	14	4	10	0.71
ktiyrap كتيرة	13	3	10	0.77
saArap سارة	12	12	0	0.00
$aATrap شاطرة	12	5	7	0.58
Almaskarap المسكرة	11	11	0	0.00
Alt~ijaArap التجارة	10	9	1	0.10
siyrap سيرة	10	0	10	1.00
Aln~imrap النمرة	10	7	3	0.30
AlHaraArap الحرارة	9	8	1	0.11
Alfujayrap الفجيرة	9	5	4	0.44
HaArap حارة	8	8	0	0.00
tan~uwrap تنورة	8	8	0	0.00
safrap سفرة	8	8	0	0.00
AlT~ay~aArap الطيارة	8	8	0	0.00
xsaArap خسارة	8	8	0	0.00
mubaA$arap مباشرة	7	7	0	0.00
mfakrap مفكرة	7	5	2	0.29
EibaArap عبارة	6	6	0	0.00
samiyrap سميرة	6	6	0	0.00
AlHur~ap الحرة	6	6	0	0.00
xibrap خبرة	6	6	0	0.00

Table 3 (continued)

Als~afaArap السفارة	6	6	0	0.00
xudrap خضرة	6	6	0	0.00
msak~arap مسكرة	6	1	5	0.83
iswaArap إسوارة	6	6	0	0.00
qadrap قادرة	5	5	0	0.00
naATrap ناطرة	5	5	0	0.00
muniyrap منيرة	5	5	0	0.00
AlT~anjarap الطنجرة	5	5	0	0.00
Ald~aktawrap الدكتورة	5	5	0	0.00
AlEamaArap العمارة	5	5	0	0.00
fat~uwrap قتورة	5	5	0	0.00
idaArap إدارة	4	1	3	0.75
sikaArap سيجارة	4	4	0	0.00
Aln~am~uwrap النمورة	4	4	0	0.00
Alt~azkarap التذكرة	4	4	0	0.00
xiyrap خيرة	3	0	3	1.00
AlGSaArap القصارة	3	3	0	0.00
Als~ufrap السفرة	3	2	1	0.33
vawrap ثورة	3	3	0	0.00
xiTrap خطرة	3	3	0	0.00
mastuwrap مستورة	3	3	0	0.00
>Al i$rap القشرة	3	3	0	0.00
>aTrap قطرة	3	3	0	0.00
bawdrap بودرة	3	3	0	0.00

Table 3 (continued)

AlmanaArap المنارة	3	3	0	0.00
daw~aArap دوارة	3	3	0	0.00
Al*~urap الذرة	3	3	0	0.00
mitzakrip متذكرة	3	2	1	0.33
naD~aArap نظارة	3	3	0	0.00
Al>aSiyrap القصيرة	3	2	1	0.33
AlmaraArap المرارة	3	3	0	0.00
mHaDrap محضرة	3	3	0	0.00
Aljmayrap الجميرة	3	2	1	0.33
Aljaziyrap الجزيرة	3	2	1	0.33
maHZuwrap محظورة	3	2	1	0.33
mitwafrap متوفرة	3	2	1	0.33
>ibrap إبرة	3	0	3	1.00
HaAr~ap حارة	2	2	0	0.00
giyrap غيرة	2	0	2	1.00
Almudiyrap المديرة	2	2	0	0.00
xamiyrap خميرة	2	0	2	1.00
daAyrap دايرة	2	1	1	0.50
zaEbarap زعبرة	2	2	0	0.00
zahrap زهرة	2	2	0	0.00
masTarap مسطرة	2	2	0	0.00
samaArap سمارة	2	2	0	0.00
Al>i$aArap الإشارة	2	2	0	0.00
mitTawrap متطورة	2	1	1	0.50

Table 3 (continued)

عورة Eawrap	2	2	0	0.00
قزدورة Gazduwrap	2	2	0	0.00
ظاهرة ZaAhrap	2	0	2	1.00
الغبرة Algabrap	2	2	0	0.00
مكسرة mkasrap	2	0	2	1.00
منورة mnawrap	2	0	2	1.00
حاضرة HaADrap	2	2	0	0.00
الجورة Alj~awrap	2	2	0	0.00
الأخيرة Al>axiyrap	2	1	1	0.50
دكاترة dakaAtrap	2	1	1	0.50
هترة hitrap	2	0	2	1.00
مسكرة msakrap	1	1	0	0.00
مأثرة m>avrap	1	1	0	0.00
مؤثرة m>asrap	1	1	0	0.00
المستأجرة Almista>jirap	1	1	0	0.00
متأخرة mit>ax~irap	1	1	0	0.00
لأسرة Al>usrap	1	1	0	0.00
الأميرة Al>amiyrap	1	1	0	0.00
أمورة >am~uwrap	1	1	0	0.00
إمارة >imaArap	1	1	0	0.00
الباخرة AlbaAxrap	1	0	1	1.00
بربارة burbaArap	1	1	0	0.00
بشرة ba$rap	1	1	0	0.00
البقرة AlbaGarap	1	1	0	0.00
البيرة Albiyrap	1	0	1	1.00

Table 3 (continued)

majbuwrap مجبورة	1	1	0	0.00
mjad~arap مجدرة	1	1	0	0.00
Hajra حجرة	1	1	0	0.00
Ha*irap حذرة	1	1	0	0.00
Hasrap حسرة	1	1	0	0.00
maHSuwrap محصورة	1	1	0	0.00
AlHumrap الحومرة	1	1	0	0.00
miHtaArap محتارة	1	1	0	0.00
AlmuxaAbarap المخابرة	1	1	0	0.00
xityaArap ختيارة	1	1	0	0.00
xTuwrap خطورة	1	1	0	0.00
Ald~aA>irap الدائرة	1	1	0	0.00
zaEtarap زعترة	1	1	0	0.00
Alz~afrap الزفرة	1	1	0	0.00
Almasxarap المسخرة	1	1	0	0.00
sarsarap سرسرة	1	1	0	0.00
sakrap سكرة	1	1	0	0.00
sik~arap سكرة	1	1	0	0.00
Als~ikirtayrap السكرتيرة	1	1	0	0.00
Als~anyuwrap السنيورة	1	1	0	0.00
msayTirap مسيطرة	1	1	0	0.00
Al$~ajarap الشجرة	1	1	0	0.00
Al$~axtuwrap الشختورة	1	1	0	0.00
$aT~uwrip شطورة	1	0	1	1.00
$aT~uwrap شطورة	1	1	0	0.00

Table 3 (continued)

