



Yaghnobi: an example of a language in contact*

Ľubomír Novák

Charles University in Prague

ABSTRACT:

Yaghnobi, an Eastern Iranian language spoken in northern Tajikistan, shares many important features with Sogdian and is generally believed to be descended from a nonliterary dialect of Sogdian. The language has historically been in intensive contact with Tajik and more recently with Russian; in addition, some features connecting Yaghnobi with the Pamir language area may be identified. Contact with Tajik has introduced thousands of lexical items, many of which have become an integral part of the Yaghnobi lexicon. There are also examples of Tajik grammatical loans: some of these features, such as the *izafet* construction or past participles in *-gí*, have become part of the Yaghnobi grammatical system, while other features, such as the direct object marker *-ro* or verbal imperfective prefix *me-*, are usually used in Tajik-dominated mixed speech communities. The intensity of language contact nowadays differs in the various areas inhabited by the Yaghnobis; the present article is based on material recorded in the Yaghnobi-speaking community in the Lower Varzob area, north of the Tajik capital Dushanbe.

KEYWORDS:

Yaghnobi, Sogdian, language contact, Tajik, Tajikistan, contact linguistics

1. HISTORY OF YAGHNOBI AND ITS LANGUAGE CONTACTS

Yaghnobi (*yaɣnōbí z'vók*) is an Eastern Iranian language spoken in areas of northern Tajikistan. Originally the language of the Yaghnob (Yagh. *Yáynōb*) valley, at present the language is also spoken in the Zafarobod district, in the Ziddeh and Takob valleys in Upper Varzob, in Lower Varzob, in Dushanbe, in the Hisor and Kofarnihon valleys, and in the Ghonchī region (see Figure 1). We may suppose that in past centuries the language was spoken in a wider area of the Yaghnob and Zeravshan valleys, as documented by some place-names and Tajik dialectal lexicon shared with Yaghnobi.¹ According to Hromov, some settled the southern slopes of the Hisor Range in

* I would like to thank Desmond Durkin-Meisterernst, Viktor Elšík, Jan Bičovský and Ronald Kim for proofreading and valuable comments. This article was made possible thanks to the support of the Charles University Grant Agency, Project Nr. 51807: "Electronic Lexical Database of Yaghnobi."

1 Mallickij observed that some of the Tajik-speaking inhabitants of Takfon used numerals '2', '3', '4' and '5' similar to those found in Yaghnobi: *ɖyau, tɪraɪ, tɪʃɔr, pɔndʒ* (Mallickij 1924), cf. Yagh. *du, saráy / t'ráy, tafór / t'ʃór, panč*, Taj. *du, se, čor, panj*. These archaic numer-



the 16th-17th centuries (Hromov 1972: 6); some Yaghnobis probably also settled in the Ghonchī region at the same time (Buzurgmehr 2005: 121). Much later, in the 1950's, they migrated to areas in the vicinity of Dushanbe and to Dushanbe itself. In 1970-71, the Yaghnobis were forced by the Soviet government to leave the Yaghnob valley and to settle in the Zafarobod district;² after the 1980s, some of them moved back to the Yaghnob valley. According to the data provided by Sayfiddin Mirzozoda, a Yaghnobi scholar of the Tajik Academy of Sciences, there are at present about 12,500 Yaghnobi speakers in Tajikistan (Mirzozoda, pers. comm.). For the external history of the Yaghnobi people, I recommend the following works: Buzurgmehr 2005; Bashiri 1997; Loy 2005; Mirzozoda 2007; Junker 1930; Panfilov 1991, Milovskij 1991; and Yorzoda 2007.

The Yaghnobi language is considered to be a descendent of Sogdian by many scholars (cf. Bogoljubov 1956; Klimčickij 1935), and both languages share many similar features in lexicon, grammar and phonology. However, the direct relationship of Yaghnobi to literary Sogdian has not been proven, the main difference between Yaghnobi and Sogdian being the different development of stress and the operation of the so-called Rhythmic Law in Sogdian, but not in Yaghnobi (cf. Sims-Williams 1989: 173; Bielmeier 1989: 480). It seems that Yaghnobi developed from an unattested dialect of Sogdian. There is a suggestion that Yaghnobi could be a descendent of a dialect of ancient Ostrushana (Hromov 1987: 645; Buzurgmehr 2005: 117), although we have no exact idea how that dialect differed from the Sogdian literary language.³ After the fall of Sogdiana, the language fell under the influence of the Persian language, but we cannot tell precisely when Persian influence began; according to Vinogradova, heavy influence of Persian on Yaghnobi began in the second half of the 19th century (Vinogradova 1999: 309). In the present stage of knowledge of the language, very strong influence from Persian can be observed especially in the lexicon, but some grammatical features have been borrowed as well. In addition to Persian, Yaghnobi also has many borrowings from other surrounding languages, including Arabic, Uzbek and Kyrgyz, and in the 20th century from Russian and, via Russian, internationalisms from Western languages.

als are no longer used in Takfon, but they raise the possibility that some other dialects related to Yaghnobi might have been spoken also in the vicinity of the Yaghnob valley in the past.

