University of Pennsylvania Working Papers in Linguistics

Volume 25
Issue 1 *Proceedings of the 42nd Annual Penn Linguistics Conference*

Article 15

2-27-2019

A Corpus Phonetic Study of Contemporary Persian Vowels in Casual Speech

Taylor Jones
University of Pennsylvania

A Corpus Phonetic Study of Contemporary Persian Vowels in Casual Speech

Abstract
Contemporary Iranian Persian is described as having a six vowel system; however, there is currently very little work that characterizes the exact nature of these vowels. Previous works posit substantially different vowel spaces, and conflicting accounts of the extent to which historical length distinctions are still relevant in Persian -- these distinctions then affect phonological theorizing, especially with regards to vocalic assimilation (sometimes referred to as "vowel harmony"). This study uses a corpus of over 60 hours of casual telephone speech among 104 speakers to describe the vowels of Persian. It is demonstrated that historical length distinctions no longer obtain, that previous descriptions of the vowel space of Persian are no longer necessarily accurate, and that the low back vowel may no longer be a steady state vowel for all speakers nor as low or as rounded as previously described.

A Corpus Phonetic Study of Contemporary Persian Vowels in Casual Speech

Taylor Jones

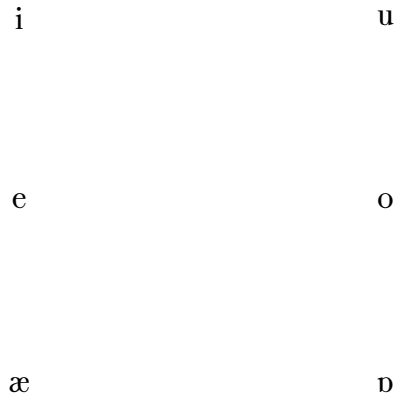
1 Introduction

Contemporary Iranian Persian (CIP) is generally argued to have six phonemic vowels, however their exact characterization has been a source of controversy in the literature. The goal of this study is to provide a phonetic foundation for the analysis of Persian, both for future sociolinguistic study and to better inform phonological theorizing about CIP.

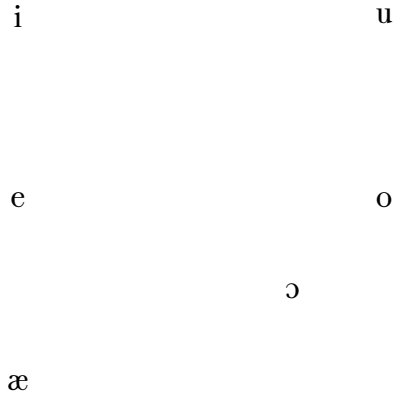
The vowel system of Contemporary Iranian Persian is a relatively straightforward 6 vowel system, but with two sources of controversy:

- Are historical length distinctions preserved? Are they relevant for the phonology?
- What is the low back vowel?

There are a number of different vowel spaces that have been proposed for Persian. More accurately, there are a number of different competing vowel spaces that are *assumed* a priori in the literature on Persian, which are then taken as the basis for phonological theorizing. Lazard (1992), Toosarvandani (2004), and Miller (2013) use an abstract schematization which assumes roughly even distribution of the vowels in the vowel space, and posits the low back vowel is /ɒ/ (consistent with older literature on Persian, Figure 1a). Both Ansarin (2004) and Aronow et al. (2017) claim the vowel space is somewhat different, especially with regards to the back vowel, which is claimed to be something closer to /ɔ/ (Figure 1b). The International Phonetic Association (1999) claims yet a third vowel space for CIP (Figure 2).



(a) Persian Vowels, according to Lazard (1992), Toosarvandani (2004), and Miller (2013).



(b) Persian Vowels, according to Ansarin (2004) and Aranow et al (2017).

Figure 1: Proposed Vowel Spaces for Persian.

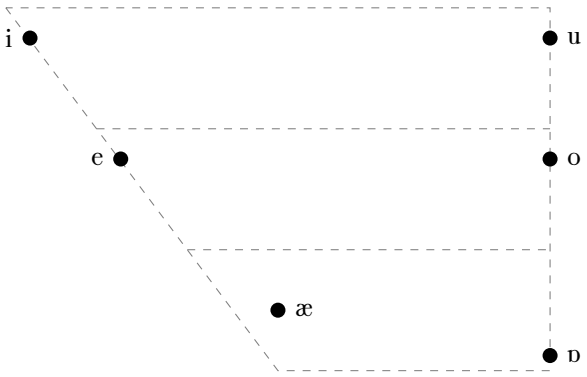


Figure 2: Persian vowels according to the *JIPA*.

