The Emergence of New Emphatics in Moroccan Arabic

Aaron Freeman*

I propose to study the phonological behavior of secondarily pharyngealized and velarized consonants in Colloquial Moroccan Arabic (CMA), focusing on the liquid /r/ and labial obstruents /b f m/. In certain contexts, these sounds independently acquire phonetic characteristics similar to those of the existing emphatic 'pharyngealized' phonemes /ṭ ḍ ṣ ẓ/, and for many speakers a combination of borrowing and analogy has extended the context of emphatic variants outside of the original conditioning environment, resulting in new emphatic phonemes. Through interviews with individual speakers, I aim to establish the mechanisms of the phonological change and to evaluate the phonemic character of these segments through processes associated with phonological emphasis, as well as investigating the effect of individual differences in phonetic distribution on systematic differences in an individual's grammar.

1. Background and Theoretical Context

1.1 Emphasis and Pharyngealization

The consonantal feature known as emphasis is one of the more idiosyncratic characteristics of the Arabic language. It consists of a secondary consonantal articulation distinguishing emphatic consonants from 'plain' counterparts, which is generally identified phonetically as pharyngealization. In Modern Standard Arabic (MSA), phonemically contrastive emphasis is restricted to coronal obstruents, such that there is a phonemic contrast between /t d s ð/ and /ṭ ḍ ṣ ð̣/. The coronal obstruents /θ z/ do not have emphatic counterparts in MSA, but in CMA /θ/ has been lost and emphatic /ẓ/ has developed into a separate phoneme, so that all coronal obstruents in CMA with the exception of post-alveolars exist as both plain and emphatic phonemes. The exclusion of post-alveolars /ʃ ʒ/ from this class is not particularly surprising, since CMA /ʒ/ is historically derived from a velar, and there is evidence that both /ʃ/ and /ʒ/ were pronounced as palatals in earlier stages of Arabic (Al-Nassir 1993).

If the history of emphasis is traced back to proto-Semitic, it will be found that the feature was originally restricted to voiceless lingual obstruents, rather than coronal

* I would like to especially thank my proposal committee, Rolf Noyer, Mark Liberman, and Meredith Tamminga, and my advisor Don Ringe for bringing this project to where it is today, as well as Bill Labov, Stuart Davis, Mbarek Sryfi, Alex Magidow, and Mohamed Maamouri for their valuable comments and suggestions. Among my fellow students, I am particularly grateful to Akiva Bacovcin and Aletheia Cui for their invaluable programming help, and to Amy Goodwin Davies and Betsy Sneller for their unflagging support, patience, and insightful feedback.

obstruents. The MSA emphatics /ð̣/ and /ḍ/ derive from the pSem emphatic voiceless interdental and lateral sibilant respectively, and the Arabic uvular stop /q/ is historically the emphatic counterpart of /k/. It has been suggested, based on this pattern and comparative data from Ethiopic and non-Semitic Afro-Asiatic languages, that the emphatics may originally have been glottalic ejectives rather than pharyngealized consonants (Watson 2007; Diakonoff 1965). Both articulations are attested in modern languages inheriting the feature, while in some languages, such as Aramaic and modern Hebrew, it has been lost entirely.

Even within Arabic, the articulatory correlates of emphasis have been debated. While a majority of studies have indicated that constriction of the upper pharynx is the essential feature of consonantal emphasis, others have identified emphasis with velarization or uvularization instead (Zawaydeh & de Jong 2011). In addition, specific emphatic consonants have been associated with other features, such as labialization for /ḍ/ (Zeroual et al. 2011) and glottalization for /ṭ/ (Schroepfer 2015). All previous research has straightforwardly identified CMA emphasis with pharyngealization, but with a locus of pharyngeal constriction that is distinct from the that of primary pharyngeal consonants /ħ/ and /ʕ/. Based on laryngoscopic evidence, Esling (1996) claims that secondary pharyngeal articulation involves lingual constriction towards the upper pharynx wall, whereas /ħ/ and /ʕ/ involve constriction of the lower pharynx by the epiglottis and associated structures. Because of these ambiguities, and in order to avoid reification of the term in phonetics, I will continue to refer to the distinctive phonological feature here as emphasis (or consonantal emphasis) rather than pharyngealization, and will adopt the convention of transcribing it as a subscript dot.

Although contrastive or primary emphasis is typically restricted to coronal obstruents in Arabic dialects, the pharyngealization associated with emphasis has a strong tendency to spread to adjacent segments, giving rise to pharyngealized allophones of both vowels and other consonants. Pharyngealized consonantal allophones resulting from emphasis spread, which may include labials [b̪ f̪ m̪] and coronals [ɾ̪ l̪ n̪ ʃ̪ ʒ̪], are typically referred to as 'secondary emphatics'. The vowel allophones found in proximity to emphatic consonants are lowered and/or backed, and provide the clearest perceptual cues for emphasis. Heath describes the CMA allophone of /a/ roughly as [ɑ], of /i/ as [ɨ], and of /u/ as [o].

Patterns of emphasis spread vary widely across dialects. Sibilants, pharyngeals, and high vowels have all been observed to block emphasis spread (Al-Masri and Jongman 2004), while in other dialects spread of pharyngealization has been observed to be restricted to one syllable rightwards but may extend further to the left (Watson 2007). In CMA, only certain morphological boundaries are consistently reported to block emphasis spread, particularly boundaries with clitic pronouns and verbal inflections (Heath 1987). This broad range of spread has led authors such as Dell and Elmedlaoui (2002) to propose that emphasis is a word-level suprasegmental feature, rather than a consonantal feature. However, a significant subset of Moroccan dialects, typically Southern, observe emphasis dissimilation within a stem, and Heath (1987) further argues that even in CMA dialects with unfettered spread of

emphasis there may be co-occurrences of plain and emphatic consonants within the same stem.

The segments under consideration here are of interest because they are neither primary (phonemic) nor secondary (regularly conditioned allophonic) emphatics, but are consonantal variants with features similar to emphasis arising independently of interaction with emphatic coronal obstruents. One, the pharyngealized [ɾ̞], is present across Arabic dialects as an allophone adjacent to the vowel /a/ and a variable set of other sounds, including the vowel /u/ and guttural consonants /q χ ɣ ʕ ħ/. In certain CMA dialects, the conditioning provided by these environments is erratic enough for Heath (1987) to consider [ɾ̞] phonemically distinctive. The primary mechanism of phonemicization is levelling across morphological paradigms with variable vocalic environments, as in /ṣɣir/ 'small (sg.)' versus /ṣɣaɾ̞/ 'small (pl.)', where [ɾ̞] would be expected in the plural but not in the singular. In this particular case, [ɾ̞] is reported to be generalized across both forms in central urban dialects, but not in the Saharan region.

The case of the labials /b f m/ is more phonetically complex, and less clearly involves phonemicization. While secondary emphatics [b̪ f̪ m̪] arising from emphasis spread do exist, there are also idiosyncratic pronunciations of these consonants in specific words with a secondary articulation variously described as velarization, labio-velarization, or emphasis (Zeroual et al. 2011), and which I will transcribe here as velarization. These words typically exhibit surface gemination of the labials, and often contain underlying clusters with following /w/, as in [mˠmˠagən], the plural of [magana], which uses the morphological template /C_1waC_2əC_3/. In other cases it is much harder to make the case for an adjacent /w/, although some labiovelar feature is influencing the pronunciation. A common example is [mˠmˠi], 'my mother', derived from /ʊmm/ 'mother' and the first-person singular clitic pronoun /-i/. In addition, some clusters of labials with /w/ persist in CMA, so this is not a straightforward conditioned allophone. Heath (1987: p. 225 ff.) attempts to associate these labial variants in his analysis with a process generating labialized velars in CMA from a floating labialization feature originally associated with short /u/, and a similar line of thought leads Harrell (1962) to call them 'labialized labials.'

1.2 Phonetics and Phonology in Sound Change

The issue of phonemicization, which I am considering here for the marginal emphatics, has a central and controversial place in phonological theory. While the existence of discrete abstract categories mapping onto sounds (in spoken language) has been a basic tenet of linguistic analysis since at least Saussure, the precise nature of these units has been the subject of much debate, particularly with respect to the role of phonetics in their representation and their ontological status when undergoing change. The phonological feature model which has been dominant since at least Jakobson and Halle (1965) assumes that phonemes are specified by clusters of features grounded in phonetic perceptual and articulatory cues, but opinions differ widely as to how abstract or concrete these features

need be. While authors such as Hale and Reiss (2008) believe that features consist of abstract divisions in a fixed featural space, Mielke (2008) considers features to 'emerge' from phonological class groupings in a language, and other frameworks, such as articulatory phonology (Brownman and Goldstein 1986) and the work of Ohala (1990), interpret features more as reflections of concrete phonetic information such as articulator gestures. A more radical view, typified by Evolutionary Phonology (Blevins 2006), treats dimensions of phonological difference as artifacts of historical change with a dubious synchronic representational status.

These divergent views respond to different, apparently conflicting, aspects of phonological systems. On the one hand, Saussurian oppositional contrasts in the form of minimal pairs and non-overlapping distributions form the basis for defining and differentiating phonological categories, both at the phonemic and the featural (natural class) level. On the other, these categories correspond closely to measurable physiological and acoustic properties of speech, and undergo processes which are often explicable by very concrete, not to say extralinguistic, means. A typical example is palatalization of velars before front vowels, in which the consonant's place of tongue constriction is moving nearer to the tongue position of the adjacent vowel. Yet this process, evidenced in early Romance, clearly became detached from vocalic coarticulatory effects somewhere along the way to modern French, in which the reflex of **k* before a (high) front vowel is an apical sibilant, but remains [k] in other contexts. At some point, the difference between palatalized-*k* and non-palatalized-*k* became phonological rather than phonetic, allowing the two sounds to have different histories. Emphasis spread in Arabic is a similar case in which a coarticulatory phonetic process has become a phonological rule subject to structural, non-phonetic constraints, yet still retains its phonetic transparency (Shahin 2002, Bin-Muqbil 2006).

In the palatalization example, the shift described was from a purely phonetic effect to an allophonic rule (at least initially). New phonemic categories can also emerge in languages, primarily through the process known as 'secondary split', whereby an interaction of processes results in phonological opacity obscuring the distributional relationship between two variants (Hoenigswald 1960). A secondary split could be said to have occurred in CMA between [ɾ] and [r], since the merger of short /a/ and /i/ to schwa or null in non-word-final position resulted in a removal of the conditioning environment for one or the other variant.

On a grand historical scale, changes like these seem relatively straightforward, but identifying phonological rules and phonemic categories in the process of emergence is a different matter. As testified by the entire body of work in sociolinguistics, changes in progress are typified by broad variation at both the individual and community level, with different social groups and interactional contexts favoring different variants (Labov 2001). Also, changes can operate through lexical diffusion, such that words come to be learned with different variants independent of phonetic context.

A famous example is the so-called 'tense short *a*' variant of /æ/ in the Philadelphia dialect of English, which, in addition to being conditioned by a miscellany of fricatives and nasals, appears in the words *mad*, *bad*, and *glad* to the exclusion of other words with near-

identical phonetic contexts, such as *sad*. This could be considered sufficient evidence to consider /æ:/ and /æ/ separate phonemes, and most authors do indeed approach the sounds as separate categories (Labov et al. 2013, Bermúdez-Otero 2007). Even so, a large degree of regularity and lack of overlap is preserved in the patterning of the two segments, and an allophonic system tends to re-emerge in the speech of speakers with the so-called 'nasal' system, which has the tense variant before /n/ and /m/ and the lax variant elsewhere. This variant is less stigmatized, and exists in variation with the traditional Philadelphia system.

Another well-studied instance of a phonemicization puzzle in English is the phenomenon known as 'Canadian raising', in which the diphthongs /ai/ and /au/ exhibit a raised nucleus before tautosyllabic voiceless obstruents. The process becomes opaque when a following coronal stop /t / or /d/ neutralizes to [ɾ] before an unstressed vowel, yielding pairs like [rʌiɾər] 'writer' and [raiɾər] 'rider'. As long as surface morphological alternations with more transparent forms exist, as for example between 'writer' and [rʌit] 'write', the process can be claimed to remain allophonic, but the situation is indubitably ambiguous, and lends itself almost equally well to an analysis where raised /ʌi/ is the underlying representation of the vowel in 'write' (Joos 1942). In fact, the raised variant has been observed to occur in some single-morpheme words with a flap historically derived from /d/, as in [spʌiɾər] *spider* and [snʌiɾər] *Snyder*, suggesting that it is spreading as a separate category via lexical diffusion. However, as Idsardi (2007) notes, these cases could be handled by positing underlying /t/ in their representations.