الشعرة Al$~aErap	1	1	0	0.00
مشكورة ma$kuwrap	1	1	0	0.00
مشهورة ma$huwrap	1	1	0	0.00
الشارة Al$~aArap	1	1	0	0.00
مصرة muSir~ap	1	1	0	0.00
مصفرة mSafrap	1	1	0	0.00
صفورة Sfuwrap	1	1	0	0.00
مضطرة miDTar~ap	1	1	0	0.00
ضهرة Dahrap	1	1	0	0.00
ضاهرة DaAhrap	1	1	0	0.00
الطائرة AlT~aA>irap	1	1	0	0.00
الظفرة AlZ~afrap	1	1	0	0.00
مظاهرة muZaAharap	1	1	0	0.00
عباقرة EabaAqirap	1	1	0	0.00
عسكرة Easakarap	1	1	0	0.00
عصيرة EaSiyrap	1	1	0	0.00
العصفورة AlEaSfuwrap	1	1	0	0.00
المغارة AlmagaArap	1	1	0	0.00
فرفورة farfuwrap	1	1	0	0.00
فقرة fa>rap	1	1	0	0.00
مقدرة m>ad~irap	1	1	0	0.00
قارة qaAr~ap	1	1	0	0.00
مقررة mqar~irap	1	0	1	1.00
المقررة Almuqar~arap	1	1	0	0.00
قزدرة Gazdarap	1	1	0	0.00

Table 3 (continued)

كرة kurap	1	1	0	0.00
كزبرة kizibrap	1	1	0	0.00
مكسورة maksuwrap	1	1	0	0.00
كفرة kafrap	1	1	0	0.00
كندرة kindarap	1	1	0	0.00
بالمرة biAlmar~ap	1	1	0	0.00
لنشرة Aln~a$rap	1	1	0	0.00
النطرة Aln~aTrap	1	1	0	0.00
هبرة habrap	1	1	0	0.00
البيرة Alyabrap	1	1	0	0.00
ميسرة myas~arap	1	1	0	0.00

As shown in Table 3, in spite of the fact that بكرة bukra is by far the most frequent -ra word in those 252 conversations, exhibiting a total of 321 tokens, not a single instance of word-final *imaala* is found in any of those tokens. The same is true of the second most frequent word, مرة mar~ap with 192 all unraised tokens. However سيرة siyrap, which was uttered only ten times, was raised in every single token, and حارة HaArap, which was uttered eight times, was not raised once. Therefore there seems to be no obvious correlation between word frequency and taa-marbuuTa raising in the data as it is presented in Table 3.

The total number of tokens presented in Table 3 is 1544, 184 of which actually exhibit taa-marbuuTa raising. Table 4 lists only those 184 tokens and has them

organized by percentage of word-final *imaala* (i.e. how often each word-form undergoes *imaala*), from most frequently raised to least frequently raised. By looking at the word-forms that exhibit the highest percentage of word-final *imaala*, one may be able to discern phonological patterns or word classes that seem to allow for taa-marbuuTa raising.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
siyrap سيرة	10	0	10	1.00
>ibrap إبرة	3	0	3	1.00
xiyrap خيرة	3	0	3	1.00
giyrap غيرة	2	0	2	1.00
xamiyrap خميرة	2	0	2	1.00
ZaAhrap ظاهرة	2	0	2	1.00
mkasrap مكسرة	2	0	2	1.00
mnawrap منورة	2	0	2	1.00
hitrap هترة	2	0	2	1.00
AlbaAxrap الباخرة	1	0	1	1.00
Albiyrap البيرة	1	0	1	1.00
$aT~uwrap شطورة	1	0	1	1.00
mqar~irap مقررة	1	0	1	1.00
zgiyrap صغيرة	41	5	36	0.88
msak~arap مسكرة	6	1	5	0.83
kbiyrap كبيرة	32	6	26	0.81

Table 4-All instances of word-final *imaala* organized from most frequently raised to least frequently raised

Table 4 (continued)

ktiyrap كتيرة	13	3	10	0.77
>idaArap إدارة	4	1	3	0.75
msaAfrap مسافرة	14	4	10	0.71
$aATrap شاطرة	12	5	7	0.58
mitTawrap متطورة	2	1	1	0.50
daAyrap دايرة	2	1	1	0.50
Al>axiyrap الأخيرة	2	1	1	0.50
dakaAtrap دكاترة	2	1	1	0.50
Alfujayrap الفجيرة	9	5	4	0.44
Als~ufrap السفرة	3	2	1	0.33
mitzakrap متذكرة	3	2	1	0.33
Al>aSiyrap القصيرة	3	2	1	0.33
Aljmayrap الجميرة	3	2	1	0.33
Aljaziyrap الجزيرة	3	2	1	0.33
maHZuwrap محظورة	3	2	1	0.33
mitwafrap متوفرة	3	2	1	0.33
SaAyrap صايرة	16	11	5	0.31
Aln~imrap النمرة	10	7	3	0.30
mfakrap مفكرة	7	5	2	0.29
Ea$arap عشرة	47	35	12	0.26
AlHaraArap الحرارة	9	8	1	0.11
Alt~ijaArap التجارة	10	9	1	0.10
sahrap سهرة	32	29	3	0.09
Als~ay~aArap السيارة	187	178	9	0.05

Table 4 (continued)

زيارة zyaArap	38	37	1	0.03
فترة fatrap	42	41	1	0.02
فكرة fikrap	55	54	1	0.02

Table 4 reveals that the raising of taa marbuuTa after raa remains highly variable overall. The rules expressed by such experts as Grotzfeld, Cantineau, Cowell etc. cannot accurately account for this variation. For instance, Cowell states that it is possible for words in the fəʕla form to undergo *imaala*,[29] but he does not make a word-to-word distinction within that category, whereas Table 3 shows very different patterns within the fəʕla form: for example, النمرة Aln~imrap 'number' appears ten times and undergoes *imaala* three times (30% of the time), in contrast with فكرة fikrap 'idea' which was uttered fifty-five times but was only raised once (0.02%) according to the transcriptions. This, of course, constitutes a significant difference. Other differences can be detected by comparing the behavior of other word-forms as well.