- 2 The Yaghnobis often refer to Zafarobod as *Zahfarōbōd*; the name was changed by folk etymology to represent the unwillingness of the people to live there (*zah* means 'mud' or 'filth' in Yaghnobi and Tajik). For the people who used to live in the high mountain valleys with clean water and fresh air, it was extremely difficult to adjust to the steppe conditions of Zafarobod with hot weather and muddy water.
- 3 Only three short Sogdian texts are known from the territory of Ostrushana from the fortress of Chilhujra south of Shahrison, Tajikistan. There is no dialectal difference between these three short texts and texts in "literary" Sogdian; they may be compared with contemporary documents from Mount Mugh near Xayrobod (cf. Livšic 2003). The Yaghnob valley was a part of the southern region of Ostrushana called *Buttām* or *Buttamān* (Sogd. *pyttm'n*, Arab. بْتَمَن or بْتَمَن).

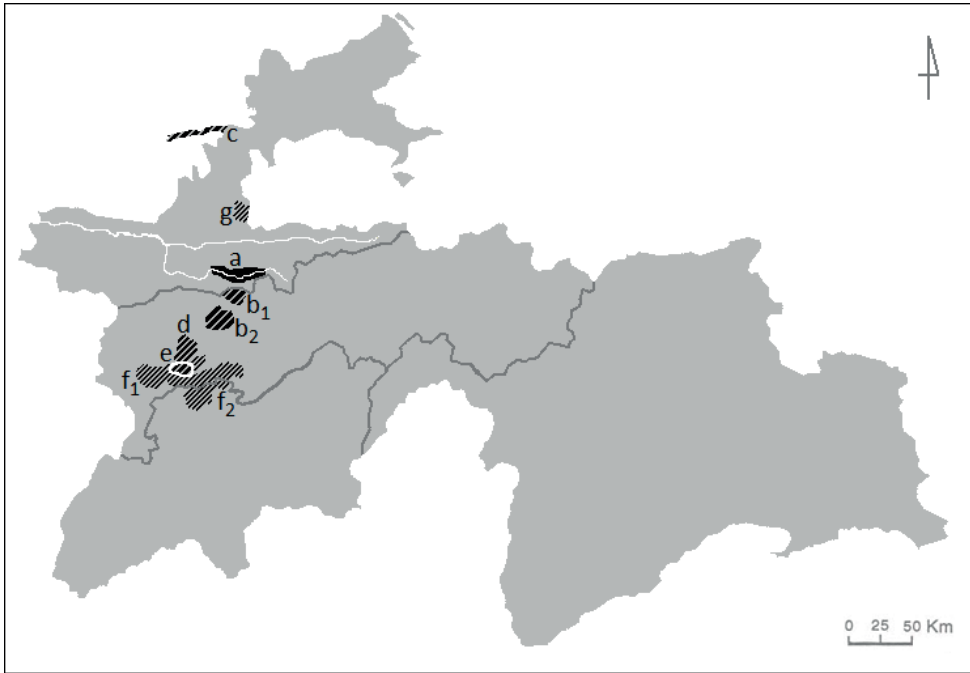


FIGURE 1. Main areas populated by the Yaghnobis in Tajikistan: a – Yaghnob valley, b₁ – Ziddeh valley (Upper Varzob), b₂ – Takob valley (Upper Varzob), c – Zafarobod, d – Lower Varzob, e – city of Dushanbe, f₁ – Hisor valley, f₂ – Kofarnihon valley, g – Ghonzhi region. (Shading indicates relative proportion of the Yaghnobi population in a given area.)

The etymological research on Yaghnobi is still in its beginnings. There are only three main sources utilizable for etymological studies: a Yaghnobi-Russian dictionary by Andrejev – Livšić – Pisarčik (1957) published together with Yaghnobi texts (Andrejev – Peščereva 1957); the Yaghnobi glossary of Salemann (Benveniste 1955); and an outline of historical grammar by Hromov (1987). In the two above mentioned dictionaries and also in a small dictionary attached to Hromov's Yaghnobi grammar (Hromov 1972), brief etymological notes are given. However, for the lexicon shared with Tajik precise etymologies are not given, so in some cases we have to undertake further study to distinguish words of Arabic or Uzbek origin.

In this paper, I will discuss the Yaghnobi language as an example of a language in contact, a topic which did not receive much attention in previous studies. Some basic data can be found in Hromov's Yaghnobi grammar in a short chapter dealing with the issue of bilingualism among the Yaghnobis (Hromov 1972: 105–118). To address this problem, it is necessary to focus on several questions concerning other languages which have been in contact with Yaghnobi. I have specified some of those languages in previous paragraphs. Let's have a look at them in greater detail:



- a) **Persian/Tajik:** Tajik⁴ may be considered the main influence on Yaghnobi today. Contemporary Tajik may be divided into four main dialectal groups: the South-eastern/Darvoz type, the Southern type, the Central/Zeravshan type and the Northern type (Rastorgujeva 1964). In the case of Yaghnobi, the Central/Zeravshan type is the relevant dialectal group. The dialects of this type are divided into two main subtypes, the dialects of the Mastchoh area and those of the Falghar region. For the Yaghnobi-speaking settlers in Upper Varzob, it is necessary to focus also on the possible relations with the Varzob dialect of the Northern dialectal group.