Lexical diffusion, sociolinguistic variation, and the additional problem of phonetic gradation in the course of a change are all elements of what Labov, Weinreich and Herzog (1968) called the 'incrementation problem' : how linguistic changes, once initiated, spread through the grammar and the community on their way to completion as stable elements of a language. An influential approach to understanding incrementation, particularly in the syntactic domain, has been the 'competing grammars' model of Kroch (1989), whereby an initial ambiguous input is variably acquired with different grammatical parameters. A central tenet of the competing grammars model is the Constant Rate Hypothesis, which states that a unitary syntactic change progresses at the same rate in all contexts for which it occurs, although some contexts may favor the change more than others (Kroch 1989: 200). Fruehwald (2013) adapted the constant rate approach to evaluate phonologization of phonetic variables in the Philadelphia Neighborhood Corpus, predicting that effects resulting from co-articulation would exhibit stable rates, though at different levels across conditions, whereas structural phonological effects would exhibit divergent rates of change. Applied to Canadian Raising, this approach succeeded in establishing that the process was allophonic rather than phonetic from an early date (p. 130).

Returning now to the issue of phonetics in phonology, we can see that phonological theory must be able to account for the difference between phonetic effects, allophonic rules, and phonemic contrasts in order to adequately model language variation and change. In the approach outlined by Bermúdez-Otero (2007), a 'modular feedforward model' of phonology

separates lexical, phonological, and phonetic representations into separate modules, which are related by both phonological rules mediating lexical and phonological representations, and phonetic rules mediating phonological and phonetic representations. Phonologization primarily involves codification of a purely physical effect as a phonetic rule, which may in turn come to alter the phonological representation of a word. This latter stage 'is implicated in the rise of so-called 'quasi-phonemes' that precedes secondary split' (p. 504), as in the case of Moroccan [ṛ] discussed here and the case of tense short-*a*, which exhibits so-called 'marginal contrast' (p. 511). The unsystematic recategorization of lexemes is a third stage of 'phonological' change, as is the fossilization of rules into morphological contrasts. Unless a change has gone to completion, some amount of variation is always present: 'The omnipresence of variation during change in progress is one of the reasons why quantitative techniques are indispensable in research into the problem of implementation' (p. 499).

The approach to phonological features discussed in section 2.4 below assumes a relatively concrete mapping of features onto phonetic correlates, while maintaining the modular distinction between phonology and phonetics and the abstract character of the phoneme. A further aspect of Sylak-Glassman's (2014) model which is of interest from the perspective of sound change is the proposal of 'phonetic subfeatures', described in terms of articulatory gestures, which can be used to predict some phonological behavior but are not contrastive. For example, the subfeature 'double bunching epilaryngeal configuration' is derived from the behavior of pharyngealized vowels in the Caucasian language Lak, which also exhibit palatalization and spread both secondary articulations to adjacent consonants. Velarization in Moroccan Arabic may be found to have some similar status as an inchoative low-level feature, if it is found not to pattern phonologically as emphasis.

2. Marginal Phonemes, Allophony, and Place Features

With respect to phonological change and ambiguity, a number of significant questions may be raised concerning the two sets of Moroccan Arabic segments under consideration. For the emphatic liquids, which include both /ṛ/ and a variant of /l/ attested sporadically in similar environments, the main issue at stake is the degree of regularity in their distribution and how this affects their perception as distinct categories in the grammar. For the velarized labials, a further important question is whether the articulatory phonetic differences separating them from secondary emphatics arising from emphasis spread are sufficient to keep the two categories phonologically distinct. A final set of questions revolves around what the role of guttural consonants, particularly uvulars, in conditioning these sounds can tell us about the organization of phonological place in the lower vocal tract.

2.1 Ambiguous Distributions and Phonemicization : The Case of [ṛ]

The question of whether secondary emphatics, including [ṛ] and [l̪], are separate phonemes has been asked for a number of dialects, with highly variable and often inconclusive results. For instance, in Watson's description of Cairene and San'ani phonology (2007), she describes [ṛ] in Cairene as a segment 'whose distribution is sometimes determined by phonological context and for which several (near-) minimal pairs with plain /r/ are attested' (p. 16). As for [l̪], she claims that for many dialects it 'is found exclusively in *allah* 'God' and derivatives.' While Watson chooses to write these sounds between slashes, it is not clear from such a description that they merit consideration as phonemes, and she ultimately classes them, together with [ḅ] and [ṃ], as 'marginal phonemes,' often attested only 'among certain speakers of the dialect' (p. 21). Such a characterization of the situation requires more nuanced characterization from a structural perspective. Does a 'marginal phoneme' only have a vaguely different mental representation from similar segments in the language? Or is it marginal in the sense that it idiosyncratically phonemic or allophonic depending on the speaker, or varies by context?

These questions are never really answered in most descriptions of marginal emphatics, even by writers such as Younes (1994) who address their patterning in detail. The difficulty, of course, arises from the ambiguity of distributional evidence, and the paucity of alternative sources of evidence for the distinct categorization of these segments. As described by Younes for northern Palestinian Arabic, [ṛ] and [r] would seem to be allophones of a single phoneme which is underlyingly emphatic in northern Palestinian Arabic, but he notes that other authors interpret them as allophones of an underlyingly non-emphatic segment (Al-Mozainy 1981, for Hijazi), as separate phonemes (Broselow 1976, for Cairene), or as allophones undergoing phonemic split in the case of Heath's (1987) analysis of Moroccan Arabic. As support for his own view, Younes cites participation of /r/ in emphasis spread and association with emphatic vowel allophones, as well as its association with emphatics and dorsals in a verbal class with theme vowel /u/, except when it is 'de-emphaticized' to [r] before /i/ or a coronal or dorsal consonant.

In Moroccan Arabic, Heath's analysis of /r/ proposes that phonemic split is occurring by means of levelling across morphological paradigms, but he also claims that different areas of the country exhibit different levels of productive alternation. In his 2002 dialectological study, he divides Moroccan dialects into three main types, northern (pre-Hilalian), central (koiné), and Saharan. In the central koiné, phonemicization of [ṛ] is quite advanced: "either plain *r* or pharyngealized *ṛ* generalizes to most or all ablaut forms of a given stem" (p. 9). The northern sedentary dialects also exhibit a high degree of levelling, but in the southern, Saharan dialects "a respectable number of *r* ~ *ṛ* alternations are preserved in ablaut derivation, even when the original vocalic basis for the allophony has become opaque" (p. 7). A small number of alternations, however, are preserved even in the supposedly phonemicizing dialects, such as /ħmaṛ/ 'donkey' versus /ħmir/ 'donkeys' and /kbir/ 'big' with plural /kbaṛ/.

As these examples show, differences in vocalic environment underlie the productive alternations that exist in CMA, such that plain [r] occurs adjacent to /i/ and emphatic [ṛ] occurs adjacent to /a/ or /u/. An adjacent schwa may condition [r] if it is historically derived from short /i/, as in the Saharan example /ʃaːrəb/ < /ʃaːrib/ 'drinking' (Heath 2002, p. 7). Consonantal influences do not produce any allophonic alternations, because the set of consonants making up a stem is stable across any given paradigm. The only consonantal effect noted by Heath is a tendency for neighboring uvulars /q ɣ χ/ to favor [ṛ], 'but this factor is not always decisive' (2002, p. 151). He cites /ṛqba/ 'nape' and /qḍəṛ/ 'be able to' as cases in which /q/ favors [ṛ], but a plain variant /qdər/ is dominant in the eastern part of Morocco. For stems with /χ/, he cites /mnχəṛ/ 'nostril' and /lχχəṛ/ 'last', both typical of northern Morocco with plain variants around Marrakech and in rural areas farther north (p. 153). The only example with /ɣ/ is /ɣaṛ/ 'cave', which often exhibits generalization of [ṛ] to the plural /ɣiṛan/ despite the presence of a high vowel. Heath considers this to be primarily due to the overall trend towards leveling of alternations, but believes that "it got some additional phonetic support from the consonantism" (p. 155). Interestingly, pharyngeals /ħ ʕ/ did not favor [ṛ] in the same way, and there is no indication of a coronal depharyngealization effect as described by Younes for Palestinian Arabic.

For [ḷ], Heath records only a handful of words in CMA, including the ubiquitous use of the sound in /ḷḷah/ 'God'. The others are /tħḷḷa/ 'take care', a verb used in a formulaic greeting, sporadic occurrences near Marrakech in /ltlata/ 'Tuesday', and words such as /gəlb/ 'heart' where the /l/ is adjacent to a /g/ derived from /q/ (2002, p. 157). If, as some evidence suggests, the /g/ is taken to be emphatic, this could be understood as pharyngealization spread. Either way, it is hard to use this data to draw any conclusions about CMA /ḷ/, except that it is indeed marginal and occasionally appears without a clear conditioning environment. In Gulf and Iraqi dialects of Arabic, [ḷ] often appears as an allophone conditioned along lines similar to [ṛ], and the CMA pattern may simply be the last vestiges of such a system. For /ḷḷah/ and /tħḷḷa/, one might also suggest that the forms are cultural borrowings from an artificial formal register (Classical or Standard Arabic) and do not form part of the core CMA grammatical system.

Taking this data into consideration, it is easy to appreciate the position of many researchers that /ṛ/ and /r/ are almost, but not quite, separate phonemes. Within the same dialect, it is easy to find both allophonic alternations and overlapping distributions, creating what Heath calls a 'structural tension' (1987, p. 298) that has yet to be resolved. Surface ambiguity may indeed characterize the language at the community level, yet within the context of variation, it is likely that individual speakers organize their mental grammars in a more principled way. I intend to investigate the hypothesis that individual speakers of CMA will choose either an allophonic or phonemic representation of emphatic [ṛ] based on the patterns of distribution in their own language. My expectation is that this choice will be reflected in the degree of participation of [ṛ] in emphasis spread, such that phonemic /ṛ/ will serve as a locus for the long-range spread of emphasis to other segments in the same way as

/ṭ ḍ ṣ ẓ/, whereas allophonic [ṛ] will exhibit weaker, shorter-range effects attributable to phonetic interaction with adjacent segments.

2.2 Phonetic Aspects of Secondary Place : The Case of Velarized Labials

Although these same considerations apply to the labials [bˠ fˠ mˠ], one must also account for the fact that these sounds are less phonetically identifiable as emphatics than [ṛ], and might be expected to pattern differently from emphatic consonants even if they were phonemicized. Not even [ṛ] is phonetically identical to the coronal emphatics, according to articulatory phonetic studies (Younes 1994). Its secondary articulation is characterized by a less extreme retraction of a larger part of the tongue dorsum, including uvular and/or velar constriction as well as constriction in the upper pharynx. Ghazeli (1977) even suggested that in Tunisian Arabic, [ṛ] is retroflexed rather than pharyngealized. Despite these differences, the evidence overwhelmingly suggests that [ṛ] exhibits pharyngealization, and it is typically described as an emphatic. The same cannot be said for the velarized labials, especially since they coexist with secondary emphatic allophones [b̪ f̪ m̪] in the grammar.

The velarized labials are tentatively labeled emphatic-like in most descriptions of Moroccan Arabic, including Harrell (1962) and Heath (1987). Heath sidesteps the question of their status somewhat by naming their defining property 'Special Labial Pronunciation' (SLP). He claims that SLP consonants are typically geminates, that 'some velarization or pharyngealization is almost always present; and in the geminate cases there is often a faintly labialized release' (p. 225). Vowel effects are found to be variable, with /a/ ranging from quite fronted to quite backed, /i/ ranging from lowered to a centralized diphthong with an effect "similar to that of Russian *y* in *ty*", and /u/ remaining unaffected. While SLP clusters "rarely occur before short Vs", they would appear to modify unrounded schwa to [ʊ] as a result of their labialization (p. 226). From my own impressionistic experience, I can confirm that the Russian-like centralized /i/ associated with these sounds is their most salient and unusual characteristic to a native English speaker.

In some respects, this description diverges enormously from the vowel effects associated with primary emphasis: backed /a/, lowered /i/, and lowered /u/. However, there is a large degree of overlap, particularly for /a/ and /i/, and it is not clear from Heath's description whether we are dealing with inter- or intra-speaker variation. It may be that for some speakers, the velarized labials are structurally emphatic allophones, and have similar effects on the adjacent vowels, while for others this is not the case. An excellent example of the possible relationship to pharyngealization is the word [ṛ bˠbˠi] 'my Lord' < /rabbi:/, cited by Heath (1987, p. 227). There is no plausible source here for a (labio-)velarization feature, so it is likely that emphasis spread from the neighboring emphatic [ṛ] (originally conditioned by the historical short /a/) and was reinterpreted as 'SLP' due to the association of that feature with geminates and with similar words like [mˠmˠi] 'my mother'. In most words exhibiting velarized labials, however, a geminate velarized labial can be traced back to either a cluster with following /w/, or a nearby historical short /u/. In almost all

cases, there are productive morphological alternations with forms exhibiting no velarization, as in [fmˠmˠək] 'your mouth' versus [fʊmm] 'mouth'.