One may reasonably ask why this matters. The answer is that it is relevant to a number of unresolved issues in the analysis of Persian, which then have implications for broader linguistic theorizing. Foremost among these issues, the exact categorization of Persian vowels will affect the interpretation of Persian vowel assimilation (sometimes loosely referred to as "harmony"). It is not settled whether place or (possibly historical) vowel length is the key factor in triggering vocalic assimilation in Persian. If place is potentially relevant, we obviously need to know what the places are. Assimilation is always regressive, and generally described as occurring in bisyllabic words. An example of vocalic assimilation in CIP is given in Figure 3.

In fact, the continued existence of a historical length contrast is still the subject of ongoing debate in the Persian literature. Lazard (1992) claims length distinctions present in Middle Persian still obtain in speech (Figure 4). Toosarvandani (2004) and Rahbar (2009), among others, appeal to length distinctions directly to explain phonological processes of assimilation. However, none of the (phonological) works claiming length distinctions are relevant to the phonology of Persian empirically demonstrate the existence of such a distinction.

Currently, there are only two phonetic studies of Contemporary Iranian Persian vowels. Ansarin (2004) recruited 12 female undergraduates from Tabriz, Iran, and asked them to read word lists. Aronow et al. (2017), in a pilot study, recruited two speakers, one male and one female, both from Tehran, and also used word lists. Both studies have relatively small sample sizes — Aronow et al. (2017), for instance measured 45 vowels from each speaker, for a total of 90 observations. The results of both studies differ from previous accounts of Persian vowels, and from each other (although we should expect regional variation between Tabriz and Tehran, which are 633km apart).

The present study makes use of the CALLFRIEND corpus (Canavan and Zipperlen 1996) from the Linguistics Data Consortium (LDC) to determine the vowel space of CIP, and finds that there is no significant length distinction between vowels, that there is regional variation in the vowel space of CIP, that the low back vowel is higher than often assumed (consistent with Ansarin 2004, Aronow et al. 2017), and that the low back vowel may be better characterized as a diphthong.

1.	*devist*	→	*divist*	'two hundred'
2.	*forush*	→	*furush*	'sale'
	fozul	→	*fuzul*	'impertinent'
	sholuq	→	*shuluq*	'crowded
3.	*jahân*	→	*jâhân*	'world'
	maʔâsh	→	*mâʔâsh*	'livelihood'

Figure 3: Vocalic Assimilation in Persian.

i, ī	u, ū
e	o
æ	a, ā

Figure 4: Middle Persian vowels, following Lazard (1992).

2 Methods

The present study makes use of the CALLFRIEND Farsi corpus from the LDC. The corpus consists of over 60 hours of casual telephone speech between native Persian speaking friends and family. There are 104 speakers in total, from multiple cities in Iran covering the main population centers (Figure 5). The recordings were made at 8Khz, and transcribed by native speakers in both Arabic script and a consistent romanization.

Figure 5: Cities represented in the corpus.

The corpus was used to train an HTK-based forced aligner (Young et al. 2002), using the McGill Prosody Lab wrapper (Gorman et al. 2011). First, the data were processed. This entailed splitting the data into utterances. Utterances with English words or code switching were discarded. Utterances that included laughter, coughing, or other noises (present in the transcription as {LAUGH} {COUGH}, etc. were discarded. This left 21,002 unique utterances used to train the aligner. In training the aligner, the low back vowel was represented as <A>. It will be represented as such below to avoid making an a priori claim that it is /ɑ/, /ɔ/, or some other vowel. Vowel extraction and measurement was performed using scripts written in Praat and in R.

3 Results

Alignment was essentially perfect at the word level, and excellent at the phone level (a sample alignment is presented in Figure 6). The end result was 70,711 force-aligned vowel observations — four orders of magnitude greater than previous studies.