A study of the Tajik dialects in comparison with the historical development and lexicon of Yaghnobi yields some interesting results. In particular, the Zeravshan-type dialects share many features with Yaghnobi, so that their precise study can offer an outline of some possible mutual influences and be compared with the historical development of Yaghnobi. This question will be discussed below in section 2.

- b) **Arabic:** The extent of Arabic borrowings in Yaghnobi is comparable to that of Arabic borrowings in Tajik: many cultural words were borrowed that cover many semantic fields of the lexicon, mainly connected with religion and law. The influence of the Arabic lexicon on the Central Asian languages is considerable. Many of these loanwords are no longer considered to be borrowings, partly because they have been phonologically adapted to the borrowing language, partly also because the replacement of the Perso-Arabic alphabet by the Latin and later the Cyrillic alphabet caused the loss of some etymological information.⁵ In addition, many typical features of Arabic (mainly the so-called broken plurals) were lost or replaced by the grammatical structure of the borrowing language, e.g. Yagh. pl. *kitóbt* (sg. *kitób* 'book') instead of the Arabic broken plural *kutub* (sg. *kitáb*), cf. Taj. pl. *kitobhó*, Pers. pl. *kitābhá* (or *kutúb*), Uzb. *kitāblar*, but Chaghatay pl. *kütüblär*. In contrast to Tajik, Yaghnobi preserves the distinct pronunciation of Arabic *w*, which merges with *v* in Tajik. Yaghnobi and the Northern and Central Tajik dialects also distinguish both pharyngeal sounds *ħ* and *ʕ*, while in standard Tajik *ħ* merged with *h*. The Arabic sounds *ʔ* and *ʕ* merge into *ʕ* in Tajik, but this sound is preserved only before or after a consonant; word-initially and between vowels it is lost.⁶
- c) **Other Iranian languages (mainly Pamir languages):** Some Yaghnobi words appear to be of native Iranian origin, but have no generally accepted Sogdian equiva-

4 I use the term *Tajik* to denote the Tajik dialects of (Classical) Persian with their specific dialectal features that differ from the development of Persian in Iran and Afghanistan. By Persian I mean Classical Persian (i.e. *fārsī-yi darī*); when speaking about its modern variant in Iran I will refer to it as to Farsi.

5 In comparison to the original Perso-Arabic alphabet used for writing Persian/Tajik or Chaghatay/Uzbek, the Cyrillic letter *χ* represents *و* and *ح* in Arabic, the letter *з* can represent as many as four letters of the Arabic script: *ز*, *ذ*, *ض*, *ظ* and so on. The Arabic letters *ع*, *ء* are represented as *ѣ* in Cyrillic or are omitted in writing altogether.

6 In colloquial Tajik, *ʕ* is phonetically usually [ʔ] or *∅*.



lent. Some of those words have an etymology consistent with one or more of the Pamir languages. This may be evidence of early contacts.⁷

- d) **Turkic languages:** Words of Turkic origin also form a major part of the borrowed lexicon of both Yaghnobi and Tajik. Many of these are early borrowings from the period of Turkic expansion in the 14th and 15th centuries. Turkic words also appeared in some numbers in Sogdian, but the majority of Turkicisms come from the later period of intensive contacts between Iranian speakers and the Karluk and Kypchak Turks. Because Yaghnobi has been recorded for less than 100 years, we can hardly tell whether Turkic words came into the language directly from Turkic languages or whether they were borrowed via Tajik: loans from Uzbek could have come into Yaghnobi directly, whereas Kyrgyz and also Kazakh or Uyghur words were probably borrowed via Tajik. Through Turkic, some Mongol words were also introduced into Tajik and Yaghnobi.
- e) **Russian:** Russian influences in Yaghnobi and Tajik became especially strong after the October Revolution in 1917. During the Soviet period, many words derived from Latin and Greek came in their Russian form into Tajik. Russian borrowings tend to keep their original pronunciation in Tajik, but are usually adapted to native pronunciation in Yaghnobi. The usage of Russianisms may differ among speakers of these languages due to their perceived prestige compared to native Yaghnobi or Tajik expressions.
- f) **English and Western internationalisms:** In the most recent period English has begun to play a greater role in the lives of many Central Asians. Nowadays its influence is visible in Tajik, and to a lesser extent in Yaghnobi. Many English words are adopted via Russian in Russianised forms.

2. DOMAINS OF CONTACT

2.1 PHONOLOGICAL EVIDENCE FOR LANGUAGE CONTACT

Yaghnobi shares some phonological features with the Zeravshan dialect group of Tajik. The development of Proto-Iranian vowels and diphthongs in Yaghnobi is given in Tables 1 below, while Tables 2 and 3 shows the development Classical Persian vowels in Zeravshan Tajik and in the Tajik literary language:

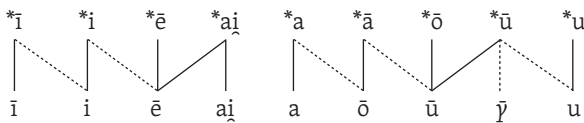


TABLE 1. The development of Yaghnobi vowels.^{8,9}

7 Such lexicon is represented by words like *γayk* 'girl, daughter' (cf. Yazghulami *yačág* and Shughni *γāc* 'girl, daughter'), *k'šók* 'bull' (cf. Ishkashmi *kəjuk* 'bull', Sarghulami *kišó* 'cow'), *ód'ma* 'Saponaria plant' (cf. Khufi *wuđm*), etc.