The articulatory phonetics of velarized labials have been studied in some detail for the eastern Moroccan city of Taza by Zeroual et al. (2011). This EMA and ultrasound study determined that "MA labialised labials are produced with labial-velarisation", whereas "emphatics /ṭ, ḍ/ are pharyngealised and not velarised, and /ḍ/ has a slight degree of labialisation" (p. 295). The authors further considered the phonetic quality of /a/ following (labio-)velarized labials to be velarized [aˠ] as opposed to pharyngealized [ɑ] after /ḍ/, on the basis of F2 lowering which was not combined with F1 raising. Considering, however, that F1 raising has not consistently been associated with emphatic vowel variants across Arabic dialects whereas F2 lowering has been (Bin-Muqbil 2006), this argumentation must be treated with circumspection. What is clear from the study is that the phonetics of the two articulations are objectively different, with emphasis located in the pharyngeal cavity and 'SLP' located in the oral cavity. This concretization of the phenomenon allows us to frame the problem of 'SLP' as the question of whether pharyngealization and velarization have different featural specifications in the phonology, a question which can be addressed both distributionally and by examining whether velarized labials ever induce long-range spread of their secondary articulation to other sounds, in a manner similar to emphasis spread. My hypothesis is that the distributional differences between velarized and pharyngealized labial allophones will maintain the categories as distinct, and that adjacent vowel effects of velarization are mostly, if not entirely, phonetic in nature. However, it is also possible that the gemination associated with velarized labials has resulted in the construal of velarization as the variant of emphasis associated with labial geminates, as suggested by the word [ṛbˠbˠi] above.

2.3 Phonological Place in the Lower Vocal Tract : The Question of Uvularization

The issue of whether velarization and pharyngealization are phonologically distinguished, together with the effect of uvulars favoring emphatic [ṛ], relates to a further set of issues surrounding the organization of place features in the lower vocal tract. Although pharyngeal and uvular articulations have been historically underrepresented in the literature due to their typological rarity, there is an important thread of literature accounting for their featural representation. The most comprehensive recent survey and analysis of lower vocal tract phonology is Sylak-Glassman (2014), and issues of featural organization in Arabic are discussed in detail by Bin-Muqbil (2006) and Youssef (2013).

The guttural feature geometry of McCarthy (1994) is one of the most influential modern analyses of pharyngeal place. On the basis of primarily Semitic data, McCarthy argues for the existence of a natural class of 'gutturals' comprising uvular and pharyngeal consonants, which are characterized by the place feature [pharyngeal]. Since uvulars share some properties with velars, they are specified by a double place specification of [dorsal] and [pharyngeal]. Emphatics, as might be expected, are specified as both [coronal] and

[pharyngeal], but also have a third place specification as [dorsal], since according to Ghazeli's (1977) X-ray tracings they appear to be more uvularized than pharyngealized. All of these features attach directly to the place node, with no hierarchy of primary versus secondary place. Bessell (1992) uses primarily data from Salishan languages to derive a similar system, but with [tongue root] in place of [pharyngeal] and no [dorsal] specification for emphatics.

While the systems of McCarthy and Bessell account for the existence of the natural class of gutturals and successfully model processes such as Arabic emphasis spread as feature spreading of the radical/pharyngeal feature, they are not quite descriptively adequate. Two major problems relevant to Arabic are the association of pharyngeals and emphatics with the same place feature, and the specification of uvulars as doubly articulated 'dorso-pharyngeals'. A more general typological problem addressed by Sylak-Glassman (2014) is the inability of these systems to account for pharyngealized uvulars or uvularized pharyngeals, both of which are attested in Salishan and Caucasian languages.

As mentioned earlier, emphatics and pharyngeals have markedly different phonetic and phonological effects in Arabic. Pharyngeals, for instance, characteristically raise the first formant of an adjacent vowel, whereas emphatics raise the second formant. While first-formant effects are sporadically claimed for emphatics, the backing effect is never observed for pharyngeals (Bin-Muqbil 2006). This acoustic observation is backed up by a body of recent instrumental work, such as Moisik (2013) and Esling (1996), which demonstrates that primary pharyngeals /ħ ʕ/ are primarily articulated by structures in the lower pharynx such as the epiglottis and arytenoid cartilages rather than the tongue root, while secondary pharyngeals are articulated in the upper pharynx by the tongue root and the pharyngeal wall. McCarthy's assignation of both [dorsal] and [pharyngeal] to emphatics, while capturing the notion that secondary pharyngealization is higher up, fails to account for this fundamental difference in the articulatory gesture from primary pharyngeals. A similar argument holds for the uvulars, although at least the fricatives /χ/ and /ɣ/ sometimes pattern phonetically with pharyngeals. An additional problem with the treatment of uvulars is that there is no evidence that they have a phonetically complex articulation in the sense of [k͡p̠]. McCarthy does recognize that there are articulatory differences between primary and secondary pharyngealization, but he argues that these are phonologically irrelevant due to the lack of sensorimotor precision in the pharyngeal region (1994: p. 201), a claim which is no longer admissible in light of recent phonetic work. To account for the phonological differences, he suggests that /χ ɣ ħ ʕ/ may all actually be approximants (p. 222), a view which is also adopted by Bin-Muqbil (2006).

An alternate approach, first taken by Czaykowska-Higgins (1987), postulates separate place features for the upper and lower pharynx. In the original conception, these are binary features dominated by a 'tongue root' node, but in the more recent proposal of Sylak-Glassman (2014), based on Esling (2005), the asymmetry in active articulator is taken into account, and a feature system proposed based on lingual and epiglottal gestures. Uvular and upper pharyngeal constriction is characterized by the feature [±retracted], and lower pharyngeal or epiglottal constriction is characterized by the feature [±constricted epiglottis]

([±ce]). The [+retracted] feature (similar to [RTR], but without restrictive reference to the tongue root) also characterizes low and low-mid back vowels /ɑ ɔ/ which involve retraction of the tongue body by the hyoglossus muscle, whereas upper back vowels involving 'movement of the tongue by the styloglossus upward and backward' share a feature [+raised] with dorsal consonants (Sylak-Glassman 2014, p. 137). A final distinctive feature is [±open], which correlates with jaw lowering and characterizes both pharyngeal and epiglottal consonants and low and low-mid vowels. Low front vowels are [+open] and [-retracted].

Under the analysis that emphatics are [+retracted] but not distinctively [+ce], pharyngeals are characterized by [+ce] and [+open] but not [+retracted], and uvulars are [+raised] and [+retracted], the vowel effects fall out naturally from feature spreading to vowels in this system. Pharyngeals should cause vowel lowering but not backing, emphatics should cause vowel backing, and only some lowering, and uvulars should cause only backing. A possible criticism is that this system does not account for the patterning of /χ ɣ/ with pharyngeals, but this asymmetry between uvular fricatives and stops is not accounted for by the McCarthy-Bessell system either.

In both the Sylak-Glassman and McCarthy systems described above, the upper pharyngeal constriction of emphatics is theoretically indistinguishable from uvular constriction, and its differentiation in the Bessell system results in an unhelpful conflation with primary pharyngeals. If emphasis is phonologically just uvularization, it would make sense for primary uvulars to exhibit properties similar or identical to emphatics, such as vowel backing and long-range conditioning of secondary emphatics. Heath (1987) does observe that uvulars trigger a sort of emphasis spread, with the odd caveat that 'roughly, a uvular counts as half a [+PH] value in its allophonic influence' (p. 306). This is meant to account for an intermediate vowel backing effect, and for the fact that 'in the immediate vicinity of a uvular, /ṛ/ predominates but there are some cases of plain /r/' (p. 307). Examples of the former include words like /χ̟rif/ 'autumn' in which the vocalic environment disfavors [r̠], and the latter include /qrd/ 'monkey' and /rɣəb/ 'request'. Intriguingly, Heath observes that when /r/ is between two uvulars in the same word as in /ɣṛəq/ 'drown', the non-emphatic variant is never attested.

Given this information, when [r̠] occurs in the same root as a uvular, it could possibly could be interpreted as a secondary emphatic allophone resulting from emphasis spread, rather than a separate phoneme. In terms of acquisition, inputs like [r̠qba] from a speaker for whom the uvular stop is structurally emphatic and [r̠] is an allophone resulting from emphasis spread would be interpreted as evidence for phonemic /ṛ/ by a language learner for whom /q/ has not been acquired as emphatic. The data as reported is ambiguous, and this would be a useful dimension of individual variation to consider.

A further implication of Sylak-Glassman's system with respect to [bˠ fˠ mˠ] is the lack of distinction between velarization and uvularization in his model. Both secondary articulations are treated as [+raised], since the feature distinguishing velars from uvulars, [±open], is not lingual and its articulatory correlate of jaw lowering is not found in pharyngealized anterior obstruents. The author does not consider this to be a disadvantage,

since "no language phonologically distinguishes velarization from uvularization" (Sylak-Glassman 2014, p. 138). Under this view, [bˠ] and [ḃ] would be phonologically identical 'dorsalized' labials, and SLP should not be distinct from emphasis except from a strictly phonetic perspective. If the velarized labial consonants are in fact found to behave like emphatics with respect to emphasis spread, this would make a strong case for Sylak-Glassman's view.

2.4 Phonetic Evidence and Research Questions

Putting this all together, we see that the marginal emphatic consonants in Moroccan Arabic raise several questions regarding phonemicization, phonologization, and the phonological organization of the lower vocal tract. First, the ambiguous distribution of *ṛ* raises the question of whether it is a phonemic or allophonic category with respect to plain *r*, and if it is not clearly either, how it is expected to behave as a so-called 'marginal phoneme'. Second, while the velarized labials do appear to be allophonic, it remains to be seen whether their phonological representation is distinct from that of secondary emphasis, and whether their effect on adjacent vowels is structural or merely co-articulatory in nature. Lastly, it must be determined if uvulars are in the same phonological class as coronal emphatics, since this would suggest that occurrences of *ṛ* and *ḷ* in their environment is merely conditioned emphasis spread.

Key to effectively answering these questions is moving beyond distributional data to consider the behavior of these ambiguous segments as compared to primary emphatics. The process of emphasis spread provides an ideal avenue for such an investigation, since it is a phonological process specifically targeting emphatics which lends itself well to acoustic phonetic measurement. Strong correlation between a lowered second formant and emphasis, with a less consistent effect of raised first formant associated mainly with high vowels, is upheld by a number of phonetic studies (Bin-Muqbil 2006, Zeroual et al. 2011, Shoul 2009, Ghazeli 1977), allowing us to observe differences in both the intensity and scope of vocalic emphasis spread. This in turn can help to distinguish between phonetic and phonological effects, as F2 differences which are quantitatively smaller or which are restricted to adjacent segments are differentiated from categorical long-range effects. By analyzing vowel variation conditioned by nearby emphatic or velarized consonants, we can determine whether this influence is a local effect explicable by co-articulation or a long-range effect attributable to feature spread. This analysis can then be assessed with respect to distributional patterns for a given speaker in order to evaluate the phonological status of emphatic liquids and velarized labials.

Contextualizing the problem of marginal emphatics in terms of the interaction between a concrete phonetic effect and lexical distributions allows the formulation of concrete, testable hypotheses concerning their behavior. Based on previous literature, I predict (1) that individuals with a more regular distribution of [ṛ] will exhibit an attenuated vowel-backing effect attributable to co-articulation, whereas for individuals with phonemic

/ṛ/, the segment will be a locus for phonological emphasis spread indistinguishable from that associated with /ṭ ḍ ṣ ẓ/; (2) that despite slight phonetic differences, [bˠ fˠ mˠ] will be found to exhibit similar vowel backing effects to /ṭ ḍ ṣ ẓ/, except for speakers for which all occurrences of SLP are attributable to phonologically transparent /w/; and (3) that speakers with allophonic [r̠] may exhibit more typical emphasis spread in words containing both [r] and a uvular.

3. Research Methods and Prospectus

In order to appropriately test these hypotheses, it will be necessary for me to collect thorough and individualized phonetic datasets from Moroccan Arabic native speakers. A variety of factors will need to be taken into account to ensure the quality of the data, including providing comprehensive, well-structured word lists, controlling for stylistic context in conversational speech, and recruiting balanced and age-stratified subject pools within a community. While some interviews may be conducted in the U.S., a period of travel to Morocco will be necessary to acquire the controlled community-level data that is needed to understand variation.

In preparation for this proposal, a preliminary interview was conducted with a Moroccan Arabic speaker living in Philadelphia, eliciting a short list of words containing the target segments. In this remainder of the proposal, I will first present the results arising from that interview, before describing an improved methodology for future data collection addressing the factors mentioned above. Finally, I will outline my plan for scheduling fieldwork and dissertation completion.

3.1 Analysis of Preliminary Data

My preliminary interview was conducted with a Moroccan contact affiliated with the University of Pennsylvania. The speaker lived in multiple parts of Morocco before moving to the United States, but is native to the area of Settat, a dialect which is reported by Aguadé (2013) to have phonemic 'velarization' of [r̠]. The interview was conducted in English and consisted primarily of elicitation of individual CMA words reported to contain marginal emphatics, as well as some full sentences and control vocabulary.

In total, the interview provided 111 tokens of 79 distinct forms, with 43 forms including a rhotic and 11 forms including velarized labials. Only 2 word forms in the interview contained emphatic coronal obstruents. 10 of the remaining forms contained plain labials with no emphatics, uvulars, or rhotics in the same word, so a comparison was possible between these and the velarized labials. Among the words with /r/, 14 have an adjacent /a/, 13 have an adjacent /ə/, 9 have an adjacent /i/, and 5 have an adjacent /u/, while 3 words exhibit word-initial syllabic /r/ with no adjacent vowel. With respect to non-adjacent vowels in the same stem as /r/ or a velarized labial, not enough words were present

in the data for comparison. A number of tokens exhibited word-final /a/, but since *a* in this position is regularly backed to [ɑ] by a phonological rule unrelated to emphatics, it is not appropriate for emphasis spread analysis. Emphatic /l/ was also not considered in the interview since at most one token of the segment was observed (in the word /tħlḷa/).

