# NOTES ON THE ARABIC DIALECT SPOKEN IN THE BALKH REGION OF AFGHANISTAN

RV

#### ABDUL-SATTĀR SĪRAT Kabul, Afghanistan.

#### Annotated by

## EBBE EGEDE KNUDSEN University of Oslo.

In 1926 it became known through a Soviet census that there were 27.977 Arabs in Uzbekistan. In a book published the same year Professor Georg Morgenstierne quoted an administrative authority in Afghanistan stating the presence of an Arabic speaking community in that country. The discovery of the Arabic dialects of Uzbekistan was a scholarly sensation and the discovery of another dialect in Afghanistan was no less so. The first specimen of Soviet Central Asian Arabic appeared in print in 1930. However, it was only in the last years before World War II and in the post-war period that real progress in the study of the Soviet dialects was achieved through the writings of G. V. Cereteli and I. N. Vinnikov. From the pen of the latter we have even a dictionary. In 1961 Wolfdietrich Fischer summed up our knowledge of Soviet Central Asian Arabic. There are two clusters of dialects.

a For a reference, see the contribution of Abdul-Sattar Sirat below note 1.

b N. Burykina and M. Izmailova, La langue des arabes des kichlaks d'Uzbekistan Jugara et Jeinau (in Russian), reviewed by H. S. Nyberg, Le Monde Oriental 24 (1930), 121-126. The population figure as quoted above is given by Nyberg p. 121.

c Slovar' dialekta buxarskix arabov, Paleslinskij Sbornik 10 (Moscow 1962).

d Die Sprache der arabischen Sprachinsel in Uzbekistan, *Der Islam* 36 (1961), 232–263. References to earlier literature, mainly in Russian and Georgian, are given by Fischer p. 232 and by Haim Blanc in *Current Trends in Linguistics I* (ed. Thomas A. Sebeok, The Hague 1963), p. 38142. Fischer's 1960 paper at the Congress of

Genetically there is an apparent relation to the dialects of the Mesopotamian area.e

Research on Afghanistan Arabic has been much less elaborate. According to Dr. Ravan Farhadi Arabic is spoken in the village of Zâri, about 100 km south of Balkh in the region of Keshen-Deh, by a semi-nomadic population of about 5.000 persons. Some time ago the manuscript printed below was handed over to Professor Georg Morgenstierne to consider it for publication. It is the work of a young Afghan Arabist and religious historian with no particular background in modern Arabic and General Linguistics. Professor Morgenstierne and I agreed that the manuscript was an important contribution to Arabic Studies, however preliminary and fragmentary it was. It contained the first specimen of Afghanistan Arabic to be treated in a Western language. So we recommended it for publication in the 'Acta Orientalia'. We are aware of certain discrepancies in the rendering of Arabic words, but we decided to have these words printed exactly as they appeared in the manuscript. However, the system of transcription was changed on a few points so as to conform with accepted usage. Due to the distance between Scandinavia and Afghanistan the annotator took over the responsibilities of seeing the manuscript through the press.

The new material reveals a language that is very closely related to Soviet Central Asian Arabic. Even a superficial comparison of the dialects confirms this in itself not unexpected view. In terms of classification we may state that in its present forms Central Asian Arabic constitutes a clearly defined dialect group. The characteristic features in this group may be summed up as follows:

- (a) All adult male speakers are at least bilingual Arabic Tajiki.
- (b) For the northern dialects Fischer considered the possibility of classifying the vowels i and u as one phoneme (p. 235 f.). The

Orientalists in Moscow Die Stellung des Arabischen Zentralasiens innerhalb der neuarabischen Dialekle seems to have remained unpublished, cf. Trudy 25. meždunarodnogo kongressa voslokovedov II (Moscow 1963), p. 136. For a narrative account of Soviet research on the Central Asian dialects, see I. J. Kratschkowski, Die russische Arabistik (Leipzig 1957), pp. 236-238 and 252-254.

e See the references given by Blanc p. 381 f.43.

f For a reference, see the contribution of Abdul-Sattar Sirat below note 2.

southern material as described by Mr. Sirat is hardly conclusive in this respect. Apparently the author often found it difficult to decide on a transcription *i* or *u*. In his manuscript he corrected cases of *izn* 'ear' and *izrab* 'beat (imp.)' into *uzn* and *uzrub* respectively. The imperfect *muydī* '(I) go' occurs along with *maydī* with the same meaning, cp. also the morphologic (morphophonemic) variants *mektib* '(I) write' and *mašūf* '(I) see'.