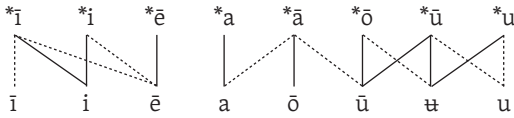


Table 2. The development of vowels in the Zeravshan dialect group of Tajik.¹⁰

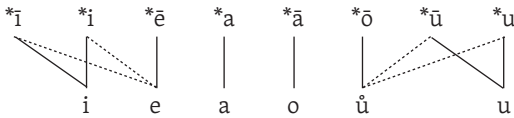


TABLE 3. The development of standard Tajik vowels.^{11,12}

The question is whether Yaghnobi as a substrate could have influenced similar changes in the Zeravshan Tajik dialects, or if the parallel development of the vowel system is a Tajik dialectal feature which has influenced Yaghnobi. The same question arises in connection with other phonological features shared by the two languages. Thus Biellemeier suggests that the Yaghnobi change $*\delta > d$ and the raising of prenasal $*\bar{a} > \bar{o} > \bar{u}$

- 8 The development $*\bar{u} > \bar{y}$ has not yet been successfully explained, but it seems that every historical stressed $*\bar{u}$ changed to \bar{y} . It also seems that this change has not taken place in all dialects (Mirzozoda, pers. comm.). On the other hand, Hromov notes that every Yaghnobi native \bar{u} is stressed (Hromov 1987: 649), so we can suppose that every $*\bar{u}$ gave \bar{y} in the dialects where this change appeared, but we see this change also in words of Tajik origin, e.g. *kabýd* : *kabúd* < Pers. *kabūd* ‘green, blue’. In words of Tajik origin, \bar{y} appears in the same places as u in the neighbouring Zeravshan dialects of Tajik.
- 9 In Yaghnobi there are three short vowels: *i*, *a*, *u*; the other vowels (*ī*, *ē*, *ō*, *ū*, \bar{y}) are all long. Short *i* or *u* can be reduced in open syllables before a stressed syllable: this is marked as ⁱ, ^u. Historical $*a_i$ (or $*\bar{a}$) changes into *ē* in Eastern Yaghnobi, but remains *a_i* in the western dialect.
- 10 The sound I transcribe u is realised as [ɣ] in the dialects of Upper Mastchoh and in some dialects of Falghar, and as IPA [ɔ] in the dialects of Lower Mastchoh and in the majority of Falghar dialects (cf. Rastorgujeva 1964, 25–6; Ido 2009: 68).
- 11 The vowels *i*, *a*, *u* are underlyingly short, but they can be slightly lengthened under the influence of the word stress. The vowels *e*, *o*, *ū* are always long. This distinction of long and short vowels can be observed in some Tajik dialects, while in the standard literary language there is no difference at all (cf. Ido 2009). In the Tajik alphabet the letter \bar{u} (transliterated *ī*) is used to distinguish word-final (stressed) *i* from the *i* of the izafet.
- 12 The changes $*\bar{i} > e$ and $*\bar{u} > \bar{u}$ (IPA [ɔ:]) take place before *h* (i.e. $*h$ or $*\bar{h}$) and \bar{r} (i.e. $*\bar{r}$ or $*\bar{r}$) in closed syllables, e.g. *kūhná* < Pers. *kuhná* ‘old, ancient’; *mehmón* < Pers. *mihmān* ‘guest’; *rūh* < Arab. *rūh* ‘soul, spirit’; *mūšmín* < Arab. *mušmin* ‘faithful’; *ešlón* < Arab. *išlān* ‘announcement’; *šabēh* < Arab. *šabih* ‘alike, equal’. This change was rarely documented in historical Perso-Arabic orthography, e.g. *mehmón* was written *ميهمان* instead of the more common *مهيمان*. The change $*\bar{u} > \bar{u}$ is also recorded in the vicinity of the Arabic pharyngealised consonants in some Tajik words, e.g. *sūfī* (Pers. *šūfī*) < Arab. *šūfī* ‘Sufi, dervish’, *arūz* (Pers. *šarūz*) < Arab. *šarūd* ‘meter (in poetry)’.



were triggered by contact with Tajik (Bielmeier 1989; 2006), and there is also a similar development of the Iranian ending **-ākā-* > *-a* in Yaghnobi and Tajik (cf. Sogdian **-aka-* > **-ē-*, **-ākā-* > **-ā-*). These changes show some phonetic similarities between Yaghnobi and Tajik, but they do not tell us whether the Yaghnobi changes were affected by Tajik or occurred independently. Regarding the change **δ* > *d*, Hromov writes that we cannot posit **δ* as a part of the early Yaghnobi sound inventory, because we would then expect different developments in Yaghnobi dialects: **δ* > *ʃz* (E) : *ʃd* (W), cf. the documented change **θ* > *s* (E) : *t* (W) (Hromov 1972: 121).¹³ The change of **δ* > *d* has also taken place in the Persian borrowing *guḍar-* : *guḍaštān* > *g^udāarak*, Tajik *guzar-* : *guzaštān* ‘to pass’ (Farsi *gozār-* : *gozāštān*; cf. Andrejev — Livšic — Pisarčik 1957: 255).