The interview was prepared for analysis using the phonetic analysis program Praat (Boersma & Weenink 2015). All Moroccan Arabic speech in the interview was transcribed and segmented, with segmentation of vowels and sonorants restricted to intervals exhibiting clear formant structure. Praat's automatic tracker was used to take formant measurements, set to find 6 formants below 6000 Hz using the default window length of 25 ms. Formant measurements with associated segment and word transcriptions were then imported into the statistical environment R (R Core Team 2013) and coded manually for phonological environment.

3.1.1 Vowels Adjacent to /r/

With respect to [r] and [ṛ], the preliminary data confirms the existence of an originally allophonic distribution which has acquired a phonemic character through paradigmatic generalization. For /a/, the best-attested vowel adjacent to /r/ in the dataset, the data generally conform to a pattern whereby an /a/ adjacent to an /r/ which is also adjacent to an /i/ is raised and fronted while other tokens of /a/ adjacent to /r/ are lower and backer, with higher F1 and lower F2. This conforms to the allophonic generalization that an adjacent /i/ conditions plain [r], while /r/ adjacent to /a/ is otherwise emphatic and will thus be followed by the emphatic allophone of /a/. However, there are notable exceptions in the case of *byar*[1], the plural of *bir* 'well', *ɣiran*, the plural of *ɣar* 'cave', and *ḳiran*, the plural of *ḳar* 'bus'. / *byar* clusters with *ʕaris* 'bridegroom' in having a high-front 'plain' /a/, while the /a/ in *ɣiran* and *ḳiran* has a back pronunciation despite the presence of an *i* adjacent to the preceding *r*. This pattern can be explained by generalization of the variant of *r* found in the singular of each noun to the plural form -- plain [r] in the case of *bir*/*byar*, and emphatic [ṛ] in the case of *kar*/*kiran* and *ɣar*/*ɣiran*. This data provides good evidence for a phonemicizing emphatic *r* system, since three singular/plural pairs exhibit generalization of either [ṛ] or [r] across vocalic context.

Figure 1 illustrates this acoustic distribution, with the color of datapoints indicating whether there is an /i/ adjacent to /r/ in the word. The front and back clusters of /a/ vowels form clearly distinct distributions with regard to both first and second formants. The /a/ in *biar* and *ʕaris* has a mean F2 of 1823 Hz, with a standard deviation of 87.6, while the /a/ adjacent to *r* in other words has a mean F2 of 1288 Hz (sd = 108.5). A Welch's two-sample t-test comparison confirms that the F2 distributions are significantly different between the two groups of words ($t(8) = -12.8$, $p < 0.0001$). A significant difference between

[1] Here and elsewhere where italicized transcriptions are used for ease of reading, an Arabist transcription system is used which differs in some respects from the IPA. In addition to subscript dots representing emphasis, note that *e* = /ə/, *š* = /ʃ/, *j* = /ʒ/, *y* = /j/, and *ḥ* = /ħ/.

the groups also exists for F1 (t(8) = 5.43, p = 0.0007), although the means are closer and there is some degree of overlap. For *biar* and *ʕaris* the mean F1 is 612 Hz (sd = 54.6), while for the other words the mean is 746 Hz (sd = 55.8).

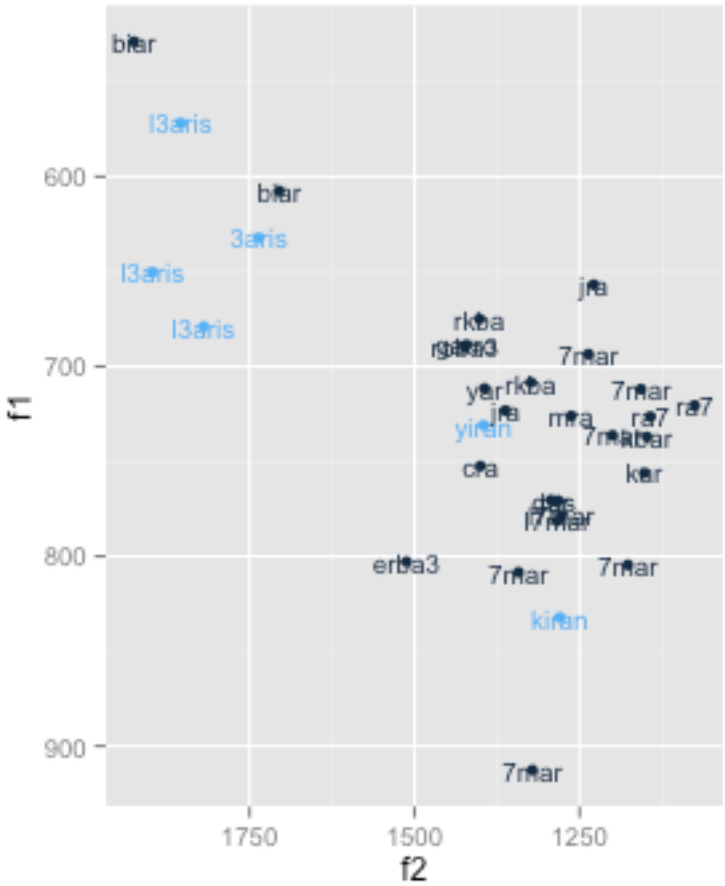

Figure 1: Vowel midpoint measures for /a/ adjacent to /r/. Words colored light blue also have an /i/ adjacent to /r/, and thus would be expected to exhibit fronted a under an allophonic distribution of [ɾ].

The distribution of /i/ tokens adjacent to /r/ was more distributionally homogeneous. Figure 2 shows the distribution of /i/ for vowel midpoint measures. Midpoint measurements for *kiran* and *ɣiran*, which are expected to have backed and lowered /i/ based on their patterning as emphatic with respect to /a/, exhibit higher than average F1 without significantly departing from the distribution of other instances of /i/ adjacent to /r/ in the words *ħmir*, *bir*, *sir*, *ɣir*, *riħa*, *ʕaris*, and *jrit* (F1: t(2) = 1.54, p = 0.28; F2: t(2) = -0.71, p = 0.57). However, vowel endpoint (75%) measures for the /i/ in *kiran* and *ɣiran* do indicate significant lowering of F2 (t(10) = 2.92, p = 0.01), at 1947 Hz and 1876 Hz respectively as compared to an overall mean of 2156 Hz for other words with /i/ adjacent to /r/. This indicates that the emphatic [ɾ̞] in *kiran* and *ɣiran* has only a local coarticulatory effect on preceding /i/, although it spreads its emphasis categorically to following /a/.

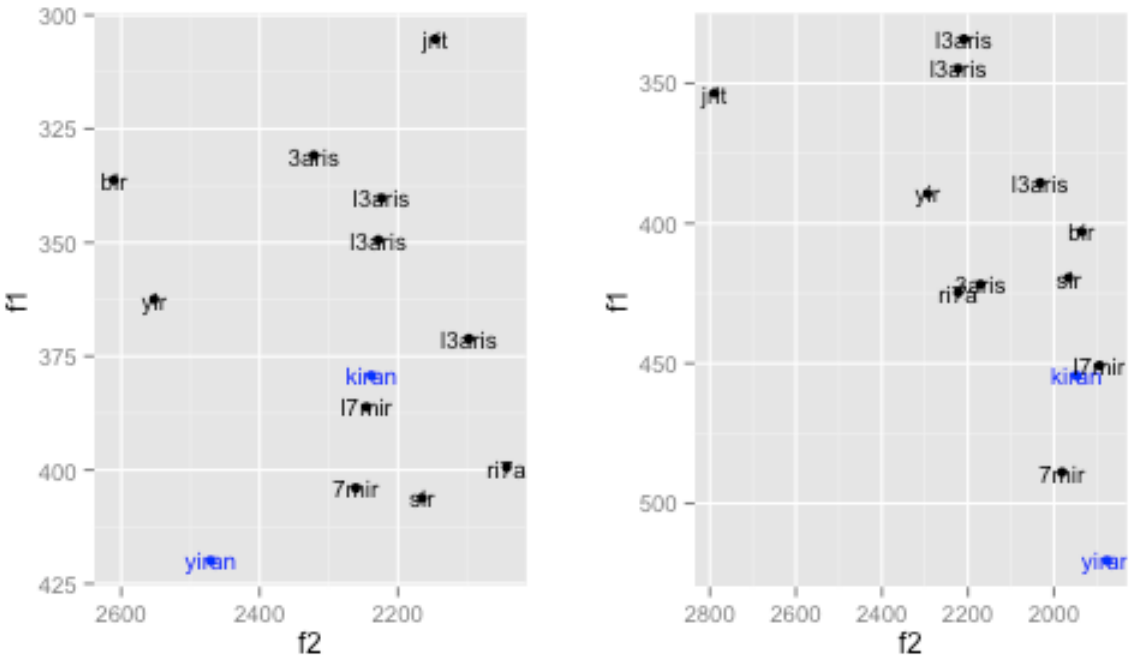

Figure 2: Formant measures at the midpoint (left) and 75% point (right) for /i/ adjacent to /r/.

In the endpoint data for Figure 2, note also that the /i/ in *ḥmir* 'donkey (pl.)' has similar phonetic characteristics to the /i/ in *ḳiran* and *ɣiran*, being characterized by high F1 and low F2. This effect cannot be explained as influence from the primary pharyngeal /ħ/, since the /i/ in *ʕaris* does not behave similarly. This may be taken as evidence that the *r* in *ḥmir* is emphatic [ɾ̪], having generalized from the singular form *ḥmar*, in which the /a/ is firmly within the low-back emphatic [ɑ] range (see Figure 1). If *ḥmir* is included with *ḳiran* and *ɣiran* as 'emphatic' instead of 'plain', the F1 difference between the two word classes is found to be significant ($t(3) = -4.75$, $p = 0.01$) as well as the F2 difference ($t(14) = 4.15$, $p = 0.0008$). This observation strengthens the conclusion that [ɾ̪] is phonemic for this speaker, especially since the productive *ḥmir*/*ḥmaɾ̪* alternation is one which Heath reported to be widespread or even exceptionless in Moroccan dialects.

For the remaining vowels, since there was less bimodal behavior within the *r* tokens, I compared tokens adjacent to *r* to those adjacent to other segments. The data for /u/, shown in Figure 3 below, indicate that /u/ has significantly higher F1 when adjacent to *r* (blue datapoints) than when adjacent to other segments ($t(5)=-3.4$, $p = 0.02$). This accords with the observation that emphatic [ɾ̪] tends to be conditioned by /u/, and with the F1 effect for emphatic /u/ noted by Shoul (2009). However, both attested lexical items in which /u/ was adjacent to *r* also contained a primary pharyngeal consonant, so the first formant effect could also be explained as influence from the pharyngeal.

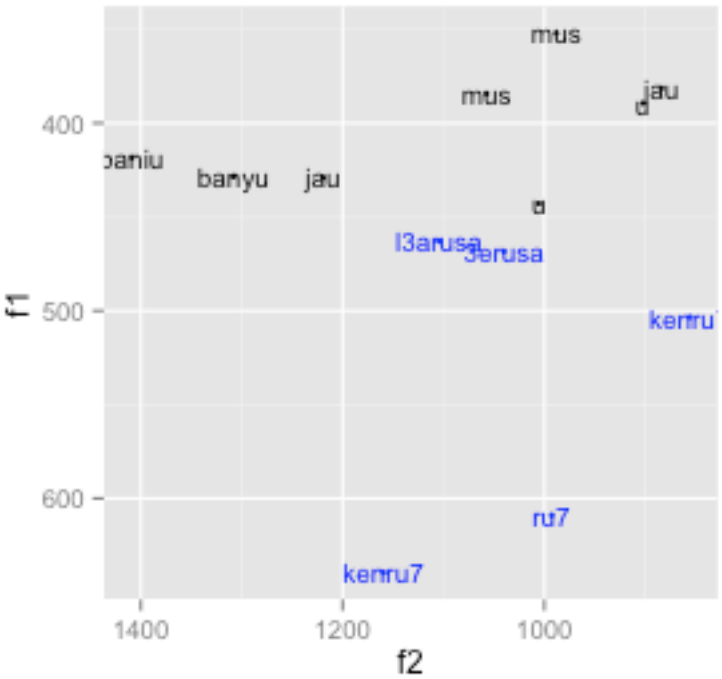

Figure 3. Midpoint formant measurements of all /u/ tokens in the preliminary interview, with tokens adjacent to /r/ colored blue.