Within this sample of tokens exhibiting taa-marbuuTa raising (Table 4), one can identify patterns and phonological environments which allow for word-final *imaala*. A total of four *imaala*-friendly word classes can be established. Here is a brief description

[29] Cowell, *Reference Grammar*, 138.

of each: 1) words that end in long i + ra (long I will be represented by ii), 2) words that end in short i + ra, 3) words that end in the diphthong ay + ra, and 4) the fiʕla form. It seems like in every case, the high front vowel ii/i or the high front glide y impedes the tafkhiim of r and thus allows for vowel raising after r. Table 5 below shows the data from Table 4 organized according to those word classes. Twelve tokens from Table 4 do not fit in any one of those word classes and therefore will not appear in the following discussion. They will be dealt with in a later section of this paper. Keep in mind that although the words in Table 5 are all represented in their unraised form, they all exhibit at least one raised token in the LDC.

ii+r+a	i+r+a	ay+r+a	fiʕla form
giyrap غيرة			
xamiyrap خميرة	ZaAhrap ظاهرة		
Albiyrap البيرة	AlbaAxrap الباخرة		
zgiyrap صغيرة	mqar~irap مقررة		
kbiyrap كبيرة	mitTaw~rap متطورة		
ktiyrap كتيرة	msaAfrap مسافرة		
xiyrap خيرة	$aATrap شاطرة		
siyrap سيرة	dakaAtrap دكاترة	daAyrap دايرة	Al>ibrap الإبرة
Al>axiyrap الأخيرة	mitzakrap متذكرة	Alfujayrap الفجيرة	hitrap هترة
Al>aSiyrap القصيرة	mitwafrap متوفرة	Aljmayrap الجميرة	Aln~imrap النمرة
Aljaziyrap الجزيرة	mfakrap مفكرة	SaAyrap صايرة	fikrap فكرة

Table 5-Word Classes that Allow for Word-Final Vowel Raising

The -iira word class does not present any complications. As mentioned earlier, it is a known fact that i (short or long) impedes r from becoming emphatic. In contrast, looking at the -ira word class leads to an interesting observation: many of the forms in this word class actually do not contain any i sound right before the ending -ra. The reason for this is that it is very common for Levantine dialects to drop kasras and dammas where they are not absolutely necessary to syllable structure and the breaking of consonant clusters. Thus a form like dakaAtirap 'doctors' is usually pronounced dakaAtrap, and asaAtizap 'professors' becomes asaAtzip. Of course, this phonological rule affects many active participles since those take kasra as their stem vowel: faa3il in form I (e.g. saakin 'living,' Taaliʕ 'going out, turning out'), m(u/i)fa33il in form II (mʕazzib 'torturing, bothering,' mbayyiD 'whitening, making happy'), mu/itfa33il in form V (mitlabbik 'confused, hesitant,' mitzakkir 'remembering'), etc. In the feminine, the word-final marker taa marbuuTa creates an additional syllable in these active participles. In turns, speakers often delete the kasra vowel, which is the trademark of active participles, possibly to retain the same number of syllables as the masculine, thus yielding fa3la/e <faa3ila/e in form I (sakne, Talʕa), for instance. In forms II and V, this kasra dropping creates a new consonant cluster which is then reduced by way of degemination of the doubled consonant: m(u/i)fa3la/e <m(u/i)fa33ila/e in form II (mʕazbe, mbayDa), and mu/itfa3la/e <mu/itfa33ila/e in

form V (mitlabke, mitzakra/e). It is therefore not surprising that all of the words in the -ira word class but one (dakaAtrap) be based on the active participle form. This all means that as long as the non-raised final a is the original, which is the consensus among Arabists and Semiticists, then rule ordering reveals that word-final *imaala* has to have taken place before short i was dropped, which is a nice diachronic insight into the history of the Levantine dialects. Taa marbuuTa raising in the -(i)ra word class takes place even in the vicinity of emphatic consonants and other sounds like long u which are known to make r emphatic: مقررة mqar~irip, متطورة mtTaw~rip, شاطرة $aATrip.

The -ayra word class is in agreement with Barbot's tafkhiim rules, which state that r remains non-emphatic when it is in contact with e<i, or e:<ay, or y.[30] It appears that in some instances, this rule takes precedence over the presence of sounds that are known to lead to the tafkhiim or f, such as Saad in صايرة SaAyrip.

Finally, the fiʕla noun class presents no particular difficulties. The presence of i immediately preceding the consonant cluster impedes tafkhiim, which can lead to the raising of final a.

It seems that the four word classes established in this section agree with the sound rules expressed in studies such as Cantineau's, Grotzfeld's or Barbot's in the

[30] Barbot, *Evolucion Arabe Contemporain*, 396.

sense that those studies successfully identify the role of ii, i and y in preventing the velarization of r, which in turns allows for variation in the production on taa marbuuTa in the word-final -ra sequence. However, as pointed out earlier, these sound rules only partly account for what is really happening with word-final *imaala* in Levantine Arabic; all they do is single out phonological environments that can allow for taa-marbuuTa raising after raa. They teach that when the r in -ra is velarized, word-final *imaala* cannot take place, but when it is not, then word-final *imaala* can take place. The raising does not *necessarily* take place in those environments, rather it is simply not impossible. Now, there is, of course, a certain degree of dialect-to-dialect variation, but some trends are shared by all Levantine dialects and can be identified as common features between them: for example, the raising of taa marbuuTa in kbiire or ktiire is very common, whereas it is extremely uncommon in words such as mudiira or amiira. How can that be? Is this an exception to the sound rules, where r is velarized in spite of the presence of ii in mudiira and amiira, thus preventing the *imaala*, or does this have nothing to do with the velarization or non-velarization of raa, and possibly something to do with other factors such as lexical diffusion and word frequency?

Table 3, which contained all tokens of words ending in -ra both raised and unraised, showed no obvious correlation between word frequency and *imaala*. But some of those word-forms presented phonological environments which made taa-marbuuTa raising phonetically very unlikely or even impossible, because of the

velarization of r. Maybe focusing solely on those word-forms which present a phonological environment that allows for variation in the production of final taa marbuuTa can give us a slightly different picture of the situation. Table 6 presents all such word-forms organized by decreasing word frequency, regardless of whether those word-forms actually exhibit any instances of *imaala* in the data. All words in this table are represented in their unraised form, but the "Percentage of taa-marbuuTa raising" column indicates whether or not each word-form exhibited *imaala*.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
fikrap فكرة	55	54	1	0.02
zgiyrap صغيرة	41	5	36	0.88
kbiyrap كبيرة	32	6	26	0.81
layrap ليرة (and liyrap (ليرة	18	18	0	0.00
SaAyrap صايرة	16	11	5	0.31
msaAfrap مسافرة	14	4	10	0.71
ktiyrap كتيرة	13	3	10	0.77
$aATrap شاطرة	12	5	7	0.58
siyrap سيرة	10	0	10	1.00
Aln~imrap النمرة	10	7	3	0.30
Alfujayrap الفجيرة	9	5	4	0.44
mfakrap مفكرة	7	5	2	0.29
xibrap خبرة	6	6	0	0.00