As for the alternation *ō* : *ū* < **ā* before a nasal, this phenomenon is observed not only in the Southern and Central Tajik dialect groups (Rastorgujeva 1964: 28), but also in many other modern Western Iranian languages and dialects.¹⁴ To take an additional example, the lowering of historical **i* before **h* in a closed syllable in Yaghnobi is similar to the treatment in Tajik, so that the root of **dih-ak* ‘to hit’ **dih-* changed into *deh-* before an ending beginning in a consonant: *man dih-ōmišt* ‘I hit’, *ax deh-tišt* ‘he hits’. In the contemporary Yaghnobi language this change has been completely generalised to all positions, so the root **dih-* has been replaced by *deh-*.¹⁵

In contemporary Yaghnobi, the opposition of long vs. short vowels is disappearing under the influence of Tajik. Vowel length is related to word stress, similarly to contemporary Tajik (Vinogradova 1999: 291). Due to the shortening of historical long vowels, length is not consistently marked in Yaghnobi dictionaries and textbooks printed in Tajikistan (cf. Mirzozoda 2007; 2008; Mirzozoda — Alavī 2008; Mirzozoda — Qāsemī 1995). On the other hand, the disappearance of length did not influence the quality of the vowels. In the scholarly notation of the language, it is necessary to mark the (historical) length to signal the different vowel qualities in the archaic Yaghnobi pronunciation.¹⁶

¹³ We do not know the status of **δ* in early Yaghnobi. Based on the observations of Junker (as one of the earliest scientific investigators of Yaghnobi), we can tell that the change **δ* > *d* occurred earlier than the change **θ* > *t* : *s*. Junker writes that the older generation in the Yaghnob valley kept the pronunciation *θ*, but the younger speakers substituted *t* in the western dialect and *s* in the eastern dialect for it, but he does not mention **δ* (cf. Junker 1930, 128). These two sounds often do not behave as a class: cf. Old Persian and the West Germanic languages.

¹⁴ The raising of prenasal **ā* > *ō* > *ū* seems to differ in the Yaghnobi dialects. In the dialect of Gharmen (Eastern Yaghnobi), there are some words that have undergone the change, but in many other words *ō* has remained unchanged. On the other hand, in the pronunciation in the dialect of Zumand (Western Yaghnobi) I recorded that this change has been fully realised, e.g. *nūm*, *nūn*, *tafarōmišt*, *mūnōmišt*, *°xōna* (Gharmen) : *nūm*, *nūn*, *ʃfarūmišt*, *mūnūmišt*, *°xūna* (Zumand) < **nām*, **nān*, **θbarāmišt*, **mānāmišt*, (°)*xānā* ‘name, bread, I give, I put, house’; cf. *nūm*, *nūn*, *xūnā*, *dandūn* < **nām*, **nān*, **xānā*, **dandān* ‘name, bread, house, teeth’ in the Zeravshan dialects (Rastorgujeva 1964, 28).

¹⁵ Unfortunately this feature cannot be illustrated with other examples because the sound *h* is rare in genuine Yaghnobi words.

¹⁶ The quality of Yaghnobi short *i* and *u* differs according to the adjacent consonants, but this is not the case for *ī* and *ū*. If the length marks were omitted, Yaghnobi words would be



The proto-Yaghnobi phonemic opposition of unvoiced stop vs. voiced fricative¹⁷ (*p : *v < *θ, *k : *γ, *č : *ž but *t : *d < *δ) disappeared with the introduction of voiced stops due to numerous Tajik borrowings (cf. Bielmeier 2006). In contrast to Persian and Sogdian, the consonantal groups *nd*, *ng*, *mb* are not allowed in Yaghnobi, which has only *nt*, *nk*, *mp*. Place names of Sogdian or Yaghnobi origin in Tajikistan treat these groups as a combination of nasal and voiced stop, but we do not know whether these borrowings retain the original pronunciation or if they have undergone the process of voicing of stops that we know from the historical development of Persian.

Some early borrowings from Tajik were assimilated to the Yaghnobi opposition of unvoiced stops vs. voiced fricatives. Hromov gives as examples two Tajik loans where we can observe the change of Tajik *g* into Yaghnobi *k*: *kūr* < Pers. *gōr* 'barrow'; *k^usēlak* < Pers. *gusēl kardān* 'to send' (Hromov 1987: 565). Other examples of devoicing of voiced stops are *xapār* (E) : *x'pār* (W) < Pers. *xabār* < Arab. *ḥabar* 'information' (Hromov 1972: 85); *lakát* (E) : *l'kát* (W) < Pers. *lagád* 'a kick'; and *otaškírák* < Taj. *otašgirák* 'pincers'. Some words show the change *b* > *m* or *m* > *b*, such as *kás'm* < Pers. *kasb* < Arab. *kasb* 'craft, occupation'; *mūs* < Pers. *bōsá* 'kiss' (in the dialect of Mastchoh *mūs*); *falaxbán* < Pers. *falaxmān* 'sling'; *tamassúm* < Pers. *tabassúm* < Arab. *tabassum* 'smile'; the change **m* > *b* can also be found in the Yaghnobi word *b'dōn* 'middle', cf. Sogd. *myδ'n* */*miδān*/, Pers. *miyān* < Iran. **madyāna*-. Whether **d* was borrowed as *t* we cannot tell with certainty. Based on these facts, it appears that the opposition **t* : **δ* changed to **t* : **d* at an earlier stage of the language (Bielmeier 1989: 482), so the voiced dental stop became an integral part of the phonetic system, but other voiced stops appeared in Yaghnobi much later.