For /ə/ adjacent to /r/, shown in Figure 4, the vowel was consistently backed and lowered compared to /ə/ adjacent to labials and plain coronals. Given that all schwas adjacent to /r/ in the dataset with the exception of *kber* 'to grow' < /kabira/ and *ʕers* 'wedding' < /ʕirs/ are derived from SA short /a/, we would expect the phonological environment in these words to favor emphatic [ṛ] and lowering of F2 in the adjacent vowel. As expected, the schwa in *kber* and *ʕers* is further front, suggesting non-emphatic [r] in these words.

rkeb 'to ride' < /rakaba/ also has a front and high /ə/ compared to other /r/ words. The behavior of /rkəb/ admits of multiple explanations, which the formant measurements of /i/ adjacent to /r/ and of /r/ itself (see section 3.1.2 below) can help to distinguish between and evaluate. First, it is possible that [ṛ] cannot spread emphasis rightward to a non-adjacent vowel even if the vowel is tautosyllabic. It could also be the case that the /r/, originally emphatic and conditioned by following **a*, has undergone a lexically idiosyncratic shift to plain [r] in this word. Finally, the /r/ in /rkəb/ could be exempt from allophonic emphaticization, since although it was historically followed by **a*, it is not adjacent to any surface vowel in CMA.

The behavior of /i/ preceding /r/ provides some indirect evidence in favor of the first hypothesis, since any lowering and backing of /i/ in *kiṛan* and *ɣiṛan* occurs only at the endpoint of the vowel and can be explained as a phonetic effect. This suggests that an attenuated pattern of emphasis spread is associated with [ṛ], with leftward spreading to adjacent vowels being partial or absent. The consonantal formant data, however, indicate that the /r/ in *rkeb* has similar phonetic properties to /r/ adjacent to both /i/ and /ə/ < **a*,

indicating that either the second or third hypothesis is likely to be correct. Since the /ɾ/ in [ɾ̪ bˠbˠi], which is also historically adjacent to an **a* which has deleted, has a low F2 of 1412 Hz falling within the emphatic range, the second hypothesis of lexical shift is most likely to be correct.

Schwa is also the only vowel in the data for which comparison with coronal emphatics /ṭ, ṣ/ was possible, in the words *weṣṭ* 'middle' and *meṣṣ* 'small knife'. The /ə/ in these words exhibited a low F2 similar to that in words containing *q* or *r*, but dissimilar from the /ə/ in words containing only plain labial, coronal, or velar consonants. The patterning of the uvular stop in *qelb* 'heart' and *deqq* 'taste' with emphatic [ɾ̪] and pharyngeals provides evidence that /q/ is phonologically similar to emphatics, since it effects adjacent vowels similarly.

The *r* in *merfeq* 'elbow', the only word containing both /ɾ/ and /q/ in the dataset, is expected to be emphatic regardless of its consonantal context. However, it is worth noting that the vowel in this word exhibited the lowest F2 of any /ə/ adjacent to /ɾ/, 1173 Hz as compared to a mean of 1336 Hz (sd=145). This would appear to indicate that a uvular emphaticizing effect exists for this speaker. However, there is no significant difference in the distribution of schwa F2 between tokens containing emphatic [ɾ̪] but no /q/ and the tokens containing one of /q ṭ ṣ/ ($t(50) = 1.20$, $p = 0.24$). From this we can conclude that the vocalic target of /ə/ is the same after emphatics (and *q*) as after *ṛ* -- in other words, emphatic *ṛ* affects following /ə/ in the same way as other emphatic consonants do.

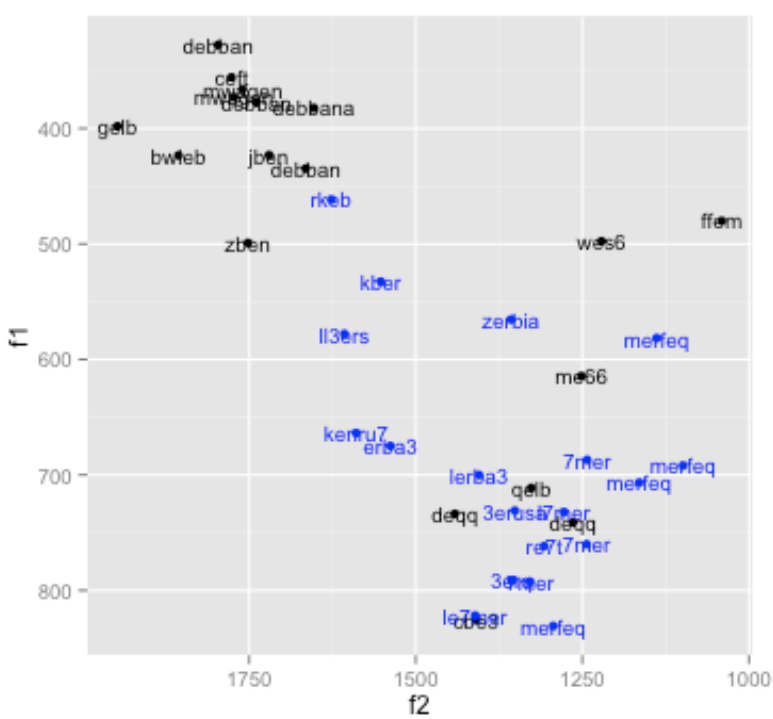

Figure 4. Midpoint formant measures for all /ə/ tokens, with tokens adjacent to /ɾ/ colored blue.

3.1.2 Formant Properties of /r/

In addition to adjacent vowel measurements, I considered the formant properties of the rhotic consonant itself, since identifying an acoustic correlate of emphasis on the consonant itself would reduce the need to extrapolate from the behavior of adjacent vowels. Figure 5 shows the distribution of midpoint *r* formant measures in F1-F2 space. The points highlighted in blue are those words in which *r* appeared to behave as plain based on /i/ or /ə/ measures: *ḥyar*, *ḳber, ʕers*, *rḳeb*, and all words with *r* adjacent to *i* except for *ɣiran*, *ḳiran,* and *ḥmir*. These words cluster together with F1 < 650 Hz, F2 > 1500 Hz, and F3 > 2300 Hz, and overlap with few other words containing /r/. *jra*, *šra*, and *rḳba* are only apparent exceptions, since their patterning with emphatic [ɾ] was determined on the basis of the F2 of word-final /a/ that would have been backed to [ɑ] even if the *r* were non-emphatic, but the overlap of *kiṛan* and *ḥmiṛ* with the plain-r tokens is more anomalous.

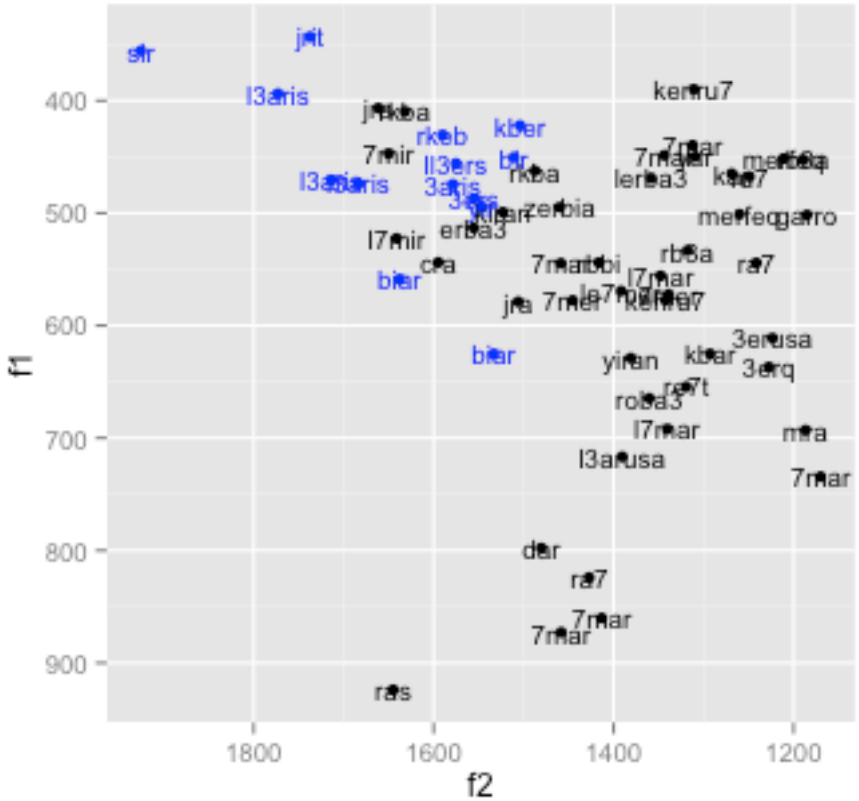

Figure 5. F1 and F2 measures for /r/. Datapoints marked in blue are expected to have non-emphatic [r] based on adjacent vowel analysis.

The patterning of the consonant formants in *kiṛan* and *ḥmiṛ* with plain-*r* words, most of which are adjacent to /i/, suggests that the difference we are seeing may be explained by the presence or absence of adjacent /i/ rather than the emphasis contrast. t-test comparisons do in fact indicate that both membership in the plain-*r* lexical set and adjacency to /i/ are

significant as predictors for all of F1 (lexical: $t(50) = -4.30$, $p < 0.0001$; adjacent-*i*: $t(31) = 3.61$, $p = 0.0011$), F2 (lexical: $t(31) = 7.63$, $p < 0.0001$; adjacent-*i*: $t(17) = -5.30$, $p < 0.0001$), and F3 (lexical: $t(36) = 4.43$, $p < 0.0001$; adjacent-*i*: $t(36) = -4.24$, $p = 0.0001$) of *r*. However, linear regressions modeling both adjacency to /i/ and membership in the lexical set as predictor variables indicate that lexical set membership is the better predictor of both F1 ($t = 2.01$, $p = 0.04$) and F3 ($t = -2.27$, $p = 0.03$), while *i*-adjacency did not approach significance ($t = 0.379$, $p = 0.706$). For F2, both variables were selected as significant predictors (lexical: $t = -3.95$, $p < 0.001$; adjacent-*i*: $t = 3.18$, $p = 0.002$). This regression data must be treated with caution due to its crucial reliance on a small number of data points, but it does suggest that high F1, low F2, and high F3 in the rhotic consonant itself may prove to be reliable indicators of consonantal emphasis.

3.1.3 Vowels Adjacent to Labials

For the velarized labials, the main effects observed were backing of adjacent /a/ and lack of any observed vowel effect when adjacent to /i/. Membership in the class was determined by the presence of a historical **Bw* cluster or, in the cases of /bˠaniu/ and /rbˠbˠi/, by reference to Heath's description of the behavior of the word. Comparison with following vowels after plain labials was possible only for /a/ and /ə/, since other vowels in the dataset occurred post-labially only under either the velarized or plain condition.

Figure 6 shows midpoint formant measures for post-labial /a/ and /ə/ in words without /r/, an emphatic, or a uvular. Comparison of the /a/ after the geminate plain labials in /dəbban/ 'fly (pl.)' and /dəbbana/ 'fly (sg.)' with the /a/ after the velarized labials in /bˠbˠa/ 'my father', /bˠaniu/ 'bath', / mˠmˠagen/ 'clocks', and /mˠmˠagni/ 'clockmaker' revealed a significant difference in F2 ($t(10) = 6.35$, $p < 0.0001$) but not in F1 ($t(10) = 0.56$, $p = 0.5877$). The mean F2 for velarized labials at midpoint was 1300 Hz (sd = 121.6), compared to 1659 Hz (sd = 88.5) for plain geminate /b/ in *debban* and *debbana*.

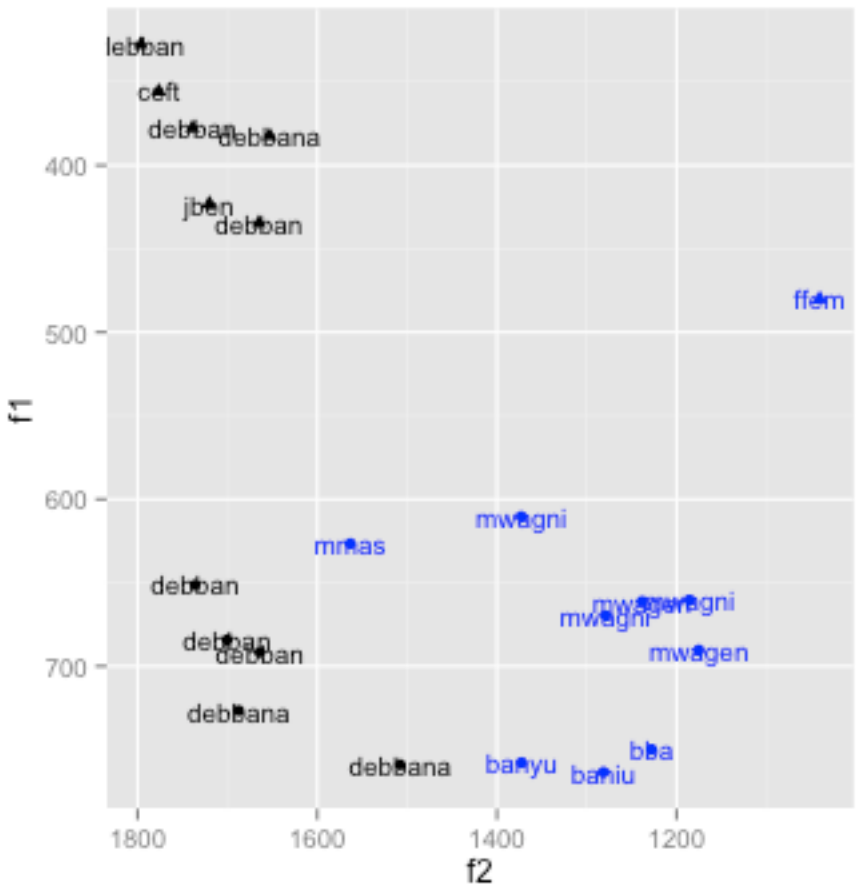

Figure 6: Comparison of midpoint formant measures on adjacent /a/ and /ə/ between plain and velarized labials (velarized tokens are marked in blue).

