

Dari genesis: closer to Persian than to Tajik

V.B. Ivanov

Institute for Asian and African Studies
Moscow State University, Moscow, Russia

ABSTRACT

Spectral analysis of Dari vowels portrays long \bar{e}/\bar{o} more closed and higher than short e/o . The situation was just the opposite thousand years ago in New Persian — the ancestor of contemporary Dari, Persian and Tajik. Contemporary Persian-Dari's e/o were short $/i/$ and $/u/$ at that time. Ancient \bar{e}/\bar{o} and \bar{a} are pronounced now in Persian as long $/i/$ and $/u/$ resp. In the Persian-Dari's past short e/o interchanged their positions with long \bar{e}/\bar{o} . Such a rearrangement did not occur in Tajik where short $/i/$ and $/u/$ united with their long neighbors \bar{i}/\bar{u} and \bar{a} . Two different processes (one common for Persian and Dari Vs the other one in Tajik) imply that New Persian was divided into two dialects: one belonging to Khorasan and the other one — to Maverannahr (two historical regions of Middle Asia). Lately Khorasan's dialect diverged into contemporary Persian and Dari while the Maverannahr's one became Tajik.

RELATIVITY OF PERSIAN, DARI AND TAJIK

New (Classical) Persian or *fārsi-ye dari* was a common language spread over the territory of contemporary Iran, Afghanistan and Middle Asia in XII — XV centuries. In the XVI century this linguistic community came to an end [1] and due to geopolitical reasons diverged into three closely related languages — Persian, Dari and Tajik. There is some evidence [2] that a certain difference in pronunciation appeared much earlier — in the XI century $/ā/$ was pronounced like \bar{o}/\bar{u} in the Maverannahr region i.e. contemporary Tajikistan, Uzbekistan etc.). The sequence of appearance of these languages is not discussed in linguistic publications and a naive native

speaker could think either they appeared all at once or still did not diverge at all, being 3 dialects of one language.

The vowel systems of contemporary Persian and Tajik were studied both articulately (by X-rays) and acoustically. But the positions of vowels in Dari vocalism were judged only by hearing. Some linguists suggest that long \bar{e}/\bar{o} (both called *majhul* "unknown" vowels, because they were not known to Arabs) are more open and lower than their short neighbors e/o [3], the other ones confirm just the opposite [4], [5]. An experimental study was necessary to make a well-founded conclusion that would help to compare the development of the three languages.

SPECTRAL ANALYSIS OF DARI VOWELS

The experimentation was based on a well-known concept that two first formants (F1 and F2) are related to the tongue position during vowel articulation. The same technique was used earlier to compare the properties of Russian and Persian vowels in bilingual pronunciation [6]. Despite common opinion that Persian $/a/$ is a front row vowel, our bilingual study proved it to belong to the middle row: Russian $/a/$ in *m'at'* "to crumple" (that was never considered to be a front vowel) is much closer to the front row than Persian $/a/$ in *madd* "tide".

Four Dari native speakers took part in a new experiment, the results of which can be seen in Table 1 and Figure 1.

Table 1. Formant frequencies of Dari vowels (in Hz)

vowels	F1	F2	F3
i	265	2125	3090
e	400	2049	2820

e	420	1875	2675
a	695	1460	2475
ā	560	1085	2080
o	440	1025	1820
ō	410	905	1660
u	280	800	1485

In Table 1 and Figure 1 long \bar{e}/\bar{o} lies between $/i/$ and $e/$. Similarly long \bar{o}/\bar{u} lies between $/u/$ and $o/$. So both *majhul* vowels belong to the upper middle rise and are more narrow than their short neighbors $e/$ and $o/$.

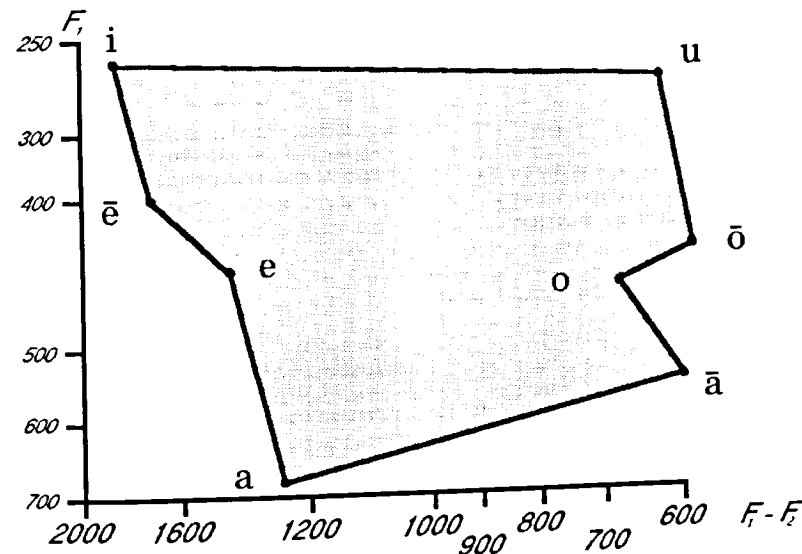


Figure 1. First and second formant positions of Dari vowels.

