

A HOLISTIC TREATMENT OF /ān/ TO [un] IN PERSIAN

Corey Miller

University of Maryland, USA

cmiller6@umd.edu

ABSTRACT

One of the most commonly mentioned features of informal spoken Iranian Persian in contrast to its formal or written form is the pronunciation of /ān/ as [un], as in [tehrun] ‘Tehran’, and [iruni] ‘Iranian’. While several aspects of this phenomenon have been reported, a comprehensive treatment of the dialectal, sociolinguistic, phonological and phonetic aspects has not appeared. This paper attempts to pose the most relevant questions, providing answers where possible, and highlighting areas where more investigation is required.

Keywords: Persian, sociolinguistics, dialectology, phonology, phonetics

1. INTRODUCTION

Persian orthography represents /ā/, long /a/, with the letters *ālef madd* (usually word-initially) and *ālef* (non word-initially). In formal speech, in Iranian Persian, it is realized as a back rounded vowel, [ɔ], in contrast to the short /a/, which is realized as [æ] [20]. In informal speech, in many parts of Iran, /ā/ is pronounced [u] before /m/ or /n/, as in /xāndan/ [xundæn] خاندن ‘to read’, or /āmad/ [umæd] آمد ‘he/she/it came’. However, even in informal styles, certain words do not undergo this change, such as /gorān/ *[gorun] قرآن ‘Koran’.

After discussing whether Persian is a diglossic or homoglossic language, this paper will first attempt to ascertain the phonetic motivation for the change. It will then explore the broader linguistic context, covering both sociolinguistic and geographical aspects. Finally, we will examine the historical context, and the conclusion will strive to offer an explanation of the phenomena observed: namely that /ān/ to [un] was a phonetically motivated “change from below” that began long ago and has never reached completion, while the current tendency to change [un] back to /ān/ in formal styles represents a “change from above” [17].

2. HOMOGLOSSIA OR DIGLOSSIA?

Persian exhibits systematic divergence between its written/formal and spoken/informal forms in pronunciation, morphology and grammar. While these differences seem much greater than those found in English, it is not clear whether they approach the canonical examples of diglossia, such as has been observed in Arabic-speaking countries or among German speakers in Switzerland [6]. [23] attempted to quantify diglossia and compare Persian to Arabic. His conclusion, based both on a quantitative examination and the historical sociolinguistic situation, is that Persian is homoglossic.

[7] discusses a developing writing style using the letter *vāv* و to represent raising. Indeed the use of such a style, reflecting the way the language is spoken, rather than the way it has traditionally been written, has ignited a “vulgarity debate” on the Persian blogosphere [4].

3. PHONETIC CONTEXT

3.1. Nasalization effects

According to [1], in the context of nasals, “vowel height becomes centralized—that is, nasalization lowers high vowels and raises low vowels”. [12] notes that a nasalized [a] can have a much lower F_1 than normal. In the Persian case, a low vowel is indeed raised, but why should it be raised several steps to [u]? [12] describes the closely spaced formants of a nasalized [a] as resembling those of [u].

The raising discussed here does not occur when the nasal precedes /ā/, so /māh/ *[muh] ماه ‘moon’, /nār/ *[nur] نار ‘pomegranate’. According to [14], syllable-final nasals nasalize vowels more than syllable-initial nasals; however, she provides a catalogue of languages whose vowels undergo either anticipatory or perseverative nasalization.

3.2. Rounding effects

The fact that /ā/ is starting in a rounded position, [ɔ], may also explain its propensity to shift to /u/, since rounding of low vowels also has the effect of lowering F₁ [28].

Note that the /ā/ formerly used in Indian varieties of Persian, and perhaps in Afghanistan, whose dialect shares many similarities with that used in India, is not rounded to the same degree as in Iran. According to [24], writing in 1894:

In the last generation, ā was generally sounded like our *a* in *ball*; but though still so sounded in parts of Persia (and especially in the Kāshān dialect), it is now becoming very common to give it the sound of our *a* in *bar*, as it had in Persia in olden times, and has, even now, in the Persian of India.

The current Iranian tendency to continue using a rounded vowel, [ɔ], indicates that the above-mentioned trend did not catch on. However, it is important to note that Indian and Afghan varieties, which lacked rounding, crucially lack the raising of /ā/ to [u] [5]. However, [25] points out that the single highly frequent word /ān/ آن ‘that’ can be pronounced with [u] in both Afghan and Tajik dialects.