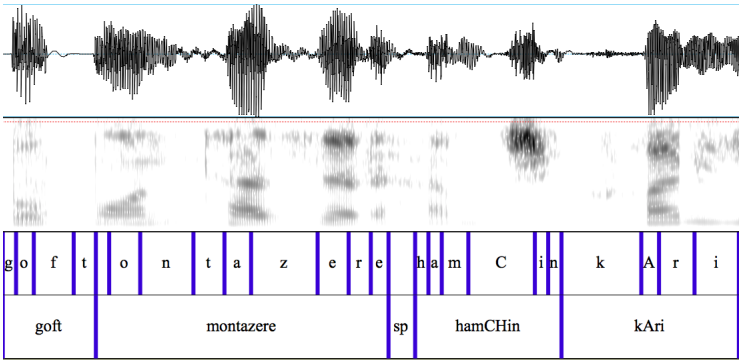


Figure 6: A sample alignment.

Because of the nature of the sample (that is, who did the most speaking), the observations were skewed in favor of male speakers (Table 1). However, there is a good volume of observations from both genders and across a range of ages (Figure 7).

speaker gender	number of tokens
Female	65,822
Male	108,739

Table 1: Observations by gender.

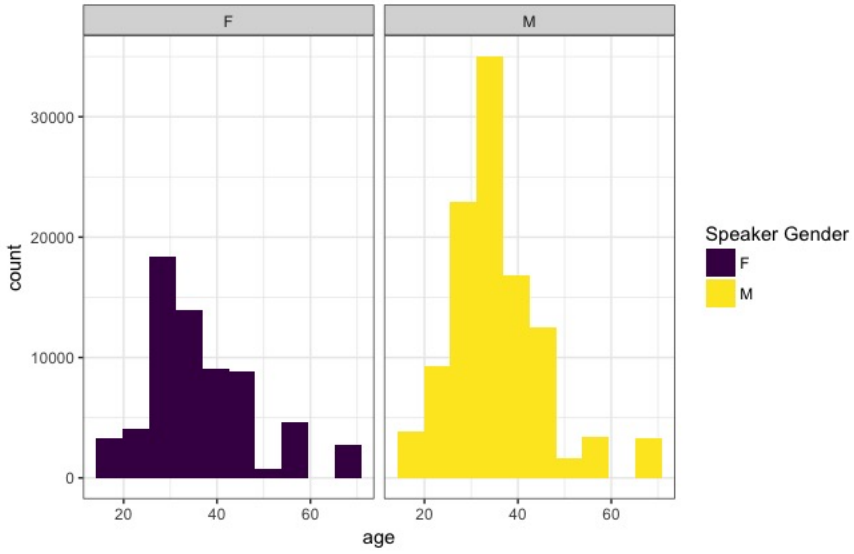


Figure 7: Observation counts by age and gender.

Vowel durations were not normally distributed, but rather, as expected in casual speech, they were significantly skewed (Figures 9a and 9b). In previous literature, /e/ has been argued to be the "default" vowel (Jahangiri 1980). Consistent with that argument, it is the most frequent in the corpus, and the most frequently reduced (Figure 8, Table 2). Vowel duration distributions were remarkably similar across vowel class, all peaking at roughly 40ms, with long tails, although there was some suggestion of a bimodal distribution within the low front vowel durations (Figure 9). While a Mann-Whitney test did find a number of the vowel classes had significantly different means from one another, the difference in means was less than 10ms and with an average sample size of

11,785 observations, one would expect *any* difference to be significant, regardless of whether that difference is linguistically meaningful.