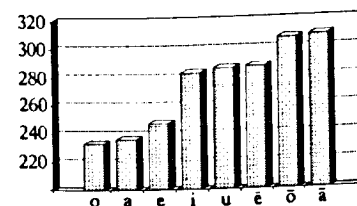


Figure 2. Mean duration of Dari vowels (in ms)

The mean duration, shown on Figure 2, divides the 8 Dari vowels into 3 classes: $e, a, o/$ are short, $i, u/$ — long, $\bar{e}, \bar{a}, \bar{o}/$ — extra long. The extra long nature of $\bar{a}/$ is caused by its

Though it may seem that the F1 difference in pairs \bar{e}/e and \bar{o}/o is not much (only 20–30 Hz) the F1/F2-difference between them on the F1/F2 plane is significant: $p < 0.1$ for \bar{e}/e and $p < 0.03$ for \bar{o}/o . The difference in F3 testifies that the *majhul* vowels are more labialized than the short $e/$ and $o/$. Beside that the vowels differ in duration. Both parameters are significant ($p < 0.03$ and $p < 0.001$ — respectively).

openness: it is the most open vowel in Dari. The more open a vowel is the longer it sounds. Thus the extra long nature of $\bar{a}/$ is a non-phonemic feature. But $\bar{e}/$ and $\bar{o}/$ are very closed and their extra long duration is phonemic. That's why $\bar{e}/$ approximates the long vowels in duration.

Persian script used for official Dari writing does not show short vowels in most cases and does not distinguish i/\bar{e} and u/\bar{o} alternatives. It brings us to a unique situation in Dari not found in Tajik or Persian: Dari native speakers' identify vowels with difficulty. The vowels in triplets $e, i, \bar{e}/$ and $o,$

u, *ō*/ can be interchanged depending on the speech style. In the official one the speakers try to use extra long vowels /*ē*, *ō*/ even if there is no historical ground for it, like in *ārōs* "bride" (it is an Arabic word and must be free of majhul vowels). The same word can be pronounced *arus* in less official cases of literary language and *āros* in colloquial. The overall tendency in contemporary Dari is to substitute long and extra long literary vowels by corresponding short ones in colloquial speech: *sotun* > *soton* "column", *budan* > *bodan* "to be", *nōzdah* > *nozda* "nineteen", *āwāz* > *awāz* "song", *āina* > *ayna* "mirror", *mēzanam* > *mezanom* "I strike" [7].

REARRANGEMENT OF EXTRA LONG AND SHORT VOWELS AFTER CLASSICAL PERIOD

In Ancient Persian there were 2 pairs of vowels. Inside each pair the vowels differed in phonological length: /*i*, *ī*/, /*ū*, *ū̄*/, and /*ā*, *ā̄*/. and 2 diphthongs /*ai*/ and /*au*/. Those 2 diphthongs were the only diphthongs possible at that time: they were made by tongue movement from lower middle position towards extreme front or back. Such movements historically precede establishment of other diphthongs like /*ui*/ because the latter is formed across a catastrophic boundary which is a more complicated movement. Catastrophic diphthongs appear after the time the more probable non-catastrophic ones are already in use.

Later diphthongs /*ai*, *au*/ turned into monophthongs /*ē*, *ō*/ resp. [8] (Figure 3). /*i*, *ī*/ and /*ū*, *ū̄*/ were articulated similarly, but /*ā*, *ā̄*/ were different even then: /*ā̄*/ was closer to back row vs. more front /*a*/. We can state it more precise that short /*i*/ and /*u*/ were non-significantly more centralized than their long counterparts, because generally it is difficult for the speaker to move the tongue during short period of time to an extreme front or back position. The central-

ization of /*i*, *u*/ allowed them later in Persian and Dari to reach the state of /*e*, *o*/ resp. It explains why it was just the short /*i*, *u*/ who did it but not the long neighboring vowels. The state of contemporary languages proves this hypotheses [9].

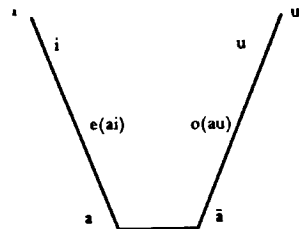


Figure 3. Vowel system of late Ancient, Middle and New Persian

After the classical period (about a thousand years ago) the previous system with 6 monophthongs and 2 diphthongs tended to be simplified. The diphthongs turned into extra long /*ē*, *ō*/. It increased the number of rises: instead of one level of the middle rise two of them appeared. The Classical Persian began to branch.