3.3. Vowel reduction

Another way of looking at the fact that [u] appears in informal situations as a reflex of [ɔ] is in the context of vowel reduction. [u] has been found to be a reduction target for /ɔ/ in both Catalan and Brazilian Portuguese [3]. In a report on vowel reduction in conversational Persian [27], [u] has also been found to be a reduction target for short [ɔ].

3.4. Effect of velars

Finally, the change does not occur before the velar nasal /ŋ/. So /bāŋ/ *[buŋ] بانگ ‘shout’, /dāŋ/ *[duŋ] دانگ ‘sixth of a property’. This exclusion of the velar nasal from the raising environments is similar to the behavior of /æ/ in the English of New York City and Philadelphia. According to [16], following /m,n/ lead to tensing and raising of /æ/ in both cities, while following /ŋ/ does not. However, [29] describes the dialect of Milwaukee, where following /ŋ/ is very much at the forefront of raising environments for /æ/.

4. LINGUISTIC CONTEXT

[22] provides one of the most detailed analyses of the classes of words that are either subject or not to the sound change under discussion in two Iranian cities: Tehran and Ghazvin, a smaller city about 165 km northwest. As will be discussed below, there is more raising in Tehran than Ghazvin, but the pattern is generally similar, with an important exception to be discussed below.

In both cities, while some words tend to undergo raising in informal speech, other words do not change regardless of style; examples include /dānefɣāh/ *[dunefɣɔh] دانشگاه ‘university’ and /pejkān/ *[pejkun] پیکون ‘a type of car’.

As a result of an investigation with four speakers, [13] note that, apart from verbs, it is rare to find a word where /ān/ becomes [un] in antepenultimate position; e.g. /hāmele/ *[humele] حامله ‘pregnant’. However, they note that when a verb stem contains /ān/, it will change to [un], whether or not it appears in antepenultimate position, due to suffixation of personal endings: e.g. /āmadam/ [umədæm] آمدم ‘I came (came.I)’

In general, [22] claims that words appearing with greater frequency are more likely to undergo raising. Among words that do not reliably exhibit raising, he identifies a class of words that, while undergoing raising much less frequently, are “in the process of being accepted as members of the group which favor the rule application once they achieve higher frequency.”

[22] identifies words where raising does not occur at all, among which we will examine a subclass of native Persian words, and foreign (non-Arabic) borrowings.

4.1. Persian words

[22] identifies linguistic constraints that may disfavor raising in a subclass of Persian words. For example, a morpheme boundary in between /ā/ and the nasal inhibits raising: /dʒā#namāz/ *[dʒunæmɔz] جانماز ‘prayer rug (place.prayer)’.

Another inhibiting factor may be dissimilation. For example, when two raising environments occur in a word, usually only the final one, which is stressed, undergoes raising: /sāmān/ [sɔmun] سامان ‘order’, /āmijāne/ [ɔmijune] عامیانه ‘colloquial’.

4.2. Foreign borrowings

Persian has borrowed extensively from Arabic since the Islamic conquest. [19] reviews several

investigations over a span of over 600 years until the 1950s, indicating that the number of Arabic types in sample texts has hovered around 40%, while the number of Arabic tokens has hovered around 20%.

While [22] finds that Arabic words, with some exceptions, undergo raising, he identifies a class of words borrowed from other languages which do not, including /lāmp/ *[lump] لامپ ‘light bulb’ and /mānto/ *[munto] مانو ‘loose outer garment for women’.

5. SOCIOLINGUISTIC CONTEXT

[22] examines the correlation of raising with the following factors: style, education, age and gender. He identified four levels of style, on a continuum from casual speaking to reading minimal pair lists. He found that levels of raising decreased as the style became more formal. The most salient drop was between speaking and reading styles. Averaging across age and educational groups, in the most informal speaking styles, approximately 70% of phonologically applicable contexts were realized with [u], in contrast to reading styles, in which raising never occurred more than 6% of the time, even among those with fewer than six years of education.

With respect to age, it was found that younger speakers tended to raise more than older. While with respect to gender, it was found that women raised more than men.

[22] considered four levels of education. In Tehran, within a given age group, the more educated tended to raise less. However, Ghazvin shows the opposite pattern, with the most educated exhibiting the most raising. This suggests that the educated in Ghazvin may derive some prestige by approximating the pattern found in the capital.