- (c) A long vowel  $\bar{u}$  occurs as a secondary reflex of Old Arabic  $\bar{a}$ , the primary reflex being  $\bar{a}$ : Northern  $m\bar{u}$ ,  $m\bar{a}y$ , dialect of Jeinau  $m\bar{u}ya$  (cf. Fischer p. 235), as against southern  $m\bar{u}ya$  'water'. The Egyptian word for 'water' phonemic /mayya/ and a large number of its cognates in other dialects such as Baghdadi /mayy/, /may/ have emphasis throughout. Historically emphasis in these forms probably results from the loss of a following w in the proto-form \*mwayya.8 Alternatively the Central Asian forms may have a similar origin.
- (d) In general ancient diphthongs have become monophthongs, though with varying realizations: N.  $y\bar{u}m$ , S.  $y\bar{o}m$  'day' < yaum, similarly N.  $t\bar{e}ra$  'hen', S.  $t\bar{e}ra$  'bird' < taira and taira 'thing' < taira old Arabic taira 'and taira 'but cf. S. taira 'garment', Old Arabic taira 'garment', Old Arabic taira 'garment', Dr. Ravan Farhadi suggests to me that items with apparently preserved diphthongs represent borrowings from Tajiki as taira 'good' and taira 'people, tribe'. If shortened the monophthongs taira is become taira as in taira 'what'.
- (e) The emphatic consonants tend to lose their secondary 'emphatic' articulation.
- (f) A number of Old Arabic phonemes show an interesting double representation in the dialect group. From this evidence Fischer inferred that two types of dialect were represented in the original

g Cf. Haim Blanc, Studies in North Palestinian Arabic (Jerusalem 1953), p. 55, the same author's Communal Dialects in Baghdad (Cambridge, Mass. 1964), p. 18f., Glselher Schreiber, Der arabische Dialekt von Mekka (Inaugural-Dissertation 1970), p. 12f., and the references given in Nada Tomiche, Le parler arabe du Caire (Paris 1964), p. 19<sup>1</sup>. Professor F. Løkkegaard of the University of Copenhagen calls my attention to a note in Wilhelm Spitta-Bey, Grammalik des arabischen Vulgärdialectes von Aegypten (Leipzig 1880), p. 99 stating the same origin (Medleval Arabic muwayya) for the Egyptian word.

population, one bedouin type of dialect and one sedentary type of dialect (p. 238). The phonemes showing this double representation are q,  $\delta$ ,  $\theta$ ,  $\delta$ , the last a merger of two Old Arabic phonemes, a voiced emphatic interdental spirant  $\delta$  and a voiced emphatic interdental lateral d. In the bedouin type of dialect q shifted to g as elsewhere in bedouin dialects and under the influence of Tajiki  $\delta$ ,  $\theta$ ,  $\delta$  became z, s, z respectively. In the sedentary type of dialect this set of phonemes was represented by q, d, t, d respectively. Except for the treatment of emphasis this is paralleled by other dialects of the sedentary type.

- (g) The Arabic definite article has disappeared. Reflexes survive in a few lexical forms as  $balb\bar{e}t$  'door'  $< b\bar{a}bilb\bar{e}t$ .
- (h) The use of the dual has been preserved.
- (i) Under the influence of Tajiki the position of the verb is usually at the end of the sentence.
- (j) Formation of verbs mediae geminatae on the analogy of verbs tertiae infirmae: N. šaddā, S. šidda 'he tied'.
- (k) The imperfect form of the verb (imperfective aspect) has borrowed the Persian-Tajiki prefix  $m\bar{\imath}$ -,  $m\bar{e}$ -. This prefix probably replaced the Common Arabic (Old Koiné) prefix bi- (Fischer p. 247³). In the northern dialects the old imperfect continued the old subjunctive in subordinate clauses. In main sentences it acquired modal function. This tendency in the development of the old imperfect is paralleled by other dialects of Modern Arabic, but it is not limited to Arabic. It is a characteristic feature in the history of the Semitic verb in general. The old imperfect or a replacement of it took over similar functions in Mishnaic and Modern Hebrew and in modern and earlier phases of Aramic and Ethiopic. Due to the lack of a real context the southern material is not quite conclusive on this point.