In the Khorasan branch /*ē*, *ō*/ went to the upper part of the middle rise, in the Maverannahr branch — to the lower one. Later the Khorasan branch was divided into Persian and Dari. The situation in Dari which is well-known for its archaic elements remained just as it was in Khorasan dialect of the Classical Persian. Western dialects — Tehran's and Isfahan's Persian — developed further: narrow /*ē*, *ō*/ lost their phonological difference from /*i*, *u*/ resp. This process is still going on in Herat dialect which is an intermediate one between Persian and Dari [10].

Table 2. Front row vowels relativity. Underlined words contain vowels tending to higher rise

Persian	Dari	New Persian	Tajik	meaning
<i>bist</i>	<i>bist</i>	<i>blst</i>	<i>bist</i>	twenty

<i>bim</i>	<i>bim</i>	<i>blm</i>	<i>bim</i>	fear
<i>xeš</i>	<i>xeš</i>	<i>xiš</i>	<i>xiš</i>	brick
<i>emruz</i>	<i>emrōz</i>	<i>imrōz</i>	<i>imrūz</i>	today
<i>mīš</i>	<i>mēš</i>	<i>mēš</i>	<i>mēš</i>	sheep
<i>riš</i>	<i>rēš</i>	<i>rēš</i>	<i>rēš</i>	wound

Table 3. Back row vowels relativity. Underlined words contain vowels tending to higher rise

Persian	Dari	New Persian	Tajik	meaning
<i>dur</i>	<i>dur</i>	<i>dūr</i>	<i>dur</i>	far
<i>dud</i>	<i>dud</i>	<i>dūd</i>	<i>dud</i>	smoke
<i>sorx</i>	<i>sorx</i>	<i>surx</i>	<i>surx</i>	red
<i>xošk</i>	<i>xošk</i>	<i>xušk</i>	<i>xušk</i>	dry
<i>ruz</i>	<i>rōz</i>	<i>rōz</i>	<i>rūz</i>	day
<i>guš</i>	<i>gōš</i>	<i>gōš</i>	<i>gūš</i>	ear

In the Maverannahr branch /*ē*, *ō*/ went to the lower part of the middle rise. It caused more narrow pronunciation of the short /*i*, *u*/ that finally in contemporary Tajik and Hazara dialect in Afghanistan merged with long /*i*, *u*/ resp. Tajik became a center row vowel, /*ā̄*/ went up to /*ō̄*/ causing former majhul /*ō̄*/ to be centralized /*ū̄*/. In both branches the upper part of middle rise in back row was unstable and disappeared.

CONCLUSION

Some common features in Tajik and Dari like final /*a*/ that is not characteristic of Persian (Persian *xāne* ~ Dari *xāna* ~ Tajik *xona* "house") lead to the conclusion that the distance between Dari and Tajik was less than between Dari and Persian. But those differences and similarities (especially the tendency to pronounce /*e*/ instead of /*a*/ in Tehran and Isfahan in quite a number of positions) are product of later development. Global position-independent tendencies to mix up the majhul vowels with the long ones described above could not have been implanted into two neighboring languages by chance. Thus Dari and Persian should be considered closer relatives than Dari and Tajik.

REFERENCES

- [1] Yefimov, V., Rastorgueva, V., Sharova E (1982). "Persidskiy, tajikskiy, dari", *Osnovy iranskogo yazykoznanija, Novoiranskiye Yazyki*, Moscow: Nauka, p.7.
- [2] Edelman, J. (1968), *Osnovnye voprosy lingvisticheskoy geografii (na materiale indoevropejskikh yazykov)*, Moscow.
- [3] Yefimov et al. Op. cit., p.26—27.
- [4] Ostrovsky B. (1994) *Ucebnyk yazyka dari. Čast 1*. Moscow, Nauka, p.19.
- [5] Kiseleva L. (1985) *Yazyk dari Afganistana*. Moscow. pp.21—22.
- [6] Ivanov, V. (1983) *Russkiye i persidskiye glasnye v proiznoshenii bilingvov*. Moscow, *Vestnik MGU. ser. Vostokovedeniye*, #2.
- [7] Farhadi R. (1974) *Razgovorny farsi v Afganistane*. Moscow: Nauka. pp. 18—25.
- [8] Edelman J. (1975) "Evolutia fonologicheskogo tipa". *Opyt istoriko-tipologicheskogo issledovaniya Iranskikh yazykov*. Moscow: Nauka, v.1, p.29
- [9] Petito-Cocorda J. (1985) *Les catastrophes de la parole de Roman Jakobson a Rene Thom*. Paris: Maloine, p.283
- [10] Ionesian Y. (1987) *Dialecte sovremennogo Dari rayona g. Herat (phonetica, morphologia)*. Ph.D. Dissertation, Moscow.