

# **Linguistic Convergence and Areal Diffusion**

Case studies from Iranian,  
Semitic and Turkic

Edited by

Éva Ágnes Csató, Bo Isaksson  
and Carina Jahani

# Linguistic Contacts in Central Asia

Guram Chikovani

## Introduction

Peripheral Arabic dialects manifest various linguistic peculiarities that are mainly due to close linguistic contacts with Indo-European, Turkic and other languages. They exhibit rich materials for the study of problems of historical and general linguistics.

The foundation for the study of Central Asian dialects was laid in the 1930s by G. Tsereteli and I. Vinnikov. The Leningrad ethnographers N.N. Burikina and M.M. Izmailova were the first scholars to publish information about the Arabic-speaking population of Central Asia. I. Kraxkovskij regarded the discovery of Arabic dialects in Central Asia as the beginning of a revolution in Arabic dialectology. N. Jušmanov, who gave a course in Arabic dialectology at Leningrad University in the 1930s, also attached great importance to the discovery and study of Arabic dialects in Central Asia. As a result of his first scholarly expedition in 1935, Tsereteli demonstrated the existence of two different Arabic dialects in Central Asia, the Bukhara and Qashqa-darya dialects. He noted that the two dialects differed considerably from each other.

The Bukhara and Qashqa-darya dialects have been separated from the rest of the Arabic speaking world for many centuries and this has been an important factor in the development of the Central Asian Arabic dialects. At the same time they have been in contact with the Indo-European languages Tajik and Dari, as well as with the Turkic languages Uzbek and Turkmen. This has favored the development of multilingualism.

Historical, geographic, social and other extralinguistic factors have played an important role in the development of Central Asian Arabic dialects and are taken into account in the program of research into these dialects. Linguistic studies have focused on phonetic, morphological, syntactic and lexical features in the speech of trilingual individuals resulting from the influence of the contact languages. Such influences have led to convergence with the structurally different Indo-European and Turkic languages. The contacts can be viewed as a long process in which quantitative changes have, at certain stages, led to qualitative changes.

Today most Arabic speakers of the Bukhara and Qashqa-darya regions make wide use of Tajik and Uzbek. Bukhara Arabs are more fluent in Tajik, whereas Qashqa-darya Arabs are more fluent in Uzbek. The current linguistic

situation is marked by the existence of coordinate and subordinate bi- or trilingualism. In the case of subordination only the native language is spoken fluently. The level of knowledge of Tajik and Uzbek is low, and the speaker tries to subordinate the foreign languages to the norms of his native Arabic. These speakers are mainly elderly people and women limited to the family domain. In certain cases young people speak Tajik and Uzbek more fluently than their native Arabic, using numerous foreign words and expressions in everyday conversations with people in their own villages. This phenomenon can be explained by modern tendencies in the Uzbek Republic. Monolingualism is not uncommon in the Arabic-speaking qishlaqs of Bukhara and Qashqa-darya. Monolinguals are mostly very old people. It is especially difficult to obtain dialectological materials from monolingual women, due to their traditional way of life.

## Features of the Qashqa-darya dialect

The Qashqa-darya dialect displays peculiarities that are both similar to and different from those of literary Arabic and other Arabic dialects. In the following, some features resulting from contact with Uzbek and Tajik will be discussed. The material demonstrates the emergence of new linguistic features as a result of long-term contacts with the structurally different languages Uzbek and Tajik.

### *Phonology*

As in literary Arabic and most Eastern Arabic dialects, the distribution of consonants and vowels in the Qashqa-darya dialect is relatively equal. In Western Arabic dialects, vocalization is minimal and the occurrence of word-initial consonants clusters is typical (*kṭeb* vs. *kataba*).

Seven main classes of articulation of consonants are found: labial, dental, alveolar, palatal, velar, pharyngeal and laryngeal.

Two phonemes not found in literary Arabic are encountered: *p* and *č*. They occur in words of Tajik and Uzbek origin as well as in purely Arabic lexemes, e.g. *polišta* 'pillow', *čai* 'tea', *harap* < *harab* 'fled, bolted, escaped', *uči* < *waġhī* 'my face'.

There is a tendency to devoice voiced consonants. Thus the voiceless aspirated *p* occurs word-initially and word-finally after vowels, e.g. *arkup* < *arkaba* 'ascended', *arap* < *arab* 'Arab', *gapa* < *gabā* 'gown'. In intervocal position *b* is sometimes replaced by the voiceless medial *b*.

The consonant *ġ* is in some cases devoiced to *č*, e.g. *hač* < *ħaġġ* 'pilgrimage'. *č* has not emerged through palatalization of *k*, which is typical of most Arabic dialects. *č* is also found in words of foreign origin, e.g. *takča* 'niche', *kuča* 'street'.