For vowels other than /a/ following labials, the most notable effect is that the schwa after velarized /f/ in /fˠfˠəm/ 'mouths' has markedly lower F2 (1040 Hz) and higher F1 (480 Hz) than other schwas adjacent to labials (mean = 1725 Hz, sd = 57.9), falling instead closer to the /u/ after /m/ in the word /mus/ 'knife' (F1 = 384, F2 = 1037). Comparison of the /i/ in /fəmˠmˠi/, /rbˠbˠi/, and / bˠbˠiəb/ to /i/ in other contexts, however, did not suggest a unique distribution for the velarized labial segments. ($t(12) = 0.1$, $p = 0.92$ for a comparison of the three SLP tokens against all other words with /i/).

To assess the diphthongization effect reported for vowels following velarized labials, I considered formant measurements at the beginning (25% point) and end (75% point) of each vowel in addition to the midpoint. For /a/, the F2 effect differentiating plain from velarized labials remained significant both early ($t(10) = 12.0$, $p = < 0.0001$) and late ($t(9) = 4.04$, $p = 0.0028$) in the vowel, but the differences between the groups in F1 were not found to be significant at any interval. After velarized labials, the mean F2 of /a/ does increase over the course of the vowel from 1197 Hz at 25%, to 1300 Hz at 50%, to 1354 Hz at 75%, but a one-way ANOVA found these differences not to be significant ($F(2) = 0.45$, $p = 0.639$).

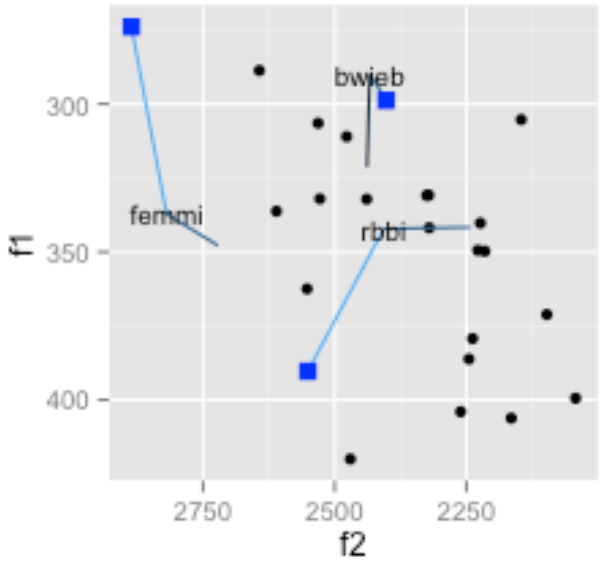

Figure 7. Vowel paths of /i/ after velarized labials, plotted with midpoint measures of the vowel in other contexts (black dots). The endpoint measure for each word is marked with a blue square.

For /i/, there were only three tokens of the vowel after velarized labials, and each token moved in a different direction over the course of the vowel. Figure 7 illustrates the trajectories of the /i/ vowel in these words, which are idiosyncratic and difficult to interpret. In none of these words is the starting point of the /i/ centralized with respect to the general distribution of /i/, and the vowels in/fmˠmˠi/ and /rbˠbˠi/ are moving in distinctly opposite directions with respect to F1. The overall conclusion appears to be that diphthongization is not present, and that the acoustic effect of the velarization is limited to the consonant itself in the case of /i/.

3.1.4 Discussion of Results

Despite the limitations of the data from the preliminary interview, it yields some promising results corroborating previous claims and supporting the direction of the present research. First, the data concerning /a/ and /i/ allophones near /r/ indicate a phonemic distribution of [r] and [ɾ], with original vowel-conditioned allophonic alternations giving way to leveling across morphological paradigms. Two examples of leveling to [ɾ] in the interview, *kaṛ/kiṛan* and *ɣaṛ/ɣiṛan*, are reported by Heath to be widespread in central urban varieties of Moroccan Arabic, while the third, *ḥmiṛ/ḥmaṛ,* levels across a paradigm which is described as typically preserving the alternation. In all cases, the /r/ variant of the singular form has spread to the plural rather than vice versa.

The distributional data from words with other vowels provide some additional support for the phonemicization hypothesis, while confirming the original conditioning environment for [ɾ]. /u/ when adjacent to /r/ is uniformly lower than /u/ adjacent to non-emphatics, supporting the allophonic generalization that adjacent /u/ conditions [ɾ]. When

/r/ is adjacent to /ə/ deriving from SA short /a/, the vowel generally has a low back pronunciation similar to its pronunciation when adjacent to emphatics. However, when /ə/ adjacent to /r/ is derived from SA /i/, the context historically conditioning non-emphatic [r], it has a higher and fronter pronunciation similar to that of /ə/ < **a* when adjacent to non-emphatics. /ə/ < **a* in the neighborhood of /r/ sometimes, however, can pattern with the non-emphatic distribution, as evidenced by /rkəb/ in this dataset.

Even though the distribution of [ṛ] and [r] was found to be phonemic, the patterns of emphasis spread exhibited by /r/ were more restricted than those associated with coronal emphatic obstruents. A lowering and backing effect was only observed on the final portion of preceding /i/ in ***kiṛan***, ***ɣiṛan***, and ***ḥmiṛ***, whereas emphatic obstruents are expected to condition a centralized and lowered allophone of /i/ in the same context. However, *ṛ* was observed to exert the same acoustic effect on following schwa as emphatic obstruents.

A major result with respect to the velarized labials was the categorical difference in F2 of following /a/ between velarized and plain geminate labials. Unlike the difference between plain and emphatic /r/, there was no F1 effect, confirming the result of Zeroual et al. (2011) with respect to vowel effects after pharyngealized coronals and velarized labials respectively. The uniform and pronounced backing of /a/ diverges, however, from Heath's account of /a/ after velarized labials, since he describes the vowel's behavior in this context as erratic, ranging from quite fronted to quite backed.

For this speaker, the /a/ after velarized labials is backed to the same degree as /a/ after emphatic *r*, and the F1 of /a/ after velarized labials and emphatic [ṛ] is also comparable. The difference between labials and *r* in whether F1 has a significant effect is due to a higher allophone of /a/ in the plain condition for *r* as compared to plain labials, rather than a different target for the velarized/emphatic condition. Velarized labials in this dataset also pattern similarly to emphatic *ṛ* in having a negligible effect on adjacent *i*, and in exhibiting a backed allophone of /ə/ in the word /fˠfˠəm/.

The overall impression with respect to the velarized labials is that they exert a backing influence on adjacent vowels which is similar in extent to the effects of emphasis spread visible in the data, but which is not similarly accompanied by any effect in the vertical (F1) dimension. This lends support to my hypothesis that labial velarization is phonologically congruous to emphasis, although it differs from emphasis in its phonetic details. Data from non-adjacent vowels would be helpful to further test this claim, but there were not enough tokens in the data for appropriate comparison.

Finally, the data suggest that uvulars, in particular the uvular stop *q*, have similar effects to emphatics on adjacent vowels, and may be phonologically assimilated into the emphatic class for this speaker. Most striking in this respect is the behavior of the word *qelb* 'heart', which exhibits a back schwa (F1=711 Hz, F2=1327 Hz) when the first consonant is pronounced as [q], but a front and high schwa (F1=397 Hz, F2=1947 Hz) when the first consonant is pronounced as velar [g]. The other word in which schwa is adjacent to a uvular, *deqq* 'to taste', also exhibits a low, back, emphatic-like variant of the vowel. The uvular fricatives /x/ and /ɣ/ do not indicate any special lowering or backing effect, but these only

occur in the data either in the same root as an emphatic/velarized segment or with following /i/, which we have seen can be unreliable as an indicator of emphatic-like behavior.

These results have interesting implications for my research questions, particularly regarding the expected emphasis-spreading behavior of phonemic /ṛ/. Here, even though the segment patterned as phonemic, it exhibited a more limited range of emphasis spread than would have been expected for a primary emphatic. This suggests that in an important way, [ṛ] remains a marginal member of the emphatic phonological class even when it is fully contrastive with a non-emphatic variant, rather than assimilating to the patterns displayed by other emphatic consonants. This result will need to be corroborated by comparison with other speakers and by analysis of primary emphatic tokens following /i/ for the speaker in the preliminary interview, but it is suggestive of a more complex and idiosyncratic pattern of behavior than had been expected for emphasis spread.

Assessing the results concerning velarized labials and uvulars, the data yielded more useful results concerning the emphatic properties of uvulars and their interaction with emphatic *ṛ*. While some generalizations concerning the behavior of the velarized labials could be made, these were not as clearly interpretable due to the limited distribution of the segments. The words *banyu*, *bba,* and *ṛbbi* all provided instances of velarized labials not attributable to adjacent /w/, so according to my hypothesis the velarized labials might be expected to behave as emphatics in their effects on following vowels. In fact, the phonetic target of /a/ and possibly /ə/ after velarized labials was found to be identical to the target of the vowel after /ṛ/, but the effect with comparison to the plain condition differed from that of /ṛ/ in having no F1 component, so it is not clear that the process is the same. In contrast, the uvular stop /q/ was found to unambiguously pattern with primary emphatics and emphatic /ṛ/ in its effect on following /ə/, supporting the hypothesis that uvulars often behave as emphatics. If future interviews fail to yield better results with respect to the velarized labials, it may be preferable to narrow the scope of the study to the behavior of /ṛ/and its interaction with the uvular consonants.

3.2 Data Collection Methods

While the preliminary data yielded some useful results, a larger dataset with a balanced set of vocabulary would have allowed for more thorough and nuanced analysis. Two major shortcomings of the interview were the limited information concerning the behavior of vowels adjacent to both plain and emphatic coronals, and the lack of vowels in comparable non-adjacent contexts across conditions, which are necessary to adequately evaluate the patterning of long-range emphasis spread. In future interviews, these problems may be avoided by preparing balanced wordlists across phonological contexts and prompts eliciting free speech on topics with a high concentration of target vocabulary. In this section, I will discuss the elicitation protocol I propose to use in future interviews, as well as measures that will be taken to investigate intraspeaker variation in the context of phonological change.

3.2.1 Interview Protocol

My plan in conducting future interviews is to combine word-list elicitation such as was used for the preliminary data with prompts eliciting free speech in the speaker's dialect. Using both of these techniques will allow the collection of larger quantities of more natural speech, while still ensuring that a balanced set of target vocabulary is elicited during the course of the interview. It will also provide a minimal dimension of stylistic variation to evaluate for each speaker, which will be important for the evaluation of intraspeaker variation.

Word lists will be designed to allow comparisons between all relevant consonantal groups with minimal differences in phonological context. The *Georgetown Dictionary of Modern Moroccan Arabic* (Maamouri 2015) will be used as a primary lexical resource for compiling word lists. Table 1 provides a sample wordlist comparing plain labials, velarized labials, plain coronals, emphatic coronals, rhotics, velars, and uvulars across vowels in CVC syllables for which the final consonant is plain /s/, allowing for controlled comparison of vowel effects across consonantal groups. Where there are gaps in the lexicon for vowels with following /s/, a minimally different coda consonant has been chosen, except in the case of velarized [mˠmˠ] and [bˠbˠ] which do not occur before either /u/ or /ə/.

	/a/	/ɪ/	/u/	/ə/ (< *a*)
/b/	bəsbas 'fennel'	sərbis 'service'	busa 'kiss'	bəssəm 'to make smile'
/m/	χmasi 'kind of jewelry'	təʃmisa 'sunburn'	mus 'small knife (sg.)'	məska 'chewing gum'
/ bˠbˠ/	bˠbˠas 'kiss (dim. pl.)'	bˠbˠisa 'kiss (dim. sg.)'	---	---
/mˠmˠ/	mˠmˠas 'small knife (pl.)'	[mˠmˠi 'my mother']	---	---
/t/	tasəʕ 'ninth'	tis 'billy goat'	[tut 'mulberry']	təsəffəjit 'doughnut-making'
/ṭ/	ṭas 'washbasin (sg.)'	ṭisan 'washbasin (pl.)'	[ṭub 'clod of earth']	ɣṭəs 'submerge'
/r/	ras 'head'	idrisi 'Idrisid'	rusi 'Russian'	rəsmal 'capital (econ.)'
/k/	kas 'cup'	kis 'coin purse'	mʕkus 'contrarian'	nkəs 'to sweep'
/q/	qas 'stain (pf.)'	nqis 'stain (impf.)'	qus 'arch'	qəsma 'division'
/χ/	χasər 'loser'	bχis 'dirt-cheap'	[χut 'brothers']	abχəs 'cheaper'

Table 1. Wordlist providing vowel comparisons by preceding consonant for vowels with following /s/.