h The linguistic shift referred to is the one termed by Frithiof Rundgren "die dritte, westsemitische Aspektverschiebung", see his Erneuerung des Verbalaspektes im Semitischen, Acta Universitatis Upsaliensis Nova Series 1:3 (Uppsala 1963), pp. 66-71 (Ethiopic), 72 (Syriac), and in general 98-100, and the same author's Das altsyrische Verbalsystem, Spräkvelenskapliga Sällskapels i Uppsala Förhandlingar 1958-1960 (Uppsala 1960), p. 68-71.

Dr. Ravan Farhadi informs me that he is presently supervising Qāri Aynī, a professor of the Asadiyya Medressa in Mazari-Sharif, who is preparing a monograph in Persian on "the language and customs of the Balkh Arabs".

EBBE EGEDE KNUDSEN

Before 1926 there was no available information in the writings of Western linguists and geographers regarding the existence of Arabic as a spoken language in Afghanistan. In that year Professor G. Morgenstierne in his Report on a Linguistic Mission to Afghanistan1 drew the attention to an Arabic speaking minority in quoting the governor of Mazar-i-Sharif "that some nomadic tribes west of that city still speak Arabic". In 1957 the Afghan delegation to the 24th International Congress of Orientalists in Munich submitted a report on the languages of Afghanistan. The report included a short statement on the presence of an Arabic-speaking community and its geographical location.2 The present writer visited the Arabic-speaking villages of the country in 1960 and published an article on the subject in 1961 in the spring issue of the quarterly journal Adab.3 Almost a decade later, in the spring of 1970, further material appeared in a report on a visit to Khoshhal-Abad undertaken by Dr. Ravan Farhadi, honorary member of the Afghan Linguistic Institute. Dr. Farhadi had already encouraged and assisted me in my 1960 study of the subject. The present article is based on all studies on the dialect made so far.

An Arabic speaking community is found scattered in a few areas of the provinces of Balkh and Jozjan. It is concentrated particularly in the following villages to the west of the historic town of Balkh, consisting totally of about 600 families:

- Village Khoshhāl-Ābād situated to the west of Daulatābād town in the Balkh province, 120 families.
- Village Yakhdān near Khoshhāl-Ābād, 100 families.

<sup>&</sup>lt;sup>1</sup> Instituttet for sammenlignende kulturforskning, Serie CI-2 (Oslo 1926), p. 6.

<sup>&</sup>lt;sup>2</sup> 'Abd-ul-Ghafur Ravān Farhādi, Akten des 24. internationalen Orientalisten-Kongresses München (1959) p. 446.

<sup>3</sup> Colloquial Arabic in the Balkh Region, Adab 10/1, p. 1-11 (in Persian).

- 3. Village Sultān-Arēgh situated in the district of Manga-Jak to the west of the city of Aqcha in Jozjān province, 200 families.
- 4. Village Hasan-Ābād situated to the north-east of Shibirghān, the provincial capital of Jozjān, 150 families.

The location of these four villages is such that the inhabitants are in mutual contact. They seem to have been allotted land collectively at a specific locality. The inhabitants of these villages are, therefore, Arabs exclusively. Arabic-speaking villagers and craftsmen have also settled on the other side of the Oxus River in Soviet Uzbekistan. For several years already Soviet scholars have been doing research on the dialects of these Arabic-speaking groups. A comparative study of the spoken Arabic of both sides of the Oxus River is now becoming possible.

A detailed investigation of the history of these Arabs cannot be made at present, since there is no information as to when and from where they have come. However, there is a tradition among the people of the northern areas that they came to Central Asia from Syria during the campaigns of Tamerlane (end of 14th century). Further some of the older people in these villages say that they originally came from Yemen.

The Arabic-speaking villages are situated in areas that are important centres for the cultivation of cotton, wheat, and barley. They are also engaged in a considerable trade in carpets, hides, and skins. The Afghanistan Arabs belong to the Hanafi School of the Sunni Muslims. They have developed no characteristic music. They do, however, sing or narrate their own traditions and stories in their own dialect.