Table 6-All word-forms that could phonologically allow for word-final *imaala*, organized by decreasing word frequency

Table 6 (continued)

قادرة qadrap	5	5	0	0.00
ناطرة naATrap	5	5	0	0.00
خيرة xiyrap	3	0	3	1.00
خطرة xiTrap	3	3	0	0.00
القشرة i$rap>Al	3	3	0	0.00
متذكرة mitzakrap	3	2	1	0.33
القصيرة Al>aSiyrap	3	2	1	0.33
الجميرة Aljmayrap	3	2	1	0.33
الجزيرة Aljaziyrap	3	2	1	0.33
متوفرة mitwafrap	3	2	1	0.33
الإبرة Al>ibrap	3	0	3	1.00
غيرة giyrap	2	0	2	1.00
المديرة Almudiyrap	2	2	0	0.00
خميرة xamiyrap	2	0	2	1.00
دايرة daAyrap	2	1	1	0.50
متطورة mitTawrap	2	2	0	0.00
ظاهرة ZaAhrap	2	0	2	1.00
منورة mnawrap	2	0	2	1.00
حاضرة HaADrap	2	2	0	0.00
الأخيرة Al>axiyrap	2	1	1	0.50
دكاترة dakaAtrap	2	1	1	0.50
هترة hitrap	2	0	2	1.00
مأثرة m>avrap	1	1	0	0.00
المستأجرة Almista>jirap	1	1	0	0.00

Table 6 (continued)

mit>ax~irap متأخرة	1	1	0	0.00
Al>amiyrap الأميرة	1	1	0	0.00
AlbaAxrap الباخرة	1	0	1	1.00
Albiyrap البيرة	1	0	1	1.00
Ha*irap حذرة	1	1	0	0.00
Ald~aA>irap الدائرة	1	1	0	0.00
Als~ikirtayrap السكرتيرة	1	1	0	0.00
muSir~ap مصرة	1	1	0	0.00
mSafrap مصفرة	1	1	0	0.00
DaAhrap ضاهرة	1	1	0	0.00
AlT~aA>irap الطائرة	1	1	0	0.00
EabaAqirap عباقرة	1	1	0	0.00
EaSiyrap عصيرة	1	1	0	0.00
mutaTaw~irap متطورة	1	0	1	1.00

Much like Table 3, Table 6 does not seem to reveal an obvious connection between word frequency and taa-marbuuTa raising; even looking only at the first few word-forms in the list makes this point clear.

However, there is yet another way to look at the relation between the data in Table 6 and word frequency, and that is through the lens of the four word classes which were identified earlier. Tables 7 through 12 will focus on word frequency within

each individual word class. Table 7 presents all word-forms from the data (both raised and unraised) that fit into the fiʕla word class.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
fikrap فكرة	55	54	1	0.02
Aln~imrap النمرة	10	7	3	0.30
xibrap خبرة	6	6	0	0.00
Al>i$rap القشرة	3	3	0	0.00
xiTrap خطرة	3	3	0	0.00
Al>ibrap الإبرة	3	0	3	1.00
hitrap هترة	2	0	2	1.00

Table 7- All word-forms of fiʕla word class organized by decreasing word frequency

When looking at Table 7, it is important to keep a few things in mind: first of all, loanwords are known to block word-final *imaala* so they should be ignored here,[31] and second, word-forms that were uttered only a few times do not show results that are statistically significant. In other words, the higher the number of tokens, the more representative the sample will be, and the more relevant the results then become. The line must be drawn somewhere, and for the purpose of this study, it was decided that

[31] Ferguson and Ani, *Damascus Arabic*, 116.

the key number would be ten: word-forms that present less than ten tokens should be ignored. By applying these principles to Table 7, one obtains Table 8.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
fikrap فكرة	55	54	1	0.02
Aln~imrap النمرة	10	7	3	0.30

Table 8-Relevant word-forms of fiʕla word class organized by decreasing word frequency

In order to show how one ought to approach the data to decide whether word frequency could be a relevant factor in the distribution of word-final *imaala*, let us apply the idea of word frequency to the results found in Table 8, knowing that two word-forms are not enough to come to any steady conclusion. From the results shown in Table 8, it seems like word frequency could possibly be playing a role in the distribution of taa-marbuuTa raising in this word class. The word فكرة fikrap is practically always pronounced without word-final *imaala* whereas النمرة Aln~imrap, which was used more than five times less frequently in the data, shows more instances of taa-marbuuTa raising. This may indicate that within the fiʕla word class, the least frequent words are changing first, but of course this theory needs more testing.

The next table presents all word-forms from the data (both raised and unraised) that belong to the -ayra word class.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
SaAyrap صايرة	16	11	5	0.31
Alfujayrap الفجيرة	9	5	4	0.44
Aljmayrap الجميرة	3	2	1	0.33
daAyrap دايرة	2	1	1	0.50

Table 9- All word-forms of -ayra word class organized by decreasing word frequency

When applying the two filters mentioned above (foreign words and word-forms with less than ten tokens), only one word-form is left in the –ayra word class (صايرة SaAyrap), which is too little data to draw any conclusions.

Table 10 is a list of all word-forms from the data (both raised and unraised) that phonologically fit into the -ira word class.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
msaAfrap مسافرة	14	4	10	0.71
$aATrap شاطرة	12	5	7	0.58
mfakrap مفكرة	7	5	2	0.29
qadrap قادرة	5	5	0	0.00
naATrap ناطرة	5	5	0	0.00
mitzakrap متذكرة	3	2	1	0.33

Table 10- All word-forms of -ira word class organized by decreasing word frequency

Table 10 (continued)

متوفرة mitwafrap	3	2	1	0.33
متطورة mitTawrap	3	2	1	0.33
ظاهرة ZaAhrap	2	0	2	1.00
منورة mnawrap	2	0	2	1.00
حاضرة HaADrap	2	2	0	0.00
دكاترة dakaAtrap	2	1	1	0.50
مأثرة m>avrap	1	1	0	0.00
المستأجرة Almista>jirap	1	1	0	0.00
متأخرة mit>ax~irap	1	1	0	0.00
الباخرة AlbaAxrap	1	0	1	1.00
حذرة Ha*irap	1	1	0	0.00
مصفرة mSafrap	1	1	0	0.00
ضاهرة DaAhrap	1	1	0	0.00
الطائرة AlT~aA>irap	1	1	0	0.00
عباقرة EabaAqirap	1	1	0	0.00
مقدرة m>ad~irap	1	1	0	0.00
مقررة mqar~irap	1	0	1	1.00
الدائرة Ald~aA>irap	1	1	0	0.00

After taking out foreign words and word-forms uttered less than ten times, only two word-forms are left, which is too few to see any kind of pattern.