A further goal of the wordlists will be to ensure the elicitation of a number of morphological vowel alternations, particularly in the case of *r* to determine the extent of emphasis leveling across paradigms. Including the singular/plural pairs *ḥmir/ḥmar* 'donkey', *tajer/tejjar* 'merchant', and *ḳbir/ḳbar* 'large' will be of particular importance, since these are reported by Heath (2002) to exhibit alternations in *ṛ* emphasis throughout most dialects, and the preliminary interview showed that at least the alternation in *ḥmir/ḥmar* is susceptible to leveling. Words reported to exhibit differences in *r* patterning in the presence of uvulars will also be elicited, such as *xrif* (pl. *xerraf*) 'autumn', *ɣrib* (pl. *ɣrab*) 'strange', and *qrd* (pl. *qrud*) 'monkey'. The other singular/plural pairs found to exhibit leveling in the preliminary interview, *ḳar/ḳiran* 'bus', *ɣar/ɣiran* 'cave', and *bir/byar* 'well' will be of interest, as will other morphologically related pairs noted by Heath as having variable leveling: *far/firan* 'mouse', *tur/tiran* 'bull', *bar/biran* 'bar', *xruf/xrfan* 'sheep', *dar*/dim. *dwira* 'house', *ṣɣir*/dim. *ṣɣiwer* 'small', and *aṣfar*/dim. *ṣayfer* 'yellow' (the latter two only for dialects in which the *ṣ* is depharyngealized).

Table 2 gives a list of additional words mentioned in Heath (2002) as exhibiting lexically idiosyncratic dialect variation with respect to *r* emphasis, which should be included in the set of target segments.

ʃrəb 'drink'	drrəg 'to hide'
ʒrana 'frog'	rqba 'nape'
gzzar 'butcher'	qdər 'to be able to'
rʒəʕ 'to go back'	mnχər 'nostril'
fkrun 'tortoise'	lχχri 'the last one'
rkba 'knee'	ħrət 'to plow'
ṣdər 'chest'	rħa 'handmill'
bərd 'coldness'	rijəħ 'to sit'

Table 2. Lexical items exhibiting dialect variation in *r*-emphasis.

For nouns, diminutives as well as ablaut plurals can help to provide complete paradigm data, as in the diminutive *qriyyed* for *qrd/qrud* and *dwira* for *dar* above, while for verbs, perfect, imperfect, participial, and verbal noun stems can be compared, as in *tmer*, *yitmer*, *tamer*, *tmir* 'to bear fruit'. For nouns, personal pronoun affixes and the attributive adjective suffix *-i* should be added to test morphological boundary conditions and non-adjacent syllable effects, while inflectional affixes can be used in the same way for verbs. Participial stems of the form $C_1aC_2əC_3$, which exist for most non-derived verbs, provide a particularly useful frame for generating polysyllabic words with a fixed vowel structure for use in testing bidirectional and non-adjacent emphasis spread.

Not all target vocabulary will be elicited via an explicit wordlist. Some words will be targeted in the free speech section, and I also plan to use a reading passage embedding target

words to shorten the necessary length of list-style elicitations. The following short passage, for instance, incorporates eight of the target words identified above, including two singular/plural pairs and one singular/diminutive pair, in addition to other polysyllabic words containing uvulars (*qeddamu*, *qtelha*), emphatics (*mufiḍa*),and pharyngeals (*waḥed*): *l-gezzar r-rusi riyeḫ ʕal arḍ u šaf fkrun waḥed qeddamu. xed l-fkrun mn r-ras u qtelha b-muṣ ṣɣir. le-mmwas le-ṣɣiwera adat mufiḍa bezzaf f-qetl el-fkaren u j-jran*. 'The Russian butcher sat on the ground and saw a tortoise in front of him. He took the tortoise by the head and killed it with a little knife. Little (dim.) knives are very useful tools for killing tortoises and frogs.' I expect that by placing the word list at the end of the interview, and implementing a reading passage elicitation of 8-10 sentences, the necessary length of the word list to establish remaining contrasts can be kept near 30 words.

The wordlist and reading passage portions of the interview will be preceded by elicitation of free speech, so that the interviewee is not primed stylistically by a formal, linguistically self-conscious, and somewhat boring task before being asked to perform conversational speech. The prompts eliciting free speech will focus on vocabulary-rich topics such as food and religious traiditions, as well as topics such as family history which will prime kinship terms with velarized labials (*bba* 'my father', *mmi* 'my mother') and uvulars (*xu* 'brother', *xet* 'sister', *xal* 'maternal uncle'). Sample questions include *ṣṣef li maklatek lemfaḍḍla f ramaḍan u kifaš ketwejjedha* 'Describe for me your favorite foods in Ramadan, and how you prepare them', and *šnu ktetdekker ʕan lʕaila dyalek wenta derri ?* 'What do you remember about your family when you were a child?'. Haeri (1997, p. 29) noted that childhood games provided a particularly successful opening topic when conducting interviews with Cairene Arabic speakers, by reinforcing the desired informality of the interview and the colloquial diglossic context, and I plan to take this approach in my own interview procedure.

Additional strategies available for use during the free speech portion of the interview include the semantic differential technique of Labov (1984) and guided interactions between two native speakers such as the map task of Anderson et al. (1991). Semantic differential elicitation targets specific words by asking speakers the difference between the target word and a semantically closely related word, thus providing repeated tokens of the target word without a linguistically self-conscious target. For instance, the word *ḥmar* 'donkey' could be elicited by asking, *šnu lfarq bin ḥmar u byla* 'what's the difference between a donkey and a mule (*byla*)?' Eliciting verbal interactions between multiple native speakers when possible will help to provide more natural and informal conversational speech, but some structuring by the interviewer will be desirable to ensure elicitation of target vocabulary. The map task is one way to accomplish this in which two speakers are each given a copy of a map with fictitious landmarks, and one speaker must coordinate with the other to reproduce a route which is only marked on one copy of the map. The interaction is designed to elicit the names of the landmarks on the map without direct interference from the investigator. In the context of this study, the map task could easily be used to elicit words such as *busṭa* 'post office' or *blaṣa* 'plaza' with secondarily emphatic [ḇ] and *gezzar* 'butcher' or *nejjar*

'cabinetmaker' with /r/, and could be expanded to include other, non-geographical vocabulary through manipulation of building or street names on the map.

Each interview will have a similar structure, beginning with obtainment of informed consent through a verbal or written protocol explaining that the purpose of the interview is to make recordings of speech for study of colloquial Arabic (*darija*) as it is spoken at home in the speaker's native dialect, that only participants in the research will have access to the recordings and the speaker's real name will not be used in any published work. After this, demographic questions concerning age, occupation, level of education, and places of residence in Morocco will be asked, to control for sociolinguistic variables (see section 3.2.2 below). The interview will then continue with a period of elicited free speech through interaction with the interviewer, proceeding by means of prompts eliciting narratives on relevant topics. If more than one native speaker is available to participate and there is sufficient time, a map task or other recorded conversation between the speakers will follow at this point in the interview. Then the interview will end with a more linguistically focused portion, beginning with a small set of semantic differential questions, moving on to the reading passage, and ending with word list elicitation. In total, each interview is expected to last between 30 and 60 minutes.

Before moving on to discussing subject sampling and recruitment, there are some residual issues related to interview methodology which need to be mentioned. First, there is the question of providing compensation for participants. Direct transactional exchanges in social settings are considered culturally inappropriate in Morocco and other Arab countries, where exchange of gifts is considered a more appropriate expression of thanks for services such as participating in an interview. I have experienced social tension in the past arising from direct financial compensation of participants in a phonetic study, and I propose to avoid this problem for this study by instead providing participants with small gifts, such as tea or food, as a way of thanking them for their time.

A second set of issues involves linguistic competence. While I have a basic conversational command of Moroccan Arabic, I am not a native speaker, and Arabic is my third language. For this reason, I am making every effort to design the interview procedure to involve minimal linguistic intervention on my part, using open-ended prompts encouraging extended narratives and using interactions between native speakers when possible. In addition, specific prompts and reading passages will be quality-checked with a native Moroccan speaker before use in interviews. Another possible way of handling this problem which is available as a last resort, and which I implemented for the preliminary interview, is using French or English as a metalanguage in which to conduct the interview. This was a successful approach for the preliminary data, but it is much more feasible for word-list style elicitation than for conversational prompts. It will be preferable for me to spend some time improving my fluency in Moroccan Arabic before conducting in-country interviews, which I should be able to accomplish given a few weeks in the country (see section 3.3).

Lastly, I need to account for the possible interference of Standard Arabic, given the diglossic relationship between it and the Moroccan colloquial. Making explicit that colloquial speech (*darija*) is the object of study, starting the interview with an informal context, and using explicitly dialectal forms in the prompts are all measures taken to reduce or eliminate this interference. Moroccan Arabic is structurally more divergent from Standard Arabic than many other colloquial dialects, and my previous experience with Moroccan speakers has been that interference from the standard is minimal even in formal word list elicitation, since the grammatical split between the two varieties is categorical. There are many instances of lexical borrowings from Standard Arabic which have been assimilated to CMA phonology, but these forms remain relevant within the context of the study since they are interpreted within the context of the colloquial grammar.

3.2.2 Sampling and Dimensions of Variation

In choosing sites for fieldwork and recruiting interview participants, sociolinguistic and dialectological considerations will be important factors. My primary dialect of interest is the central urban koiné which has arisen over the past century as the cities of Casablanca, Rabat, Meknès and Fès have rapidly grown through migration from rural areas and other parts of the country (Heath 2002; Aguadé 2003). This koiné presents a irregular mixture of features which exhibit more systematic patterning elsewhere in the country, including the [ṛ] ~ [r] alternation and labial velarization. As discussed in section 2.1, /r/ emphasis is mostly allophonic in Saharan dialects but is phonemic in the dialects of northern cities such as Tétouan; similarly, velarized labials are widespread as reflexes of labial-*w* clusters in southern dialects but are absent or lexically sporadic in the north. When these divergent systems come into contact in the koiné, it results in the phonological ambiguity which we have seen both in descriptions of the central dialects and in the preliminary interview of a central urban speaker. A core expectation of this study is that the ambiguity will tend towards resolution as either allophonic or phonemic patterning at the level of individual grammars, but the directional trend of this resolution in the community and its sociolinguistic patterning remain to be determined.

Given that phonological restructuring of /r/ emphasis is being considered here as a change in progress, it will be important to structure the sample of speakers in a way that is well-suited to evaluate its progression in the community. In sociolinguistic studies, the apparent-time construct is often used to approximate the diachronic dimension of variation (Labov 2002, Eckert 1997). By constructing an age-stratified sample with subjects ranging from adolescent to senescent, researchers are able to use the age gradient as a proxy for real-time change in trend studies under the assumption that adult speakers' grammars remain relatively stable throughout their lives. This assumption is particularly reliable for underlying structural representations -- Kerswill & Williams (1994) and Payne (1976), among others, provide evidence that older children and adults have difficulty acquiring idiosyncratic structural features in dialect contact situations.

To ensure age stratification in the present study, I will aim to have a balanced number of interviewees within specific age brackets. For a single locale, I aim to interview four speakers of each gender in the age ranges of 15-30, 30-45, 45-60, and 60+ respectively. I plan to control for the variable of social class by interviewing speakers of a similar economic and educational background in a shared social network, likely a higher educated, professional background given that my existing Moroccan contacts work in higher education (see section 3.2.3). To control for dialect variation within the central region, I plan to focus on a single major urban center as a primary fieldwork site, preferably Fès since I am familiar with this city from previous travel to Morocco and it has undergone recent leveling towards the central koiné from an older Northern-type urban dialect (Heath 2002). To ascertain the patterns of emphasis spread in dialects having typically phonemic and allophonic distributions of *r* and velarized labials for comparison, I will also recruit a small sample of 1-4 speakers in northern and southern areas of the country respectively. The city of Tétouan in the north, and of Erfoud in the south, are likely to be the most practicable local sites for conducting interviews in these regions.

Structuring the subject pool in this way will allow for consideration of sociolinguistic patterns of variation according to gender and class that may exist, as well as the evaluation of change in apparent time. Stylistic variation within the individual will be controlled by the distinction between word lists and reading passages, free speech to non-native interlocutor, and/or free speech to native interlocutor embedded in the interview structure, and may be further refined if necessary by analysis of topical shifts following the model of Labov (1984). Finally, geographical dialect variation will be controlled for by the recruitment of subjects with a homogeneous local background for the main fieldwork site and comparison with the Northern and Southern dialect areas.