The dress of the women is very long. Even the čādar worn on the head consists of one piece of cloth which is about seven metres long. The young ladies knot their hair with wool in the shape of a chain, cover their head with a piece of cloth and emboss their chest, back, and belt with coins. The dress of the men is similar to that of the Tajiks and Uzbeks.

As regards mourning after death, close female relations of the deceased wear special mourning dress, let loose their hair, and circle round the dead chanting a dirge in Arabic. These mourning ceremonies may last up till one year. As regards marriage Arab

customs have mostly disappeared. It is only in the village of Sultan Aregh that there are still a few old women who dance and sing Arabic songs on the night of the marriage.

#### Notes on Phonology

In describing the phonology of the dialect it is important to distinguish old inherited features from those influenced by Dari-Tajiki (Persian) and Uzbek. A large number of features are old inherited and well-known from Literary Arabic and from the spoken dialects all over the Arab world. Compare the following pairs of words:

'ēn	'eye'	Literary	Arabic	<sup>c</sup> ayn
bēt	'house'	_		bayt
šī	'thing'	_	_	šay', and the verbs:
tilla	'you are comi	ng out'		J
tiqrī	'you are readi	ing'		
nizhab	'we are going'			

The word riz 'rice' may be compared with literary ruzz, aruzz. They also use the word  $d\bar{e}n$  'oil' a cognate of Literary Arabic duhn. Further note the following:

iš maḥal 'what time', mūya 'water', ḥabḥab 'water melon', γadā 'he went', ḥinta 'wheat', and jēnā 'we came'.

Common Arabic h and 'are preserved. This is an important feature considering the fact that Tajiki and Uzbek do not preserve these in Arabic loan-words. Compare the following list of examples:

<b>ḥab</b> ḥab	water melon	da <sup>c</sup> mit	go away
niḥna	we	ʻanza	goat
wāḥid	one	na'ja	sheep
Ņimār	donkey	`indak	with you
ḥinta	wheat	ʻinab	grapes
ķilla	open (imp.)	ta <sup>c</sup> a	come
<u> ļ</u> āyit	wall	tiţlaʻ	you come out
		nir <sup>c</sup> a	we pasture
		nizra	we cultivate
		ša 'r	hair

The voiced affricate represented in writing by the letter  $j\bar{\imath}m$  is generally preserved. Classical Arabic  $\not q$  ( $\not z$ ),  $\not q$  ( $\not q$ ), and  $\not q$  merge with z as can be seen from words like

```
uzn 'ear' for Classical udn-
uzrub 'beat' – – idrib
zahr 'back' – – dahr-
```

However, there are a few words with d for ancient d as ard 'land',  $ardan\ nizra$ ' 'we are cultivating land'. At the end of a word Classical d becomes d as in  $\bar{a}xdo$  'take (it)', and  $n\bar{a}xud$  'we take'.

The sounds which have not been preserved include the hamza or glottal stop '. Thus the dialect word  $\hbar \bar{a} y e t$  'wall' corresponds to Literary Arabic  $\hbar \bar{a}' i t$ . In the position as last consonant of a word hamza disappears as in  $tiqr\bar{\imath}$  corresponding to Literary taqra' 'you read'.

Common Arabic t generally shifts to t as in batn 'abdomen' and  $h\bar{a}yet$  'wall', whereas t and t merge with t as in t Literary Arabic t and t basal 'onion', the Literary t basal.

The treatment of Common Arabic q is interesting. In some words it is preserved as q, while in other words it becomes g. Thus Literary Arabic qidr 'pot' may be compared to the word gidr with the same meaning. Similarly for Literary  $daq\bar{q}q$  'flour' they say  $dag\bar{q}g$ . Sometimes q alternates with '. Instead of  $daq\bar{q}q$  one may also hear  $da'\bar{q}g$ . The treatment of Common Arabic q as q may be either old inherited or due to influence from Tajiki or Uzbek, both of which preserve Arabic q in loan-words. Note the following cases of preservation of q: qarya 'village', baqar 'cow', and raqaba 'neck'.