Finally, Table 11 lists all word-forms from the data (both raised and unraised) that fit into the -iira word class.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
صغيرة zgiyrap	41	5	36	0.88
كبيرة kbiyrap	32	6	26	0.81
كتيرة ktiyrap	13	3	10	0.77
سيرة siyrap	10	0	10	1.00
خيرة xiyrap	3	0	3	1.00
القصيرة Al>aSiyrap	3	2	1	0.33
الجزيرة Aljaziyrap	3	2	1	0.33
غيرة giyrap	2	0	2	1.00
المديرة Almudiyrap	2	2	0	0.00
خميرة xamiyrap	2	0	2	1.00
الأخيرة Al>axiyrap	2	1	1	0.50
الأميرة Al>amiyrap	1	1	0	0.00
البيرة Albiyrap	1	0	1	1.00
السكرتيرة Als~ikirtayrap	1	1	0	0.00
عصيرة EaSiyrap	1	1	0	0.00

Table 11-All word-forms of -iira word class organized by decreasing word frequency

If loanwords are left out along with word-forms that exhibit less than ten tokens, Table 11 turns into Table 12 below.

Word-forms (with some minor variations in vowels)	Number of Tokens	Taa marbuuTa: unraised	Taa marbuuTa: raised	Percentage of taa marbuuTa-raising in decimals (i.e. 1 = 100%)
zgiyrap صغيرة	41	5	36	0.88
kbiyrap كبيرة	32	6	26	0.81
ktiyrap كتيرة	13	3	10	0.77
siyrap سيرة	10	0	10	1.00

Table 12-Relevant word-forms of -iira word class organized by decreasing word frequency

Although word frequency is certainly not the only factor affecting the distribution of taa-marbuuTa raising in Levantine Arabic, as demonstrated by what is happening with siyrap سيرة, Table 12 shows a nice gradation of word frequency and word-final *imaala* among the three most common words in the -iira word class. Contrary to the fiʕla word class (Table 8), within which less frequent words may be changing the fastest, the -iira word class seems to indicate that something is happening with the most frequent words (zgiyrap صغيرة, kbiyrap كبيرة and ktiyrap كتيرة). Overall, word frequency does not seem to play a clear role in the distribution of taa-marbuuTa raising in the Levantine dialects, although the possible trends revealed in Tables 8 and 12 deserve further investigation in studies that provide a larger pool of tokens of -ra words.

IV. Problematic Data

The four word classes mentioned above (-iira, -ira, -ayra, fiʕla) seem well fit to explain the process of taa-marbuuTa raising in the 29 word-forms mentioned in Table 5. However, as explained earlier, Table 4 still contains twelve additional word-forms which remain unaccounted for since they do not fit in any of those word classes. That word-final *imaala* would take place in those twelve words will seem odd to those acquainted with the Levantine dialects, since those words do not present a phonological environment which can justify the non-velarization of r. Table 13 is a list of those twelve problematic word-forms in their raised form.

Problematic Word-Forms
$aT~uwrip شطورة
msak~arip مسكرة
>idaArip إدارة
Als~ufrip السفرة
maHZuwrip محظورة
Ea$arip عشرة
AlHaraArip الحرارة
Alt~ijaArip التجارة
sahrip سهرة
Als~ay~aArip السيارة
zyaArip زيارة
fatriap فترة

Table 13-Problematic Raised Word-Forms

Keeping in mind the essential role of the high front vowel i (whether short or long) and the glide y in preventing r-velarization, one may attempt to explain some of the tokens in Table 13 by arguing, for example, that there may be some form of distant influence of i on r in words such as التجارة Alt~ijaArip and إدارة idaArip, and of y on r in السيارة Als~ay~aArip and زيارة zyaArip. But other sounds that are closer in distance to r in those words should take precedence and ensure the production of r as an emphatic. At this point, it becomes imperative to turn to the recordings themselves and ensure that these few tokens have been transcribed accurately.

One of the most important discoveries made during this study has to do with the fact that after careful examination of the recordings, a number of discrepancies related to word-final *imaala* were discovered between what the speakers actually said and what the transcribers heard and wrote down. Table 14 presents all the word-forms that exhibit taa-marbuuTa raising in the transcriptions (Table 4 data) divided into the four word classes identified earlier, plus a fifth group for those word-forms that do not fit into any of those word classes. Within each word class, words are further distributed into two columns: a word appears under "YES" if at least one of its tokens actually exhibited word-final *imaala* in the recordings. A word appears under "NO" if taa-marbuuTa raising did not actually take place in any of the tokens of that word in the recordings, even though the transcriptions say something else. Words in the "YES" column are listed in their raised form, and words in the "NO" column in their unraised form. This should make the table more clear.

	Taa-MarbuuTa Raising in Actual Recordings	
	YES	NO
ii+r+a	giyrip غيرة	Albiyrap البيرة
	xamiyrip خميرة	Al>aSiyrap القصيرة
	zgiyrip صغيرة	Aljaziyrap الجزيرة
	kbiyrip كبيرة	
	ktiyrip كتيرة	
	xiyrip خيرة	

Table 14- Realization of all word-forms transcribed as raised, as found in recordings

Table 14 (continued)

	سيرة siyrip	
	الأخيرة Al>axiyrip	
i+r+a	ظاهرة ZaAhrip	الباخرة AlbaAxrap
	مقررة mqar~irip	شاطرة $aATrap
	متطورة mitTaw~rip	دكاترة dakaAtrap
	مسافرة msaAfrip	متوفرة mitwafrap
	متذكرة mitzakrip	
	مفكرة mfakrip	
ay+r+a	دايرة daAyrip	الفجيرة Alfujayrap
	صايرة SaAyrip	الجميرة Aljmayrap
fiʔla form	الإبرة Al<ibrip	فكرة fikrap
	هترة hitrip	
	النمرة Aln~imrip	
Problematic Word-Forms	فترة fatrip	شطورة $aT~uwrap
		مسكرة msak~arap
		إدارة idaArap
		السفرة Als~ufrap
		محظورة maHZuwrap
		عشرة Ea$arap
		الحرارة AlHaraArap
		التجارة Alt~ijaArap
		سهرة sahrap
		السيارة Als~ay~aArap
		زيارة zyaArap

The first thing one notices when looking at Table 14 is the fact that mistakes were made by the transcribers on all sorts of words, some presenting one of the four *imaala*-friendly patterns identified earlier, and some not. Although it might be tempting to label these discrepancies between what was said and what was heard as simple mistakes and to subsequently ignore them altogether in our analysis, a more adequate approach would be to try and understand the possible reasons behind those mistakes. After all, the transcribers were all native speakers of Arabic, and if the recordings are unequivocal to a non-native speaker such as myself, then how could native speakers mishear so many tokens of taa-marbuuTa raising?