6. GEOGRAPHICAL EXTENT

While raising is often associated with the dialect of Tehran [18], it is in fact common in many parts of Iran. As already mentioned, [22] described raising in Ghazvin. Since its overall rate of raising is less, Tehranis consider the Ghazvin dialect “rural or bookish”. [11] describes raising of /ā/ before /m, n/, but not /ŋ/ in Khorasan, a province in the northeast of Iran, bordering Afghanistan. In the Hazaragi dialect of central Afghanistan, [15] reports that the plural animate suffix /-ān/ is realized as [u]; which is precisely an environment not found to raise in Iran, due to its perceived

formality [13]. [21] found that the Hazaragi process is actually more general, providing examples of nonplural /ān/ realized as [u]: /gerawgān/ [grawgu] گروگان ‘hostage’ and /majdān/ [maydu] میدان ‘plaza’.

7. HISTORICAL LINGUISTICS

[26] contends that using [u] for /ā/ before nasals has been an option in Persian for over 1000 years. As an early example, he gives the name of Mount Behistun, site of notable Old Persian inscriptions, whose name comes from an earlier form, Baghastān. However, in the intervening period, before the present day, there is evidence that the sound change we are discussing went through an earlier stage in which /ā/ became [o] before progressing to [u].

7.1. Early evidence for prenasal /o/

[10] describes a Persian translation of the Koran in Roman transcription by an Italian (Spaniard?) from the late 16th or early 17th century in which prenasal ā is transcribed with “o” in the *bismillah*: [benomi joda meherabon] بنام خدامهريان ‘in the name of God the beneficent’.

In a 1634 travelogue, [9] wrote [Tyroan] for Tehran, while using [aw], [augh] and [a] for /ā/ in other contexts and localities, whether prenasal or not: [spawhawn(e)] ‘Isfahan’, [shaw] and [shaugh] ‘shah’ and [mydan] ‘maidan’.

In an 1852 grammar written in French, [2] wrote that /ā/ was pronounced similarly to *a ô* in *Saône*, which at that time was likely [o:].

7.2. Evidence for prenasal /u/

[2] notes that natives of the province of Fars pronounce /ā/ as [u], and provides examples with and without a nasal environment: /nān/ [nun] نان ‘bread’, /bejā/ [beju] بیا ‘come!’, /māhā/ [muhu] ماها ‘we’.

In 1882, [8] make the same claim about Fars, and then give an initial hint about the situation still in effect today:

In the dialect of Fars all á’s become ú. thus instead of *M é d á n í l*, “Do you know,” they say *M é d ú n í l*. And certain words all over Persia are pronounced colloquially after this incorrect fashion. Such are نان *n ú n*, ‘bread,’ and آن *ú n*, ‘that.’ هان *ham ú n* for *ham á n* is another example of this.

Another historical process that has resulted in [u] in Iranian Persian is the change from the *majhul* /ō/, which is still preserved in Afghanistan

[5]. For example, کودک 'child' is pronounced [kudæk] in Iran and [kodak] in Afghanistan.

8. CONCLUSION

[13] note that the lexical items most commonly undergoing the change to [un] are deictics and pronominal suffixes and that segments in such closed class forms resist changes undergone in other parts of the language.

Such a perspective encourages us to turn the problem around: perhaps the underlying forms contain /u/, and these are optionally changed to [v] in higher registers, influenced by the presence of alef ʾ in the orthography. Indeed, according to [25]:

when a native speaker of Iranian is asked why /xa-neʾ/ must convert to [xu-neʾ] and not remain *[xa-neʾ], the response is that the word is simply [xu-neʾ] and not *[xa-neʾ]...

Based on the available evidence, our hypothesis is that /ā/ was the starting point, and it proceeded through the steps of rounding as [v], and then raising as [o] and finally [u]. The exact circumstances under which [u] arose, whether due to influence by rounding, nasalization or vowel reduction, or some combination, remain to be explored, but the incomplete nature of this change characterizes it as an example of lexical diffusion. The contemporary dialects that show alternation between [v] and [u] likely maintain [v] in higher registers due to the influence of orthography.