In  $d\tilde{e}n$  'oil' for expected dihn the treatment of the sequence -ih- as  $-\tilde{e}$ - has a parallel in Tajiki. The following reductions in the phonetic shape of words should also be noted:

```
mālnā kasir 'our possessions are abundant' (for kasīr)
ta'a 'come in!' (for ta'āl)
balbēt šidda 'close the door' (for bab-al-bēt)
```

#### **Notes on Semantics**

The necessary contact with the Tajiki-speaking people of the area has influenced the language of the Afghanistan Arabs. Cases of borrowing from neighbouring languages are common. The numerals are a case in point. While numerals one to six are counted in Arabic, numerals from seven onwards are counted in Tajiki:

wāhid one sit six tnēn two haft seven salās three hašt eight arabaʻ four no(h) nine xams five da(h) ten yãzda(h) eleven, etc.

However, instead of the Tajiki numerals yak sad '100', do sad '200', sē sad '300' they use the Arabic expressions fad yalīza, yalīzatēn, and salas yalīza respectively. They have adopted Tajiki or Uzbek words for items not known to them before their establishment in the area, as for example čāynak 'teapot', piyāla 'cup', jašn 'festival, celebration', bāzār 'market', tarkāri 'vegetables' etc. The dialect has adopted a few more words, in part Arabic words with new meanings, like māldār 'stock-breeder', dēqān 'peasant', yarībī 'one working away from the village', xuš 'happy', aḥwāl pursi 'asking about one's health', amū, amān 'that, same', barābar 'equal, equivalent', hēč 'none', agar 'if', čūnke 'because', hēč waxt 'any time', mūy safēd 'elder, lit. white-haired'. There are Arabic words in use for all matters relating to family or village life. The words for the parts of the body are all Arabic in origin:

izn	ear	raqaba	neck
rās	head	kitaf	shoulder
⟨ēn	eye	zahr	back
anf	nose	batn	abdomen
xišm	mouth	surra	navel
sin	tooth	yid	hand
ša r	hair	rijl	leg, foot

Similarly they mostly use Arabic words for whatever cattle or other animals they keep. Compare the following list:

<sup>7</sup> Acta Orientalia, XXXV

kayiš	horse	tēra	bird
faras	horse	ʻanza	goat
, bagar	bull	mirza	calf
himār	donkey	na <sup>c</sup> aja	$\mathbf{sheep}$
kalb	dog	xarūf	lamb
	•	ba <sup>c</sup> īr	camel

A few other words of daily use are:

dēna	oil	<i>ķinta</i>	wheat
basal	onion	yarab	tree
riz	rice	nufawwar	kettle
		gidr	pot, saucepan

Minor differences in idiomatic usage exist within the dialect. These differences vary from village to village, between young and old, etc. I asked an elderly Arab in Daulat-Abad how many Arab villages they were. He replied: nahn arba'a qurā 'we are four villages'. However, when I put the same question to an illiterate young man at Shibarghan, he replied: nahn araba'a qišlāq ast. The word  $qi\bar{s}l\bar{a}q$  is Uzbek  $q\ddot{s}l\dot{a}q$  'village, settlement'. The speakers themselves deny the existence of any such differences. I was told by the elders of these four villages that there was no difference whatsoever in the speech of the villages. They also maintain that the dialect used by them at present is not very different from the one used by their ancestors. However, they are aware of the existence of differences in the speech of men and women. The women are believed to speak better Arabic than the men. The reason for this would be the general lack of contact between the Arabic-speaking women and the women of the non-Arab communities of the country. All male members of the Afghanistan Arab community know Tajiki and in general Uzbek, in addition to Arabic. The children are monolingual and do not get into contact with other languages than Arabic as long as they do not enter primary school.

Sometimes the adoption of a non-Arabic word affects certain changes in the morphological form of the word. Thus the Tajiki word pagāh 'morning' appears as yalpagā (for expected al-pagā). Tajiki māldārī 'stock-breeding' geminates its last consonant and becomes māldāriyya in Arabic.

Native speakers of Arabic recite the Qur'an and ritual prayers with a Tajiki accent, apparently because traditional recitation has been transmitted through non-Arabs. However strange it may seem, there is, therefore, no direct connection between Quranic recitation and the pronunciation of the Arabic dialect of the area.