In the phonology of loans from Tajik into Yaghnobi, one observes a tendency to change **š* > *č* and **ž* > *ǰ*, e.g. *ōč* < Taj. *oš* 'pilař'; *čapalóq* < Taj. *šapalóq* < Uzb. *šapatáq* 'slap'; *ǰóla* < Pers. *žálá* 'hail'; *mǰa* < Pers. *mižá* 'eyelash'. The change *ž* > *ǰ* is also common in colloquial Tajik and in Uzbek. A example of an early borrowing may be *čuyrót* / *čuryót* 'yoghurt' < Pers. *juyrát* (dialect of Samarqand; Steingass 1892: 365), Taj. *ǰuryót* / *ǰuryót* (< Kypchak Uzbek **ǰuryát*); this example shows the change **ǰ* > *č*, whereas the variant *ǰuryót* is due to recent contacts with Tajik.

Another characteristic that Yaghnobi shares with Central Tajik dialects is the preservation of the pharyngeal sounds *ħ*, *ʕ* in borrowings from Arabic.¹⁸ In many

written in a transcription appropriate to the phonology, but some etymological information would be lost. In the ideal case, both the historical-etymological and phonetic transcriptions should be used, e.g. *mórti* 'man' ['mo:rt^he], *áwi* pron. 3. sg. obl. [ʔaβe], *wajš* (W) / *wēš* (E) 'grass' ['βajš] / ['βe:š], *mēt* (W) / *mēs* (E) 'day' ['me:t^h] / ['me:s], *Yáynōb* 'Yaghnob' [ʔjənoβ], *yar* 'mountain' ['bɑ:r], *s^utúr* 'sheep' [sø't^hu:r], *zindagī* 'life' [zinda'g'i:], *z'vók* 'language, tongue' [zi'vo:k^h], *urk* 'wolf' [ʔɔrk^h], *v'róť* 'younger brother' [vi'ro:t^h], *ēx* 'ice' [ʔɛ:χ], *uxš* 'six' [ʔoxš, ʔoxš], *vuz* 'goat' [ʔvuz], *kabýd* ~ *kabúd* 'green, blue' [k^ha'by:d] ~ [k^ha'b^u:d], *kūprúk* 'bridge' [k^hɔp'røk^h], *širín* 'sweet' [ʃi'ri:n], *tík* 'again' [t^hi:k^h] etc. (cf. Sokolova 1953).

17 The voiced fricatives developed from Proto-Iranian voiced stops in Eastern Middle Iranian languages. An analogous situation can be observed in Byzantine Greek.

18 Sokolova reports that *ħ* and *ʕ* appear in Upper Varzob Yaghnobi in the same words as in the neighbouring Varzob dialects of Tajik (Sokolova 1953). A similar feature is documented in the dialects of Mastchoh (Hromov 1962).



texts, these consonants appear as an integral part of the phonetic system (cf. Hromov 1972: 13; Hromov 1987: 651; Bielseimer 1989: 482; Bielseimer 2006; Vinogradova 1999: 292). However, there is no difference between *h* and *ħ* in the contemporary language, where both merge as *h* (Mirzozoda, pers. comm.). The pharyngeal sound *ʕ* is documented in just one Yaghnobi word, *šaʕmák* ‘soot’ (Mirzozoda 2008: 252); in other cases it disappears (*ǰumá* < Pers. *ǰumʕá* < Arab. *ǰumʕaħ* ‘Friday’) or lengthens the preceding vowel (*lālī* < Pers. *laʕlī* ‘bowl’; *ǰām* < **ǰaʕm* < Pers. *ǰamʕ* < Arab. *ǰamʕ* ‘sum, amount’; cf. Hromov 1972: 12; see also *tārīx* < Pers. *taʕrīx* (Farsi *tārīx*, Taj. *taʕrīx*) < Arab. *taʕrīħ* ‘history, date’).¹⁹ According to Mirzozoda, *ħ* and *ʕ* were pronounced only by mullahs who knew the Arabic language (Mirzozoda, pers. comm.), which would explain their disappearance in the contemporary language. The other possibility is that some Yaghnobi speakers pronounced the Arabic loans as they learned them from Tajik speakers of the Zeravshan dialects. Another interesting contact phenomenon is the different treatment of Arabic *w*, which in Tajik merges with *v* but in Yaghnobi Arabic with historical **w*, which in turn later changes to [β]. We can also assume that many Arabic words (especially religious terms) kept their original pronunciation in an archaic form of Tajik.²⁰