One possible explanation is that the transcribers have formed expectations as to what should and should not be raised based on the nature of each individual word. For example, they may be aware, probably not consciously, that the words that fit in the -iira pattern can phonologically undergo word-final *imaala*, and so when they came across the word البيرة Albiyrap they felt that the word should be raised and therefore they heard it as such.

The same applies to the -ira word class. All word-forms from that class that are actually raised in the recordings are active participles such as ZaAhrip ظاهرة and mqar~irip مقررة. When transcribers ran into other words that fit that pattern but were not raised (such as AlbaAxrap الباخرة and $aATrap شاطرة), they heard them as raised. By extension, even the word dakaAtrap دكاترة which is not an active participle but

belongs to the -ira word class, was heard as a raised token. The same goes for the other two word classes, -ayra and fiʕla: some unraised tokens that belong to those classes were perceived by the transcribers as raised tokens, probably because they belong to word classes that allow for word-final *imaala*.

Unfortunately, while this theory may explain why unraised word forms that fit phonologically in one of the four *imaala*-friendly word classes were sometimes identified as raised tokens by transcribers, it does not account for the twelve problematic tokens that do not fit in any of those word classes.

This brings the discussion to another possible explanation: perhaps the transcribers heard taa-marbuuTa raising on all of those unraised words simply because they have heard them raised before. When looking at the data from Table 14, one notices that two of the non-raised word forms that were mistaken for raised forms in the -iira word class, are in fact very commonly raised: Al>aSiyrap القصيرة and Aljaziyrap الجزيرة. It so happens that all the tokens of those words present in the LDC recordings were unraised, but it makes sense that someone transcribing the recordings would expect to hear word-final *imaala* in those two words since they are so commonly produced that way. This theory is impossible to prove or disprove without some further studies in linguistic variation and change where all variables can actually be controlled, but it is an essential hypothesis to at least consider.

In terms of the phonological likelihood of this theory, any word that falls under one of the four special word classes does not pose any particular problem since other words that present the same phonological environment are in fact raised. When looking at the twelve problematic tokens, however, further justification is needed. It is possible that word-final *imaala*, which was described in earlier days as a purely phonological phenomenon, has now crossed these phonological boundaries and by way of lexical diffusion, it has started to apply to other -ra words which do not present any of the phonological environments that have been described as allowing for taa-marbuuTa raising. This would mean that new environments that remain to be defined are now allowing for the non-emphatic production of raa and raising of taa marbuuTa. Although this would seem surprising, the concept of lexical diffusion is certainly not unprecedented in the world's languages. In addition, one problematic token from Table 14 seems to support this theory: fatrip فترة is considered problematic because based on the phonological rules established by the different scholars mentioned in the first half of this research, its phonological make-up cannot justify the non-emphatic production of raa and subsequent raising of taa marbuuTa. However, in the LDC recordings, one token of this word-form was indeed pronounced with word-final *imaala*. Of course, this could be a simple slip of the tongue, something the speaker himself found funny-sounding as it came out of his mouth, but he certainly did not correct himself. Maybe it is more than a mistake, maybe this token is just one indication of a larger sound change in progress. Perhaps in this speaker's dialect the

word fatrap/ip فترة exhibits a degree of variation and can be produced either as a raised or unraised entity. If this is true, then the same might also be true of some of the other problematic tokens listed in Table 14.

Finally, another possible explanation is that the transcribers might have formed expectations not based on individual words, but rather on individual speakers. Certainly some dialects tend to use word-final *imaala* more than others, so that when hearing someone speak one of those dialects, listeners automatically identify taa-marbuuTa raising as a strong distinguishing feature of this person's dialect. From there, listeners start expecting and hearing taa-marbuuTa raising in words where it is not actually there. If the LDC provided more complete demographic information on the speakers, and if the sample of the 483 speakers used for this study was a bit more balanced in terms of country of origin and age, then one would be able to see immediately which groups are more liberal with their use of *imaala*, and thus social meaning could be assigned to this phenomenon. In the current situation, with limited demographic data and further issues in sampling, it will not be possible to come to specific conclusions pertaining to which social factors may play a role in the use of word-final *imaala*. However, it may still be useful to try and extract some general tendencies connected to taa-marbuuTa raising in terms of social factors, since those tendencies could help form hypotheses which could be used as a base for future studies.

	YES--Raised forms	Speakers	NO--Non-raised forms transcribed as raised	Speakers
ii+r+a	giyrip غيرة	Lebanon 21 female	Albiyrap البيرة	Lebanon 37 male
	xamiyrip خميرة	Lebanon 37 female	Al>aSiyrap القصيرة	Lebanon 26 male
	zgiyrip صغيرة	Syria 29 male, Syria 26 male, Syria 34 female, Syria 21 male, Lebanon 16 male	Aljaziyrap الجزيرة	Jordan 25 male
	kbiyrip كبيرة	Syria 27 female, Syria 45 female, Syria 34 female, Syria 27 female, Syria 34 male, Syria 34 male, Lebanon 16 male		
	ktiyrip كتيرة	Syria 30 female, Syria 35 male		
	xiyrip خيرة	Lebanon 35 female, Lebanon 32 female		
	siyrip سيرة	Jordan 30 female		
	Al>axiyrip الأخيرة	Jordan 27 male		
i+r+a	ZaAhrip ظاهرة	Lebanon 19 female	AlbaAxrap الباخرة	Syria 26 female
	mqar~irip مقررة	Jordan 17 female	$aATrap شاطرة	Syria 48 female, Syria 34 female, Syria 28 female, Syria 19 female, Lebanon 27 female, Syria 27 female
	mitTaw~rip متطورة	Syria 25 female	dakaAtrap دكاترة	Syria 24 male
	mitzakrip متذكرة	Syria 24 male	mitwafrap متوفرة	Syria 25 female

Table 15- Realization of all word-forms transcribed as raised, as found in recordings, with speaker information