### Notes on Morphology and Syntax

From the viewpoint of morphology and syntax the Tajiki and Uzbek languages have greatly influenced the colloquial Arabic of the Balkh region. The language has not retained the distinction between masculine and feminine gender in substantives and verbs. This feature is paralleled by both Tajiki and Uzbek. Further there is no distinction in the expressions for the third person of the past tense and the imperative as in yūgad, 'he made a fire' or 'make a fire'. As in Tajiki and Uzbek the distinction between singular and plural is observed, compare for example ozrub 'beat' and ozrubū 'beat (pl.)'. Similarly the pronouns for the second person anta and antum express the singular and the plural respectively. Another interesting feature in the morphology is the preservation of the dual as in yidēn 'two hands', 'ēnēn 'two eyes', rijlēn 'two legs', qaryatēn 'two villages'.

The dialect has adopted the Persian verbal prefix me-. Compare the following list of forms:

muqul	I say	muqūlūn	they say
muydī	I go	muydün	they go
myakil	I eat	mya'kilūn	they eat
mektib	I write	mektibūn.	they write
ma'rif	I know	me 'arifūn	they know
me <sup>c</sup> ayūn	I see		thoy know
mašūf	I see	1 1. 1. 1.	1

A case of a calque is the Persian gap bizan 'speak up!' rendered verbatim as gap uzrub. The dialect has taken over the word gap and translated bizan as the Arabic imperative uzrub. Similarly Persian pēš-i-kī mē-xānī? 'with whom are you studying?' corresponds to Arabic qudām man tiqrī? Some words represent a combination of an Arabic and a Tajiki element with similar

meaning. As an example one may quote the word  $m\bar{a}n\bar{e}$  'no, nothing'. The word is a compound of two negatives, the Arabic  $m\bar{a}$  and the Tajiki  $n\bar{e}$ .

An interesting case of a syntactic borrowing is the fact that the substantive precedes the verb. This feature is contrary to Classical and Literary Arabic usage and may have been adopted under Tajiki influence. Thus the Arabic phrase xubz myākul 'he eats bread' parallels Tajiki nān mēxura. Note also the following phrases and sentences:

na'ia nir'a we pasture sheep zar'a nizra' we cultivate land (lit. agriculture) jabal muydi we go to the mountain ša'r ra's na'xuz we cut (lit. take) hair of the head hinta jib bring wheat 'indak mirza šidda tie the calf near you 'indak mirza hilla release (lit. open) the calf near you zāsir āxod cut your nail kayyeš jurra ta'a drag the horse and come iš maḥal min bēt titla'? when will you leave the house? bēt ruḥnā we went home iš maḥal bāzār tuydī? when do you go to the bazaar? ba'd iš maḥal tijī? when (lit. after how long) do you leave? 'indak šī da'mit jī you have something, go and bring it yalpagāh nagām we rise in the morning *iš dars tiqrī*? what lesson do you read? maktab tiqrī you read in school koll yöm tuydi you go every day maktab maydi I am going to school maktab si čandum sinf tiqrī? in which class do you read? sinf čārum baqri I read in the fourth class nahn jēna we came kitabātak tiqdir tiqrī you can read your books yom jom'a maktab lā maydī on friday I do not go to school 'arabiyya niqrī we read Arabic qarētna ismā xošāl-ābād the name of our village is Khoshal-Abad ana lihtu-biz qarya I am the white-bearded of the village si hā qarya qalīn mā yinsif carpets are not woven in this village

liwaladnā for our children dars nagūl we read lessons

nahn zar'a zar'an ni'allimum we are cultivators and teach them cultivation

dwāzda walad xamsa minum le xafin nidi'yūm twelve children, five of them we have assigned (lit. put) (to be taught) writing ismkum šuwa what is your name?

iš fādak mirīd? what do you like (lit. does your heart want)? qalbi quttiš māyrid I do not (lit. my heart does not) want anything bētkom qarib bētnā your house is near to ours

anta agar yaylō taʻ tijūn darb malqa if you come, we will find the gate

(For the Soviet dialects, see also G. V. Tsereteli: 'The Influence of the Tajik Language on the Vocalism of Central Asian Arabic Dialects', BSOAS 33 (1970), 167-70 and the bibliography given on p. 167<sup>8</sup>. A rejoinder to Fischer's article is William Cowan: 'Sound Change in Central Asian Arabic', Islam 43 (1967) pp. 134-8).

#### Section 1 Breaking

A constitution of the constitution of