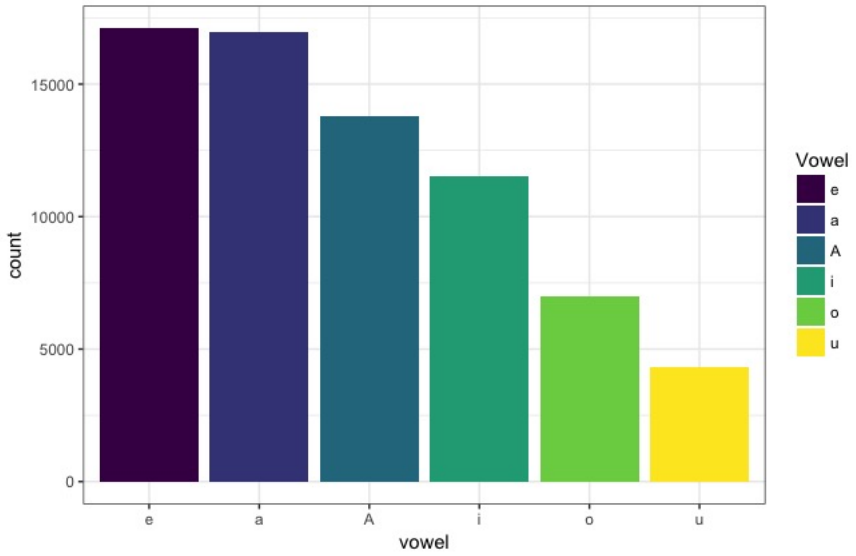
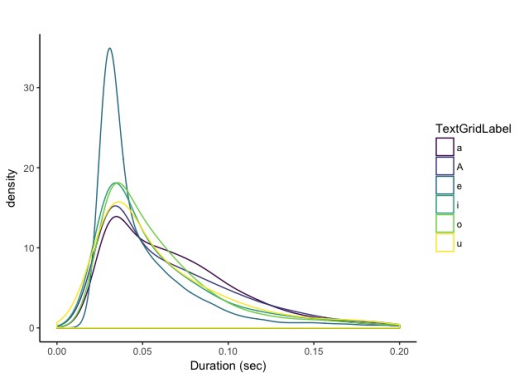
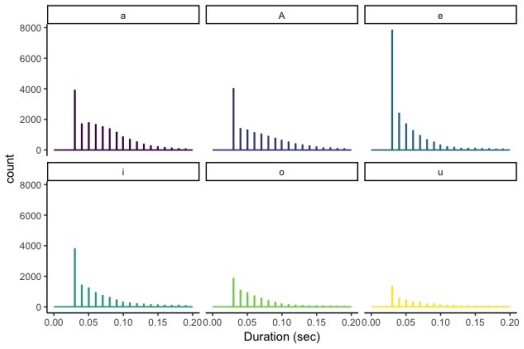


Figure 8: Vowel counts by class. "A" represents the low back vowel, "a" represents the low front vowel.



(a) Vowel Durations by class.



(b) Vowel Durations by class (excluding observations less than 30ms).

Figure 9: Distribution of Vowel Durations.

Vowel class	median duration (ms)	mean duration (ms)	tokens
æ	60	73.67	16,963
A	60	72.43	13,798
e	40	55.86	17,103
i	50	73.48	11,527
o	50	68.61	7,013
u	50	77.46	4,307

Table 2: Vowel class distribution statistics.

Persian has word final stress (with some morphosyntactically motivated exceptions). Stressed vowels are described as longer in duration (Lazard 1992). One fact that emerged from (non-automated) analysis of the corpus is that Contemporary Iranian Persian vowels significantly reduce in unstressed position. In fact, in many instances, the reduction is so extreme that there is no surface vowel. For example:

- *ruz* →[rz] 'day'
- *hich* →[ht͡ʃ] 'any'
- *chiz* →[t͡ʃs] 'thing'
- *pul* →[p^{h}l̥] 'money'
- *miforusham* →[mifrʃæm] 'I am selling'
- *miforushin* →[mifrʃn] 'you (pl.) are selling'
- *pas* →[p^{h}s] 'so'

The majority of the observations (60.29%) were from Tehran (Table 3), where there was a strong gender asymmetry in the distribution of the high front vowel, with men exhibiting a tenser, fronter /i/ than women (Figure 10). There was also evidence of regional variation, with Isfahani men exhibiting a lowering of the high back vowel, lowering of the low back vowel, and raising of the front mid vowel relative to Tehrani men (Figure 12).

raised	tokens	percent
Astara	2,628	1.48%
Babol	5,108	2.88%
California	2,957	1.67%
Hamedan	6,033	3.40%
Isfahan	9,847	5.56%
Kerman	1,227	0.69%
Mashhad	14,971	8.45%
Shiraz	20,135	11.36%
Shirvan	3,074	1.73%
Tehran	106,858	60.29%
USA	1,723	0.97%
Unknown	2,675	1.51%

Table 3: Geographic distribution of observations.