Table 1: *Minimal pairs*

<i>p / f</i>	<i>ḥalap / ḥalaf</i>	'Aleppo' / 'swear'
<i>p / m</i>	<i>harap / haram</i>	'he escaped' / 'sin'
<i>p / b</i>	<i>pūl / būl</i>	'money' / 'urine'
	<i>ḥalap / ḥalab</i>	'Aleppo' / '(he) milked a cow'
	<i>arap / arab</i>	'he escaped' / 'Arab'
<i>č / ğ</i>	<i>čōi / ğōi</i>	'tea' / '(somebody) is coming'
	<i>čarah / ğarah</i>	'roaming' / 'wound'
<i>č / š</i>	<i>čaġal / šaġal</i>	'spill' / 'jackal'

Word-final weakening of consonants may be explained by Uzbek and Tajik influence as well as by general dialectal tendencies: *ṭalaʿ* > *ṭalaḥ* 'ascend', *farah* > *farah* 'rejoice'.

The vowel system is characterized by three degrees of openness. The following vowel phonemes are identified:

Table 2: *Vowel phonemes*

<i>i</i>		<i>u</i>
<i>e</i>	<i>o</i>	
<i>a</i>		

The reduction of diphthongs has contributed to the expansion of the classical system: *ai* > *e*, *aw* > *o*. The quality of vowels is often determined by position. The environment of coronal, velar, pharyngeal and emphatic consonants affects the timbre of the vowels.

The absence of phonological length distinctions is a characteristic feature of the Qashqa-darya dialect. Occurrence of long vowels can be ascribed to various factors, the most important one being stress.

The vowel systems of the Central Asian Arabic dialects are influenced by Tajik. In both dialects *ā* mostly corresponds to *ō*, e.g. *māʿ* > *mōya* 'water', *zamān* > *zamōn* 'time', *ġawāb* > *ġawōb* 'answer'. This phenomenon, which is manifested more clearly in the Bukhara dialect, must partially be ascribed to Tajik influence (Tsereteli 1970: 168). In the Qashqa-darya dialect, *o* is a vowel between *ā* and *ō*. Many parallel forms containing *ā* and *ō* are found, e.g. *nāb* / *nōb* 'old man', *ġawāb* / *ġawōb* 'answer'. There are also many cases of *ā* without parallel forms containing *ō*, e.g. *gāl* 'come', *kān* 'was'.

The phoneme *ō* may be established by means of minimal pairs such as *kōn* 'many' vs. *kān* 'was'. It has emerged in two ways: from *ā*, e.g. *mōt* < *māt* 'died', and from the diphthong *aw*, e.g. *nōm* < *nawm* 'sleep', 'dream'.

### Morphology

As far as the structure of the adjective is concerned, the following main types may be identified: aCCaC for masculine forms and CaCCa for feminine forms. The widespread type C<sub>1</sub>VC<sub>2</sub>C<sub>3</sub>V has the following variants: C<sub>1</sub>aC<sub>2</sub>C<sub>3</sub>ān and C<sub>1</sub>oC<sub>2</sub>C<sub>3</sub>ōn, e.g. *šab<sup>c</sup>ān* 'satisfied', *ğow<sup>c</sup>ōn* 'hungry'.

The superlative degree is formed by means of *ğüdo*, e.g. *ğüdo zen* 'best', *ğüdo kbīr* 'biggest'. This formation must be ascribed to Uzbek influence. The form *ğüdo* is connected with Arabic *ğiddan*.

The structure of numerals is different from that of Literary Arabic. From 'one' to 'four' the Arabic forms are used, from 'five' to 'seven' both Arabic and Tajik forms, for 'eight' and 'nine' only Tajik forms, whereas 'ten' is expressed by means of Arabic *ašara*.

Table 4. *Cardinals A*

1	<i>fadhat-fadhate</i>	6	<i>sitta / šiš</i>
2	<i>isnēn-isnēna</i>	7	<i>sab<sup>c</sup>a / haft</i>
3	<i>salās</i>	8	<i>hašt</i>
4	<i>'orba</i>	9	<i>nūh</i>
5	<i>ḥamsa / paṅḡ</i>	10	<i><sup>c</sup>ašara</i>

The form *fadhat* 'one' is a merger of the two words *fad* < *fard* and *hat* < *'ahad*. Sometimes the word *wahad* is used. The numerals from 'eleven' to 'nineteen' have the structure 'ten' + decimal unit. This is one of the distinctive features of this dialect.

Table 5. *Cardinals B*

11	<i><sup>c</sup>ašara fadhāt / <sup>c</sup>ašar wahda</i>	16	<i><sup>c</sup>ašara sitta / <sup>c</sup>ašara šiš</i>
12	<i><sup>c</sup>ašar isnēn</i>	17	<i><sup>c</sup>ašara sab<sup>c</sup>a / <sup>c</sup>ašara haft</i>
13	<i><sup>c</sup>ašara salās</i>	18	<i><sup>c</sup>ašara hašt</i>
14	<i><sup>c</sup>ašar orba</i>	19	<i><sup>c</sup>ašara nūh</i>
15	<i><sup>c</sup>ašara ḥamsa / <sup>c</sup>ašara paṅḡ</i>		

The numerals for 11, 15, 16 and 17 exhibit parallel forms. When the numeral denoting the decimal unit begins with a consonant, the word *<sup>c</sup>ašara* keeps its final vowel -a. In the numerals for 12 and 14 the vowel -a is omitted, because the second word begins with a vowel following the muting of *hamza* (Chikovani 2000: 193).