9. REFERENCES

- [1] Beddor, P.S., Krakow, R.A., Goldstein, L.M. 1986. Perceptual constraints and phonological change: a study of nasal vowel height. *Phonology Yearbook* 3, 197-217.
- [2] Chodzko, A. 1852. *Grammaire Persane*. Paris: Imprimerie Nationale.
- [3] Crosswhite, K. 2001. *Vowel Reduction in Optimality Theory*. London: Routledge.
- [4] Doostdar, A. 2004. The vulgar spirit of blogging: On language, culture and power in Persian Weblogestan. *American Anthropologist* 106(4), 651-662.
- [5] Farhadi, A. 1955. *Le Persan Parlé en Afghanistan*. Paris: Klincksieck.
- [6] Ferguson, C. 1959. Diglossia. *Word* 15, 325-340.
- [7] Ghobadi, C. 1996. La langue parlée reflète dans les écrits. *Studia Iranica* 25(1), 135-158.
- [8] Haggard, W.H.D., le Strange, G. 1882. *The Vazir of Lankuran*. London: Trübner.
- [9] Herbert, T. 1634. *A Description of the Persian Monarchy*. London.
- [10] Horn, P. 1897. Aus italienischen Bibliotheken. Die persischen und türkischen Handschriften des Vatikans. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 51, 1-65.
- [11] Ivanow, W. 1925. Rustic poetry in the dialect of Khorasan. *Journal and Proceedings of the Asiatic Society of Bengal* 21, 233-313.
- [12] Johnson, K. 1997. *Acoustic and Auditory Phonetics*. Oxford: Blackwell.
- [13] Kahn, M., Berstein, J. 1981. Progress on the uN change in Persian: Dæbestan vs. tabestun. *Acta Orientalia Academiae Scientiarum Hungaricae* 35(1), 133-140.
- [14] Kawasaki, H. 1986. Phonological universals of vowel nasalization. In Ohala, J.J., Jaeger, J.J. (eds.), *Experimental Phonology*. Orlando: Academic Press, 81-103.
- [15] Kieffer Charles, M. 2003. Hazara iv. Hazaragi dialect. *Encyclopedia Iranica Online Edition*. <http://www.iranica.com/articles/hazara-4>
- [16] Labov, W. 1981. Resolving the neogrammarian controversy. *Language* 57(2), 267-308.
- [17] Labov, W. 1994. *Principles of Linguistic Change, Volume 1: Internal Factors*. Oxford: Blackwell.
- [18] Lazard, G. 1957. *Grammaire du Persan Contemporain*. Paris: Klincksieck.
- [19] Lazard, G. 1995. Les emprunts arabes dans la prose persane du Xe au XIIe siècle: Aperçu statistique. In Lazard, G. (ed.), *La Formation de la Langue Persane*. Paris: Peeters, 163-177.
- [20] Majidi, M.R., Ternes, E. 1999. Persian (Farsi). In *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 124-125.
- [21] Miller, C., Moats, N., Strong, R., Browne, E. 2011. Mapping convergence on [u] in several dialects of Persian. *Persian LEARN Conference* San Diego. (To appear.)
- [22] Modaresi-Tehrani, Y. 1978. *A Sociolinguistic Analysis of Modern Persian*. Dissertation, University of Kansas.
- [23] Perry, J. 2003. Persian as a homoglossic language. *Cahiers de Studia Iranica* 27, 11-28.
- [24] Platts, J. 1894. *A Grammar of the Persian Language*. London: Williams and Norgate.
- [25] Rees, D. 2008. *Towards Proto-Persian: An Optimality Theoretic Historical Reconstruction*. Ph.D. dissertation. Georgetown University.
- [26] Sadeqi, A.A. 1984. Tabdil-e "ān" va "ām" be "un" va "um" dar farsi-ye goftari va sabeqe-ye tarixi-ye an [The Conversion of ān and ām to un and um in spoken Persian and its historic background]. *Majalle-ye Zabanshenasi* 1(1), 52-71.
- [27] Sang Tajan, S.S., Bijankhan, M. 2010. Barrasi-e kahesh-e vake-i dar zaban-e farsi-e mahavarei [The study of vowel reduction in Persian spontaneous speech]. *University of Isfahan Researches in Linguistics* 2(1), 35-48.
- [28] Stevens, K. 2000. *Acoustic Phonetics*. Cambridge: MIT Press.
- [29] Zeller, C. 1997. The investigation of a sound change in progress: /æ/ to /e/ in Midwestern American English. *Journal of English Linguistics* 25(2), 142-155.