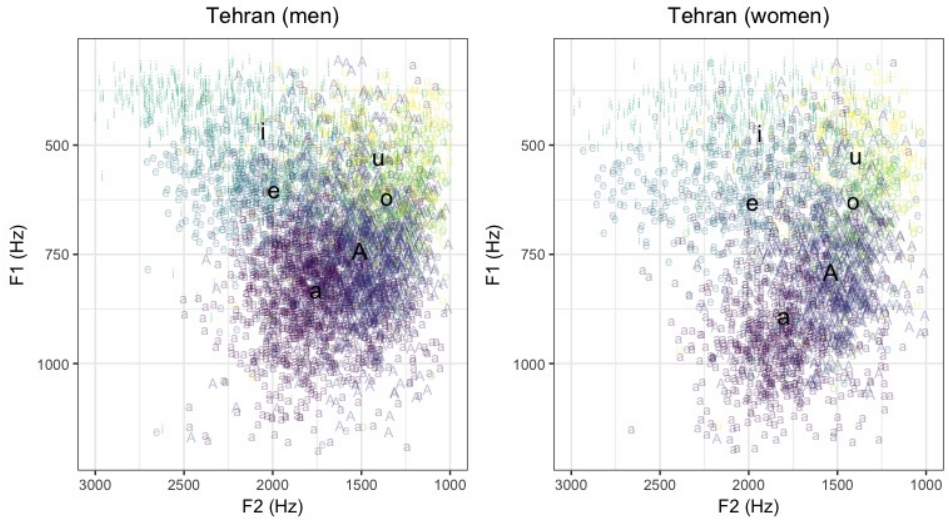


Figure 10: Tehrani Vowel Space. Left: men; Right: women.

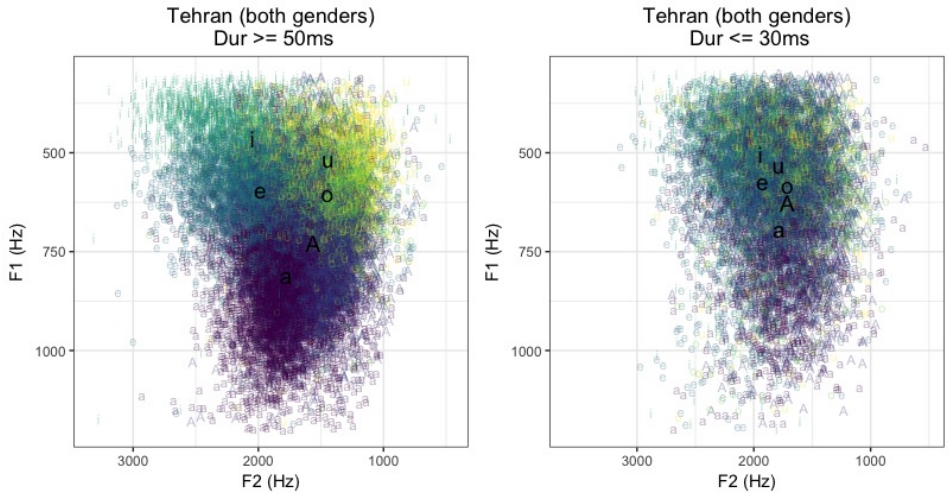


Figure 11: Tehrani Vowel Space.

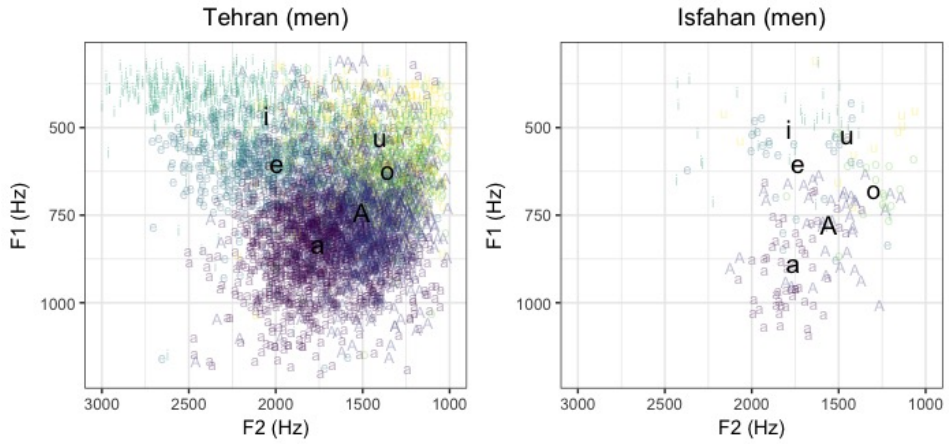


Figure 12: Vowel Spaces for Tehrani and Isfahani men.