3.3 Recruitment, Logistical Considerations, and Timeline

In contacting speakers to participate in the study I plan to take into consideration that, as Niloofar Haeri noted in her study of Cairene Arabic, 'it is culturally more appropriate [i.e. in Arab countries] to contact people, not as a stranger, but as a friend or acquaintance of their own friends or relatives' (1997, p. 23). Accordingly, I plan to expand from a small set of initial contacts to a broader network of speakers through personal introductions and referrals. Basing my research out of Fès will provide the most advantageous starting point for this process, since I have several existing contacts there from my travel in 2013, including a professor of linguistics and a master's student at the University of Fès (Université Sidi Mohamed Ben Abdellah) who has family living in Philadelphia, as well as several instructors at the American Language Center. Before starting fieldwork, I plan to supplement these contacts with others referred to me through academic networks, and to pursue institutional affiliation with either the American Language Center in Fès or the University of Fès for logistical support. Another possible institutional affiliate is the Al-Akhawayn University, 60 km south of Fès in the town of Ifrane.

I expect the duration of my fieldwork to be three months at the minimum, and no longer than nine months. In addition to conducting interviews, I will require some time to build networks of native speakers and to improve my conversational proficiency in the local dialect. Since I already have a working knowledge of the central urban koiné as it is spoken in Fès, several weeks of immersion and continual practice should sufficiently prepare me to conduct interviews. If sufficient resources are at my disposal, I also plan to enlist the help of a native speaker assistant for the free speech portions of interviews.

Availability of external funding will be an important factor in determining the scope and length of my fieldwork. I currently have an application in process for a Fulbright-Hays Doctoral Dissertation Research Abroad (DDRA) fellowship, and am also planning to submit to the National Science Foundation's Linguistics Program - Doctoral Dissertation Research Improvement (Ling-DDRI) award, for which the application is due on July 15. The Fulbright-Hays DDRA, which will report the status of my application no later than September 2015, would support travel and personal cost of living, as well as any equipment costs, for six months of fieldwork abroad. The NSF Ling-DDRI award, which will respond to applications no later than November 2015, would support only travel-related and equipment costs for the duration of fieldwork. Neither source would provide direct funding for in-country research assistants, but would provide me with the opportunity to set aside personal funds for that purpose. Even if I am unable to secure external funding, I expect the Benjamin Franklin Fellowship stipend to be sufficient to fund 3-4 months of fieldwork, especially considering the low cost of living in Morocco with respect to Philadelphia.

I intend to wait until hearing back from the Fulbright-Hays program in September to leave to conduct fieldwork. This will give me the opportunity to use the intervening period to refine interview procedures and collect more preliminary data from Moroccan contacts in the U.S. I have so far arranged for interviews with a Moroccan Arabic lecturer at Johns Hopkins University and a Philadelphia taxicab driver native to the Saharan region, and am pursuing contact with the family of my acquaintance from Fès. I plan to further expand my search for participants over the coming months by contacting the Moroccan Embassy's cultural attaché and seeking references from acquaintances researching Arabic elsewhere in the eastern U.S. Especially in the case that I have limited time to conduct fieldwork in Morocco, these interviews will provide an important supplement to the controlled sample of speakers collected onsite.

An additional source of supplemental data will be previously recorded corpora, of which there are several sources available. The Semitisches Tonarchiv of the University of Heidelberg (Bet-Sawoce et al. 2015) offers a small database of recordings from dialectological researches in Morocco, particularly those of Aguadé (2003) and Behnstedt (2004). I may be able to acquire more comprehensive data from these studies by contacting the authors directly. A second resource for recordings of CMA is the CultureTalk corpus (LangMedia 2015), consisting of interviews between native speakers collected for use in language instructional settings. The recording quality of these interviews is quite good, although some exhibit use of Standard Arabic, and the Moroccan data includes over an hour

of recordings in the dialect of the Western Sahara which could be used as a reference for the Southern dialect type. A third corpus resource which I have been made aware of is a collection of twenty hours of recordings focusing on Moroccan Arabic/French code switching collected by Rebekah Post, a doctoral student at the University of Texas-Austin, which are transcribed and currently being prepared for public dissemination. In addition, publicly available recordings of Moroccan television programs using colloquial Arabic, such as the Ramadan comedy show L'Couple (2M 2015), would provide valuable evidence for normative representations of the central koiné.

I plan to complete the data collection phase of my research no later than April 2016, and to focus on data analysis, interpretation, and theoretical implications during the months following. I expect to complete the dissertation during the fall of next year, and no later than December 2016, or nineteen months from the submission date of this proposal.

References

2M. 2015. 'L'Couple Saison 2.' Retrieved from http://www.lcouple2m.com/.

Aguadé, Jordi. 2003. 'Notes on the Arabic Dialect of Casablanca (Morocco).' In *Proceedings of the Fifth Conference of AIDA (Association Internationale de Dialectologie Arabe)*, ed. I. Ferrando and JJ. Sánchez Sandoval. Cádiz: Universidad de Cádiz (pp. 301-308).

Aguadé, Jordi. 2013. 'Zum arabischen Dialekt von Settat (Marokko)'. In *Nicht nur mit Engelszungen*, ed. R. Kuty, U. Seeger, and S. Talay. Wiesbaden: Harrasowitz.

Al-Masri, M., and Jongman, A. 2004. 'Acoustic correlates of emphasis in Jordanian Arabic: Preliminary results.' In *Proceedings of the 2003 Texas Linguistics Society Conference*. Somerville, MA: Cascadilla Proceedings Project (pp. 96-106).

Al-Nassir, A. A. 1993. *Sibawayh the phonologist*. London: Kegan Paul.

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., ... and Weinert, R. 1991. The HCRC map task corpus. *Language and speech*, 34(4), 351-366.

Behnstedt, Peter. 2004. 'Von an-'Âser (al-Qasr) nach Îgni (Îgli): Ein Vorbericht zu einigen arabischen Dialekten der Provinz er-Rashîdîya (Marokko)'. In *Approaches to Arabic Dialects: A collection of Articles presented to Manfred Woidich on the Occasion of his Sixtieth Birthday*. Leiden: Brill (p. 47 ff.).

Bermúdez-Otero, Ricardo. 2007. 'Diachronic Phonology'. In *The Cambridge Handbook of Phonology*, ed. Paul de Lacy, chapter 21, 497–517. Cambridge: Cambridge University Press.

Bessell, N. J. 1992. *Towards a Phonetic and Phonological Typology of Post-Velar Articulation*. Doctoral Dissertation, University of British Columbia.

Bin-Muqbil, Musaed. 2006. *Phonetic and Phonological Aspects of Arabic Emphatics and Gutturals*. Doctoral Dissertation, University of Wisconsin-Madison.

Blevins, Juliette. 2006. 'A theoretical synopsis of Evolutionary Phonology'. *Theoretical Linguistics* 32:117–166.

Boersma, Paul & Weenink, David (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.09, retrieved 1 June 2015 from http://www.praat.org/.

Browman, Catherine, and Louis M Goldstein. 1986. 'Towards an articulatory phonology'. *Phonology Yearbook* 3:219–252.

Czaykowska-Higgins, E. 1987. *Characterizing Tongue Root Behavior*. Doctoral Dissertation, Massachusetts Institute of Technology.

Dell, François, and Mohamed Elmedlaoui. 2002. *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht, NL: Kluwer Academic Publishers.

Diakonoff, I. 1965. *Semito-Hamitic Languages: An Essay in Classification*. Moscow: Nauka Publishing House.

Eckert, P. (1997). 'Age as a sociolinguistic variable.' In *The handbook of sociolinguistics* ed. F Coulmas. Oxford: Blackwell (pp. 151-167).

Esling, John. 1996. 'Pharyngeal Consonants and the Aryepiglottic Sphincter.' *Journal of the International Phonetic Association* 26(2): 65-88.

Esling, John. 2005. 'There are no back vowels: The laryngeal articulator model.' *The Canadian journal of linguistics,* 50(1), 13-44.

Ferguson, C. A. 1959. 'Diglossia.' *Word*, 15(2), 325-340.

Fruehwald, Joseph. 2013. *The Phonological Influence on Phonetic Change*. Doctoral dissertation, University of Pennsylvania.

Ghazeli, S. 1977. *Back consonants and backing coarticulation in Arabic*. Doctoral dissertation, University of Texas at Austin.

Haeri, Niloofar. 1997. *The Sociolinguistic Market of Cairo: Gender, Class, and Education*. London: Kegan Paul.

Hale, Mark, and Charles Reiss. 2008. *The Phonological Enterprise*. New York: Oxford University Press.

Heath, Jeffrey. 1987. *Ablaut and Ambiguity: Phonology of a Moroccan Arabic Dialect*. Albany: SUNY Press.

Heath, Jeffrey. 2002. *Jewish and Muslim Dialects of Moroccan Arabic*. London: Routledge.

Hoenigswald, Henry. 1960. *Language change and linguistic reconstruction*. Chicago: University of Chicago Press.

Idsardi, William. 2006. 'Canadian Raising, Opacity and Rephonemicization.' *The Canadian Journal of Linguistics* 51: 119-126.

Jakobson, R., and M. Halle. 1956. *Fundamentals of Language*. The Hague: Mouton.

Joos, Martin. 1942. 'A phonological dilemma in Canadian English'. *Language* 18:141–144.

Kerswill, P., & Williams, A. 199). 'A new dialect in a new city: Children's and adults' speech in Milton Keynes.' Final report to Economic and Social Research Council.

Kroch, Anthony. 1989. 'Reflexes of Grammar in Patterns of Language Change.' *Language Variation and Change* 1:199-244.

Labov, W. 1984. 'Field methods of the project on linguistic change and variation.' In *Language in use: Readings in sociolinguistics*, ed. John Baugh and Joel Sherzer (pp. 28-53).

Labov, W. 2001. *Principles of linguistic change, II: social factors*. Massachusetts: Blackwell.

Labov, William, Ingrid Rosenfelder, and Josef Fruehwald. 2013. 'One hundred years of sound change in Philadelphia: Linear Incrementaion, Reversal, and Reanalysis.' *Language* 89:30–65.

LangMedia. 2015. 'CultureTalk North Africa'. Amherst, MA: Five Colleges, Inc. Retrieved from http://langmedia.fivecolleges.edu/culturetalk/north_africa/index.html.

Maamouri, Mohamed. 2015 [in press]. *Georgetown Dictionary of Modern Moroccan Arabic: Arabic to English*. Washington: Georgetown University Press.

McCarthy, John. 1994. 'The Phonetics and Phonology of Semitic Pharyngeals.' In *Phonological Structure and Phonological Form: Papers in Laboratory Phonology II*, ed. P. Keating (pp. 191-233). Cambridge University Press.

Mielke, J. 2008. *The Emergence of Distinctive Features*. New York: Oxford University Press.

Moisik, Scott. 2013. *The epilarynx in speech*. Doctoral dissertation, University of Victoria.

Ohala, John J. 1990. *The phonetics and phonology of aspects of assimilation*. In Papers in laboratory phonology 1: Between the grammar and physics of speech, ed. John Kingston and Mary E. Beckman, 258–282. Cambridge: Cambridge University Press.

Payne, A. 1976. 'The acquisition of the phonological system of a second dialect.' Doctoral dissertation, University of Pennsylvania.

R Core Team. 2013. 'R: A language and environment for statistical computing.' R Foundation for Statistical Computing, Vienna, Austria. URL http://wwwṛ-project.org/.

Schroepfer, Jason. 2015. 'An Acoustic Investigation of /tˤ/ in Aswan Arabic'. Paper presented at the 29th Annual Symposium in Arabic Linguistics. Milwaukee, WI.

Shahin, Kimary. 2002. *Postvelar Harmony*. Amsterdam: John Benjamins.

Shoul, Karim. 2009. Formants as cues for pharyngealization perception in Oriental Moroccan Arabic. In *Proceedings of the International Workshop on Pharyngeals & Pharyngealisation*. Newcastle-upon-Tyne : Newcastle University.

Sylak-Glassman, J. 2014. *Deriving Natural Classes: The Phonology and Typology of Post-Velar Consonants*. Doctoral Dissertation, University of California, Berkeley.

Bet-Sawoce, J., C. Miller, and M. Shawarbah. 2015. 'Semitisches Tonarchiv.' Retrieved from http://www.semarch.uni-hd.de/index.php43.

Watson, J. C. 2007. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Younes, Munther. 1993. 'Emphasis Spread in Two Arabic Dialects.' In *Perspectives on Arabic Linguistics V*. Amsterdam: John Benjamins.

Younes, Munther. 1994. 'On Emphasis and r in Arabic.' In *Perspectives on Arabic Linguistics VI*, ed. M Eid, V. Cantarino and K. Walters. Amsterdam: John Benjamins.

Youssef, Islam. 2013. *Place Assimilation in Arabic: Contrasts, Features, and Constraints*. Doctoral Dissertation, University of Tromsø.

Zawaydeh, B. A., and K. de Jong. 2011. 'The phonetics of localizing uvularisation in Ammani-Jordanian Arabic: An acoustic study.' In *Instrumental Studies in Arabic Phonetics*, ed. Z. Hassan and B. Heselwood. Amsterdam: John Benjamins.

Zeroual, C., J. Esling, and P. Hoole. 2011. 'EMA, endoscopic, ultrasound and acoustic study of two secondary articulations in Moroccan Arabic: Labial-velarisation vs. emphasis'. In *Instrumental Studies in Arabic Phonetics*, ed. Z. Hassan and B. Heselwood. Amsterdam: John Benjamins.