With reference to premodern Yaghnobi phonology, we may regard the consonants *b*, *g*, *h*, *ħ*, *ǰ*, *l*, *q*, *ʕ* and the clusters *nd*, *ng*, *mb* as non-native: their presence may be considered as an indicator of lexical borrowing. Some sounds are diagnostic of the language of origin, e.g. the pharyngeal sounds *ħ* and *ʕ* point to Arabic, while the voiceless uvular stop *q* points to Arabic or Turkic. Only a few voiced stops appear in native words as a result of voicing assimilation, e.g. *b^hdúfs-* ‘to glue, to attach’ < Sogd. *pδwfs-* */*p(ə)δwfs-* / < Iran. **upa-dubsa-*; *čáǰna* ‘milking pail’: *čak-* ‘to milk’ (Hromov 1987: 656). In the Yaghnobi word *díndak*, ‘tooth, teeth’, Hromov explains the voicing of **-nt-* > *-nd-* in **díntak* > *díndak* (Sogd. *δnt’k*, *δnt’kH* */*šámdak/*) as influence of the first voiced dental stop (Hromov 1972: 128); another possibility is influence of Tajik *dandón* (ibid.). The status of *-l-* is uncertain: it can originate from a Proto-Iranian group **-θr-*, but evidence for this change is rare. The only certain example is the development of Iran. **puθra-ka-* ‘son’ > *púl(l)a* ‘boy’; less certain is the case of Iran. **āθr* > *ōl^o* ‘fire’ in verb *ōlxášak* ‘to burn’, which could be a borrowing from Tajik (Hromov 1972: 127).²¹

¹⁹ A similar compensatory lengthening has occurred in the surrounding Tajik dialects of the Zeravshan group (Hromov 1962; Hromov 1972, 195) and also in the dialects of Varzob (Rastorgujeva 1952).

²⁰ We can also assume that the Arabic words came into Yaghnobi through Tajik; the distinction *w*, *ħ*, *ʕ* was kept because of the pronunciation of the mullahs. However, this hypothesis has not been proven, and it supposes a different kind of language contact. I have recorded just one exception in a different pronunciation of the Yaghnobi word *waxt* ‘time’, which is normally pronounced as [ʕwaxt] or under Tajik influence also as [vaxt] < Arab. *waqt* (cf. Tajik *vaqt*, dialectal also *vaxt*). A similar situation can be seen in the word *aw(w)ál* ‘first’ — [ʔaʕβ(:)al] or also [ʔaʕv(:)al] < Arab. *awwál*, Taj. *avvál*. In some Yaghnobi words the sounds [β] and [v] are interchangeable, as is documented in the dictionary of Andrejev — Livšic — Pisarčik (1957).

²¹ Hromov notes that the verb *ōlxášak* could have been formed by reanalysis from Tajik *alóu/oláu/olóu* ‘fire’ (cf. Yagh. *ōlōu* from Tajik; cf. Hromov 1972: 127).



As for Yaghnobi phonotactics and syllable structure, no consonant clusters can occur word-initially; a svarabhakti vowel is inserted between (or before) a sequence of two consonants. In many cases, the svarabhakti vowel is ultra-short in pronunciation, e.g. *saráy* (E) : *tṛáy* (W) < *θray ‘three’; *xʷšift* < *xšift ‘milk’; this also affects loans as well: *tṛáktʳ* < Russ. *mpák mop* ‘tractor’; *istál* < Russ. *cmol* ‘table’.²² The sequence *-CjV(-) is not native in Yaghnobi, where it is often rendered -CijV(-) or -jCV(-), e.g. *dayró* ~ *darʷyó* < Taj. *daryó* ‘river’ < Pers. *daryá* ‘sea’ < OPers. **draya-āpa-* (cf. Av. *zraia-* ‘sea’); *dunʷyó* ~ *duynó* ~ *dunyó* < Pers. *dunyá* < Arab. *dunyā* ‘world’; *samalʷyót* ~ *samaylót* < Russ. *camonēm* ‘aeroplane’; *bis(s)ʷyór* ~ *bis(s)yór* < Pers. *bisyār* ‘much, many’.

2.2. GRAMMATICAL CONTACTS

In addition to phonological features, some features in the grammatical structure of Yaghnobi may be interpreted as results of language contact with Tajik and Persian, as well as with Turkic languages.

2.2.1. PRONOUNS

An interesting feature that is shared with the Pamir languages is the oblique case of pronouns. In some of the Pamir languages, there are different forms in the nominative and oblique in all three persons singular and in the 3rd person plural. In Sogdian the situation is comparable. In Yaghnobi there is a similar pattern except in the 1st person singular, where nominative and oblique share a single form. I propose that the replacement of older Yaghnobi ^(*)az (cf. Gauthiot — Benveniste 1929: 108–9; see Sogd. *ʷzw, zw* ^{*/əzú/}) by its oblique form ^(*)mán(a) (cf. Sogd. *mnʷ* ^{*/məná/}) was triggered by assimilation in form to the Persian 1st person singular pronoun *man* (see Table 4).

