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 /é/ and /ô/ more closed and higher than short /e/ and /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 /é/ and /ô/ were short /ê/ and /ô/ at that time. Ancient /ê/ and /ô/ are pronounced now in Persian as long /ê/ and /ô/ resp. In the Persian-Dari's past short /ê/ and /ô/ interchanged their positions with long /ê/ and /ô/. Such a rearrangement did not occur in Tajik where short /ê/ and /ô/ united with their long neighbors /ê/ and /ô/. 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 farsi-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 /a/ was pronounced like /a/ 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 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 /ê/ and /ô/ (both called majhul "unknown" vowels, because they were not known to Arabs) are more open and lower than their short neighbors /ê/ and /ô/ [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 "ridge".

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)

<table>
<thead>
<tr>
<th>vowels</th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>625</td>
<td>2125</td>
<td>3090</td>
</tr>
<tr>
<td>e</td>
<td>400</td>
<td>2049</td>
<td>2820</td>
</tr>
</tbody>
</table>

Though it may seem that the F1 difference in pairs /ê/e and /ô/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 /ê/e and p<0.03 for /ô/o. The difference in F3 testifies that the majhul vowels are more labialized than the short /ê/e and /ô/o. Beside that the vowels differ in duration. Both parameters are significant (p<0.03 and p<0.001 — respectively).

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, /ô/ long, /ê, ô, a, ò/ — extra long. The extra long nature of /ô/ is caused by its 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 /ô/ is a non-phonemic feature. But /ê/ and /ô/ are very closed and their extra long duration is phonemic. That's why /ê/ 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 /ê/ and /ô/ 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 /ê, i, ê/ and /ô,
u, ū/ can be interchanged depending on the speech style. In the official one the speakers try to use extra long vowels /e, ū/ even if there is no historical ground for it, like in arūs “bride” (it is an Arabic word and must be free of majhul vowels). The same word can be pronounced ants in colloquial. The overall tendency in contemporary Dari is to substitute long and extra long literary vowels by corresponding short ones in colloquial speech: satun > soton “six”, nodah > noda “nineteen”, awāz > awāz “song”, āina > āyna “mirror”, mēzanam > mezanom “I strike” [7].

### Long and Short Vowels

**After Classical Period**

In the Khorasan branch /e, ū/ went to the lower part of the middle rise. It caused even 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. /a/ went up to /ā/ causing former majhul /ā/ to be centralized /ā/. 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