Table 15 (continued)

	mfakrip مفكرة	Syria 23 male, Syria 30 female		
	msaAfrip مسافرة	Syria 26 male, Syria 26 male, Lebanon 18 female, Lebanon 24 male, Syria 27 male, Jordan 30 female, Jordan 27 female, Syria 17 female		
ay+r+a	daAyrip دايرة	Palestine 39 female	Alfujayrap الفجيرة	Syria 36 male
	SaAyrip صايرة	Syria 24 male	Aljmayrap الجميرة	Syria 30 female
fiʕla form	Al<ibrip الإبرة	Jordan 30 female, Jordan 40 female	fikrap فكرة	Lebanon 33 male
	hitrip هترة	Jordan 21 male		
	Aln~imrip النمرة	Syria 28 male, Syria 32 male		
Problematic Word-Forms	fatrip فترة	Syria 25 male	$aT~uwrap شطورة	Syria 27 female
			msak~arap مسكرة	Syria 28 male
			idaArap إدارة	Lebanon 19 male
			Als~ufrap السفرة	Lebanon 28 male
			maHZuwrap محظورة	Jordan 30 female
			Ea$arap عشرة	Lebanon 37 male, Lebanon 19 female
			AlHaraArap الحرارة	Syria 35 male
			Alt~ijaArap التجارة	Jordan 21 male
			sahrap سهرة	Lebanon 23 male
			Als~ay~aArap السيارة	Lebanon 33 male
			zyaArap زيارة	Syria 25 female

Table 15 provides a good insight into which specific speakers were heard pronouncing which word-forms with word-final *imaala*. Table 16 below is a summary of the speakers in the "YES" category, meaning those speakers who did in fact pronounce taa-marbuuTa as raised in some tokens containing -ra. Speakers are organized by country of origin, gender and age. Boxes that do not contain any speakers were left blank for visual effect, though they could have been filled with zeros.

Gender	Age	Country of Origin				
		Lebanon	Syria	Jordan	Palestine	All Countries
MALES	15-19	2				2
	20-29	1	11	2		14
	30-39	2	4			6
	All Ages	5	15	2		22
FEMALES	15-19	2	1	1		4
	20-29	1	3	1		5
	30-39	1	4	3	1	9
	40-49		1	1		2
	All Ages	4	9	6	1	20
ALL GENDERS	15-19	4	1	1		6
	20-29	2	14	3		19
	30-39	3	8	3	1	15
	40-49		1	1		2
	All Ages	9	24	8	1	42

Table 16-Speakers who pronounced word-final *imaala* in (some) -ra words

Earlier in this study, it was established that the sample of speakers was not a balanced one in the sense that Palestinians are grossly underrepresented. We also saw in Figures 1 to 3 that there are slightly more males than females, and that the largest age group by far is the '20-29' group (representing 49.9% of the sample), followed by the '30-39' group (representing 23.8% of the sample). Some of the implications of

these unbalances are apparent in Table 16: Palestinians are underrepresented, males get slightly larger numbers than females, and within the 'Males' categories, it is true that the '20-29' age group is the most represented, followed by the '30-39' age group.

However, two interesting trends can be noticed which cannot have been caused by those unbalances. First of all, by looking at female speakers alone, it appears that the difference between the '20-29' age group and the '30-39' age group is reversed: instead of finding twice as many '20-29' speakers as there are '30-39' speakers, one finds five speakers in the '20-29' age group and nine in the '30-39' age group. For every country, the number of female speakers in the '30-39' group is equal or superior to female speakers in the '20-29' group. This is particularly pronounced for Jordanian speakers (one female speaker in the '20-29' group vs. three in the '30-39' group). The second trend appears when comparing countries to each other. Palestinian speakers should be left out since they are underrepresented, but the other countries, which are all roughly equally represented, show that Syrians are the most prone to *imaala* after raa: for male speakers, we find fifteen Syrians, five Lebanese and two Jordanians, and for females, we find nine Syrians, six Jordanians and four Lebanese. The largest group of speakers in any category in Table 15 is Syrian males in their twenties (eleven speakers).

Although these trends definitely deserve their own well-controlled study, they seem to indicate that women (in particular Jordanian women) in their thirties may be

more likely to use word-final *imaala* after raa than women in other age groups, and that Syrians (both males and females) use word-final *imaala* the most out of all Levantine speakers (at least in words ending in -ra).

It is now time to look at a summary of the speakers in the "NO" category from Table 15. These are speakers who uttered non-raised tokens of words ending in -ra, but those tokens were heard as raised tokens by the transcribers. Speakers-related information is found in Table 17 below.

Gender	Age	Country of Origin				
		Lebanon	Syria	Jordan	Palestine	All Countries
MALES	15-19	1				1
	20-29	3	2	2		7
	30-39	4	2			6
	All Ages	8	4	2		14
FEMALES	15-19	1	1			2
	20-29	1	6			7
	30-39		2	1		3
	40-49		1			1
	All Ages	2	10	1		13
ALL GENDERS	15-19	2	1			3
	20-29	4	8	2		14
	30-39	4	4	1		9
	40-49		1			1
	All Ages	10	14	3		27

Table 17-Speakers who were thought to pronounce word-final *imaala* in -ra words but did not

Table 17 shows that the transcribers' expectations are slightly different from the reality seen in Table 16 when it comes to word-final *imaala*. According to Table 17, in the transcribers' minds Jordanians use word-final *imaala* less than they do in reality (eight speakers in the 'All Genders' category in Table 16 vs. three in Table 17). For Lebanese speakers, transcribers' expectations match the results found in Table 16

pretty closely in the 'All Genders' category, but not in the separate 'Males' vs. 'Females' categories: transcribers seem to believe that Lebanese men use word-final *imaala* after raa more than they actually do, and that women use it less than they actually do.

The first trend that was noticed in Table 16 is not visible in Table 17: females in the '30-39' age group are expected by the transcribers to use word-final *imaala* less than females in the '20-29' group, which does not reflect the image of reality painted in Table 16. As for the second trend, transcribers appropriately sense that Syrian speakers are the most prone to use taa-marbuuTa raising after raa, as illustrated by the fact that fourteen Syrian speakers are found in the 'All Genders' category, followed by Lebanese speakers (ten) and finally Jordanians with three speakers. This fits well with what was noticed about Syrian speakers in Table 16. Palestinians should be ignored since, again, they are so poorly represented in the speaker pool as a whole. The largest number of speakers in any given category in Table 17 is Syrian females in their twenties with a total of six speakers.

While listeners' expectations reflect some aspects of the data-based reality described in Table 16, such as the tendency Syrian speakers have to use word-final *imaala* after raa more than other Levantine speakers, they also differ from that same reality in several regards: use of *imaala* by Jordanian speakers, Lebanese males vs. Lebanese females, and women in the '20-29' age group vs. women in the '30-39' age

group. Whether any of these trends hold true for larger pools of speakers will have to be established by further language attitude and other sociolinguistic studies. These studies will also need to define the exact social meaning of *imaala*, which is something that cannot be achieved in this paper due to the general lack of speaker-related information in the Levantine LDC database.