Numerals of this category precede nouns in the singular form: *<sup>c</sup>ašara fadhāt walad* 'eleven boys', *<sup>c</sup>ašara salās hōiṭ* 'thirteen houses'. This phenomenon is explained by Uzbek influence.

The verbal system is constructed according to the principles of Semitic languages. Simple and derived forms are found: perfective and imperfective



forms, a well-defined subsystem of conjugation expressed by the categories of person, number and gender, with regular, geminate, hamzaed, weak and doubly weak verbal forms, and indicative and imperative moods.

The category of tense is represented by the following forms: Past perfect, Past indefinite, Present indefinite, Present definite, Simple future, Complex (immediate) future. The past indefinite is formed by adding a pronominal suffix to the form of the active participle. With respect to its content it is similar to the Georgian 'resultative'. The present definite is similar to the English present continuous. It is formed by adding the active participle forms of the verb *nām* to a form of the imperfect, e.g. *iḡbuḡ nāyim* '(somebody) is now baking (something)', *nōkul nāyiminni* 'we are now eating'. The participle of the auxiliary verb *nām* is frequently supplied with a pronominal suffix (Tsereteli 1941: 142). The presence of the past indefinite and present definite forms in the Qashqa-darya dialect could be explained by Turkic influence.

### Syntax

The main types of constituent order are the following:

Subject + Predicate + Object, e.g. *pōšō gāl i-walad* 'The pasha said to the boy ...'

Subject + Object + Predicate, e.g. *bōy i-bint ḡušrūya gāl-ki* 'The bey said to the beautiful girl ...'

Object + Predicate + Subject, e.g. *mōrati i-hama māñī-gāl-ki* 'He said: I don't give him my wife'.

Subject + Object + Predicate is the most frequent type. The free word order can be explained by contact with Uzbek.

The Qashqa-darya dialect has developed certain syntactic constructions that differ from those of Literary Arabic:

The head + modifier construction, which is typical of Arabic, is widely used. Its productivity is favored by the existence of the definite article *al*, e.g. *ṣōḡb il-bāgir* 'the owner of the cow', *rōs il-iḡmōr* 'the donkey's head'. Examples without the article: *laḡm 'anza* 'flesh of a goat', *gidīr palāw* 'pan of pilaf'.

The modifier + head construction, which is alien to Semitic languages, has developed under Uzbek influence. The head carries a pronominal suffix, e.g. *āfāndi morta* 'the effendi's wife', literally 'effendi wife-his', *hawliōy ṣōḡbu* 'the owner of the farmhouse', *waladak farasu* 'your boy's horse'. This construction is characteristic of Uzbek (Kononov 1948: 237). In order to express collectivity, the marker *-āt* is used, e.g. *na'aḡāt ṣōḡba* 'the owner of the sheep'. The use of several modifiers is possible, e.g. *faras, na'ḡa, 'anza orzum* 'the pasture of horses, sheep and goats'.

During field work in the dialect area, informants were observed to use, in code-switching, Turkic genitive constructions in which the modifier takes the genitive case suffix and the head carries the third person possessive suffix, e.g. *Dilmurād-niḡ kitāb-i* 'Dilmurad's book'.

Another modifier + head construction can be ascribed to contact with Tajik, e.g. *laḡam mibīḡ* 'meat-seller'.

The use of the personal pronoun + noun + pronominal suffix construction can be ascribed to Turkic influence, e.g. *anā hawliōi* 'my farmhouse', *ana ḡōiṡi* 'my house', *inta waladak* 'your boy'.

### References

- Chikovani, G. (2000) 'The Qashqa-Darya Arabic Dialect of Central Asia (nominal structures)', in M. Mifsud (ed.) *Proceedings of the Third International Conference of AIDA: Association Internationale de Dialectologie Arabe, Held in Malta, 29 March–2 April 1998*, Paris: AIDA / Malta: Salesian Press, 189–194.
- Kononov, A.N. (1948) *Grammatika uzbekskogo jazyka*, Taškent.
- Tsereteli, G. (1941) 'On the characteristics of the language of the Central Asian Arabs', in *Trudy Instituta Vostokovedenija Akademii Nauk SSSR* 26: 133–148.
- (1970) 'The influence of the Tajik language on the vocalism of Central Asian Arabic dialects', *Bulletin of the School of Oriental and African Studies* 33: 167–170.