Across the corpus, there was evidence that the low back vowel may not be a single steady vowel for all speakers. Rather, it looks (and sounds) like a diphthong for many speakers (Figure 13). More work is needed to tease apart regional and social factors that may be related to this phenomenon.

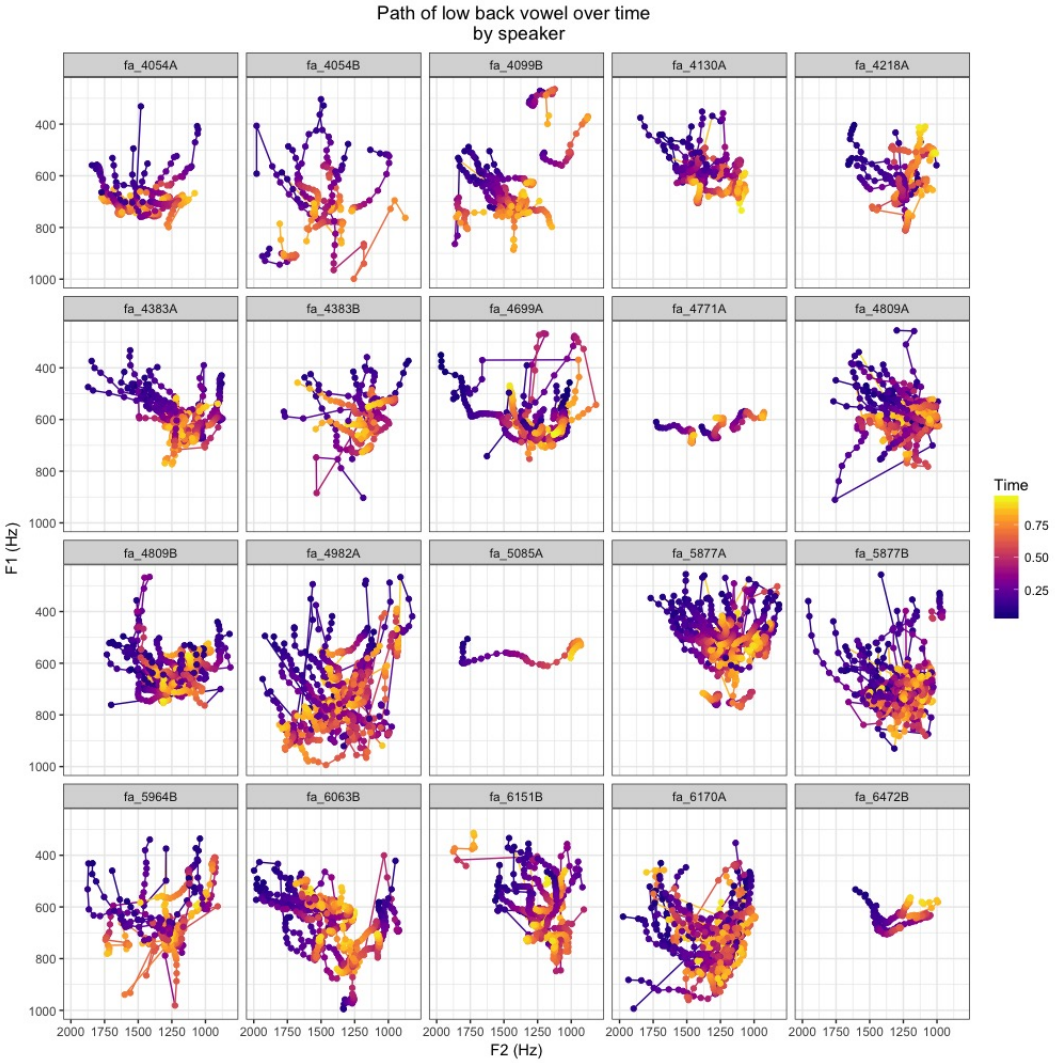


Figure 13: Low back vowel formant tracks (time normalized).