However, what can be said now is that the problematic data listed in Table 13, along with the overall discrepancies found between what speakers said and what transcribers heard, should not be dismissed as simple accidents. The fact that those mistakes happened repeatedly leaves no room for chance; something is happening, whether on the speakers' end or on the listeners' end. Even if no Levantine speaker has ever used *imaala* on any of the tokens that were misheard, which is very unlikely, it would not mean that the situation is not about to change. As a matter of fact, it is now generally accepted by linguists and most language professionals that speakers are not the only agents that participate in language change. Rather, speech-based communication involved two main elements: what the speaker says, and what acoustic message the listener happens to hear. Incomplete or mistaken perception on the listener's end has also been identified as an essential source for language variation

and change.[32] The fact that transcribers made those mistakes may be an indicator that something *imaala*-related is about to change in dialects of Levantine Arabic.

[32] John J. Ohala, "Phonetics and Historical Phonology," in *The Handbook of Historical Linguistics* (Oxford, UK: Blackwell Publishing Ltd, 2003), 672-673.

V. Conclusions

In this paper, I have shown that the phonological rules established by 19th- and 20th-century Arabists and Semiticists cannot always account for what is happening today in the language of Levantine Arabic speakers as far as word-final *imaala* is concerned. I have pointed out some of the quality issues we as a field need to remedy if we are to catapult Arabic language studies into the field of accredited scientific research. By looking at the data available through the LDC Levantine database, I have shown that a large degree of variation exists in the contemporary production of taa marbuuTa after raa, both word-form internally and word-class internally. This variation cannot be explained by phonology alone since the phenomenon of word-final raising after raa has spread to include new phonological environments. Rather, this variation extends to the realms of social meaning and identity. As noted by Labov, "social pressures are continually operating upon language, not from some remote point in the past, but as an immanent social force acting in the present."[33]

The mistakes found in the LDC in what was perceived by the transcribers as being word-final *imaala*, support the idea that we are witnessing a linguistic change in progress, and that the traditional frame used to describe the phenomenon is now insufficient. Due to a lack of control over the variables that make up the LDC Levantine database, coupled with a lack of speaker-related information, this study failed to

[33] Labov, "Social Motivation of Sound Change," 275.

establish the exact social meaning associated with word-final *imaala* in Levantine Arabic.

However, this study was successful in identifying some hypotheses which should serve as the base for a large-scale sociolinguistic study covering both speakers' evaluational reactions to word-final *imaala*, and the exact influencing factors which control variation in the production of taa-marbuuTa raising in the Levantine communities. Those hypotheses are: in the -ira class, less frequent words may be changing first; in the -iira class, most frequent words may be changing first (the study should also attempt to explain why this is the case, possibly by using some of Phillips' findings mentioned above as a starting point); female speakers in the '30-39' age group tend to use word-final *imaala* more than females in the '20-29' age group, especially Jordanians (although since the LDC Levantine database was collected in 2004, some of these speakers have probably moved up to the next age group by now); Syrians seem to be more prone to use *imaala* than Lebanese and Jordanians (in order to prove or disprove this, we will need some kind of geographical mapping of major trends in productions of word-final *imaala* in those three countries). Such a study requires that we start building a reliable database of sociolinguistic interviews with controlled variables. These interviews will have to be of sufficient quality to allow for phonetic measurements. They could be used both for form and content, since many of

them will undoubtedly shed light on local history, the social make-up of each community and issues of identity, prestige and social solidarity.

Appendix: Transcription

Arabic Letter	LDC Transcription	Transcription as Found in the Text
ء	>	ʔ
ب	b	b
ت	t	t
ث	v	th/s/t
ج	j/k	j
ح	H	H
خ	x	kh
د	d	d
ذ	*	dh/z
ر	r	r
ز	z	z
س	s	s
ش	$	ʃ
ص	S/z	S
ض	D	D
ط	T	T
ظ	D/Z	DH/Z
ع	E	ʕ
غ	g	gh
ف	f	f
ق	q/G	q
ك	k	k
ل	l	l
م	m	m
ن	n	n
ه	h	h
و	w (glide)/uw (vowel)	w (glide) /uu (vowel)
ي	y (glide)/iy (vowel)	y (glide)/ii (vowel)
ا	aA	aa
ة	ap (unraised)/ip (raised)	a (unraised)/e (raised)
ال	Al	al
fatHa	a	a
Damma	u	u
kasra	i	i
shadda	~	(letter is doubled, e.g. rr)
Other symbols: L represents emphatic laam and R emphatic raa.		

Works Cited

Barbot, M. *Evolucion de l'Arabe Contemporain*. Paris: La Sorbonne, 1981.

Cantineau, Jean. *Etudes de Linguistique Arabe*. Paris: Librairie C. Klincksieck, 1960.

Cowell, Mark W. *A Reference Grammar of Syrian Arabic*. Washington D.C.: Georgetown University Press, 1964.

Ferguson, Charles A. and Moukhtar Ani. *Damascus Arabic*. Washington D.C.: Center for Applied Linguistics, 1961.

Fischer, Wolfdietrich and Otto Jastrow. *Handbuch der Arabischen Dialekte*. Wiesbaden: Otto Harrassowitz, 1980.

Grotzfeld, Heinz. *Laut- und Formenlehre des Damaszenisch-Arabischen*. Wiesbaden: Kommissionsverlag Franz Steiner GMBH, 1964.

Grotzfeld, Heinz. *Syrisch-Arabische Grammatik (Dialekt von Damaskus)*. Wiesbaden: Otto Harrassowitz, 1965.

Labov, William. "The Social Motivation of a Sound Change." *Word* 19 (1963): 273-309.

McCarus, Ernest and Hamdi Qafisheh and Raji Rammuny. *A Course in Levantine Arabic*. Ann Harbor: the University of Michigan, 1978.

Ohala, John J. "Phonetics and Historical Phonology." In *The Handbook of Historical Linguistics*. Oxford, UK: Blackwell Publishing Ltd, 2003.

Phillips, Betty S. "Fast Words, Slow Words." *American Speech* 75, no. 4 (2000): 414-416.

Stowasser, Karl and Moukhtar Ani, ed. *Dictionary of Syrian Arabic*, Washington D.C.: Georgetown University Press, 1964.

Thackston, W.M., Jr. *The Vernacular Arabic of the Lebanon*. Massachusetts: Harvard University, 2003.