4 Discussion

The goal of this study was to provide a phonetic foundation for the phonological analysis of Persian vocalic assimilation phenomena. Rather than simplifying the analysis of these phenomena, these data create further problems for phonological theorizing. First, it is clear that Persian vowels significantly reduce in non-stressed positions. Given that all accounts of vocalic assimilation in Persian (to my knowledge) are examples of regressive assimilation, and most are essentially iambs (cf. Figure 3), the evidence for vocalic assimilation as a phenomenon that *requires explaining* is significantly weakened, even before taking into account examples like *jahân/jâhân*, where the intervocalic consonant is known to disappear in casual speech (Lazard 1992), putting vowels in hiatus, and resulting in a surface form like *jâ:n*. Second, to the extent that there is weak evidence for a difference in vowel durations, the mean durations by vowel are not consistent with the expectation that Middle Persian long vowels have remained long (cf. Table 2 and Figure 9). That is, there is scant evidence

from the above that historical length distinctions could or should be adduced to explain phonological patterns in Contemporary Iranian Persian, and the above casts some doubt on whether those patterns are sufficiently empirically established to warrant explanation anyway. Third, the low back vowel, for many speakers, is not a steady-state vowel, but is similar to the /ɔ/ with an offglide associated with stereotypical New York pronunciations of *coffee* or *dog*.

This study suggests three questions for future research. First, how empirically real, and distinct from normal reduction processes, is the phenomenon of vowel assimilation in casual Contemporary Iranian Persian? More study is needed, focusing specifically on the extent to which speakers actually exhibit vocalic assimilation, especially in careful speech contexts. Second, assuming assimilation is empirically supported, how should the phonetic patterns here inform our thinking about the phonological explanation? Historical length distinctions seem to no longer obtain in casual speech (although it may subsequently prove that such a distinction is obscured but present in the current data). And third, given a corpus with such a good balance of age, gender, education, and location, what are the patterns of sociolinguistic variation in the vowelspace of Contemporary Iranian Persian?

While the present study cannot answer those questions, it is clear that we may need to rethink our assumptions about Persian phonology, specifically assumptions about vowel length and height contrasts, especially if such assumptions are used to explain phonological contrasts that are assumed to hold across regions and despite gender differences. While much more work is needed to tease apart the sociolinguistic factors that affect Contemporary Iranian Persian variation and the phonetic and phonological factors at play with regards to vocalic assimilation, the above can serve as a starting point, providing the first large-scale, empirical description of the vowel space of Contemporary Iranian Persian.

References

Ansarin, Ali Akbar. 2004. An Acoustic Analysis of Modern Persian Vowels. In *9th Conference Speech and Computer*.

Aronow, Robin, Brian D McHugh, and Tessa Molnar. 2017. A pilot acoustic study of modern persian vowels in colloquial speech. *Proceedings of the Linguistic Society of America* 2:17–1.

Canavan, Alexandra, and George Zipperlen. 1996. *CALLFRIEND Farsi LDC96S50*. Linguistic Data Consortium.

Gorman, Kyle, Jonathan Howell, and Michael Wagner. 2011. Prosodylab-aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39:192–193.

International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

Jahangiri, Nader. 1980. A sociolinguistic Study of Tehrani Persian. Doctoral dissertation, University of London.

Lazard, Gilbert. 1992. *A Grammar of Contemporary Persian*. Mazda Publishers.

Miller, Corey. 2013. Variation in persian vowel systems. *Orientalia Suecana* 61:156–169.

Rahbar, Elham Rohany. 2009. On Contrasts in the Persian Vowel System. *Toronto Working Papers in Linguistics* 31.

Toosarvandani, Maziar Doustdar. 2004. Vowel Length in Modern Farsi. *Journal of the Royal Asiatic Society* 14:241–251.

Young, Steve, Gunnar Evermann, Mark Gales, Thomas Hain, Dan Kershaw, Xunying Liu, Gareth Moore, Julian Odell, Dave Ollason, Dan Povey, et al. 2002. The htk book. *Cambridge university engineering department* 3:175.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104
tayjones@sas.upenn.edu