2.2.2. VERBS

Contacts with Persian also introduced into Yaghnobi a feature especially typical of the Persian language, namely compound verbs. This feature can also be found in other languages in contact with Persian, whether Iranian (Kurdish, Balochi, Zazaki, Ossetian, Pamir languages, etc.; a few examples are also found in Sogdian and Khwarezmian) or non-Iranian (Turkic languages such as Uzbek, Uyghur, Chaghatay, Qashqai or Ottoman Turkish, as well as the Central Asian dialect of Arabic). Some Yaghnobi compound verbs can be regarded as calques from Tajik, e.g. Yagh. *yóđ-i nósak* or *zindađi-ĭ g^uđajrónak* ← Taj. *yod giriftán, zindađi guzarondán* ‘to remember, to spend

22 Persian, like Yaghnobi, does not tolerate word-initial consonant clusters, so a svarabhakti vowel is inserted in *CC- > CVC-. Russian loanwords in Tajik are often written in the same way as they are in Russian, but in colloquial speech the pronunciation is influenced by Tajik syllabic structure, e.g. *planetári* /p^lanetári/ ‘planetarium’ < Russ. *планетарий, stol* /s^tól/ ~ /istól/ ‘table’ < Russ. *cmol, stakán* ~ *istakón* ‘glass (for vodka etc.)’ < Russ. *стакан* (Rzehak 1999: 7); *tráktor* : /tiráktur/ : /tṛáktur/ ‘tractor’ < Russ. *mpák mop, iškóla* ‘school’ < Russ. *шкóла, ustál* ‘chair’ < Russ. *cmyn* (Kerimova 1997, 105).



		Yaghnobi	Sogdian	Roshani	Sariqoli	Yazghulami	Persian
1 sg.	nom.		əzú	az	waz	az	
	obl.	man	məná	mu	мы(n)	mũ(n)	man
2 sg.	nom.	tu	tʰyú	tu	tɛw	tow	
	obl.	taʉ	təwá	tā	ta, tɤ	tu	tu
3 sg.	nom.	ax	xō (m.) xā (f.)	γā	γɤ, γi	u, āy	
	obl.	áwi	wenē (m.) wya (f.)	way (m.) wum (f.)	wi	way, day (m.) im, dim (f.)	ōi, vaī
1 pl.	nom.		māx(u)				
	obl.	mōx	māx(i)	māš	maš	mox	mā
2 pl.	nom.		šmāx(u)	tama	tamaš	təmox	
	obl.	šmōx	šmāx(i)				šumá
3 pl.	nom.	áxtit	xā	wāδ	wod	if, dif	
	obl.	áutiti	wya	wuf	wef		ešán

TABLE 4. The nominative and oblique forms of personal pronouns in Yaghnobi, Sogdian, Roshani, Sariqoli, Yazghulami and Persian. (Note that in Sogdian, the ‘oblique’ of the 3rd sg./pl. is the genitive)

life’. However, other compound verbs in Tajik appear in Yaghnobi as normal simple verbs — e.g. *dáxšak* ‘to hurt’, but Persian *dard kardán* ‘to hurt’. From the Persian compound verb *gusēl kardán* the Yaghnobi simple verb *kʰsēlak* ‘to send’ was created; this is an early Persian loan in Yaghnobi. It is possible that in that period there were no compound verbs in Yaghnobi at all. Note also that in contrast to Tajik, the nominal part of a Yaghnobi compound verb is in the oblique case, when a verb is in the infinitive, but in the direct case in all other forms.

The Persian causative suffix *-ān-* > Taj. *-on-* was incorporated into the Yaghnobi verbal system, where *-ōn-* is always stressed (Hromov 1972: 95–6, 116). An analogous function of this suffix can also be documented in some of the Pamir languages. The Persian participial ending *-gī* is also borrowed into Yaghnobi, where its usage is the same as in Persian: cf. Yagh. *ētagī* (← past participle *ēta* to *šáwak* ‘to go’) and Pers. *raftagī* (← past participle *raftá* to *raftán* ‘to go’), Yagh. *iktagī* (← past participle *íkta* to *káarak* ‘to do’) and Pers. *kardagī* (← past participle *kardá* to *kardán* ‘to do’). As there is no difference in usage between Yaghnobi *ētagī* / *iktagī* and Persian *raftagī* / *kardagī*, this suffix can be seen as a direct grammatical loan from the surrounding Tajik dialects (cf. Hromov 1972: 45, 116).

The Yaghnobi verbal system was influenced in other ways by Tajik, specifically in two important constructions: the present and past progressive tense, and the (plu) perfect. In standard Tajik, the progressive tense is formed from the past participle with *istodá* (past participle of *istodán* ‘to stand’) and the conjugated verb ‘to be’, e.g. *man raftá istodá-am* ‘I’m going (right now)’,²³ *man raftá istodá budám* ‘I was going (right

²³ Shorter forms are used in the dialects, e.g. *man raftis(t)odám* in the Mastchoh and Falghar dialects, *man rafšodám* in Varzob dialects, *man raftestám* in Dushanbe, *man raftestúm* in Southern Tajik.

at that time)', but in the Tajik dialects neighbouring the Yaghnob valley the progressive is often formed with the infinitive and *doštán* 'to have', e.g. *man raftán dorám* 'I'm going (right now)', *man raftán doštám* 'I was going (right at that time)'. Not surprisingly, the Yaghnobi progressive is formed similarly to the second variant: *man šáwak ást(i)* 'I am going', *man šáwak ōy* 'I was going'.²⁴

Under the influence of the Persian language, the formation of some past tenses consisting of past participles together with forms of the verb 'to be' was introduced into Yaghnobi. The Tajik forms of the perfect *man raftá-am* 'I have gone' and pluperfect *man raftá budám* 'I had gone' are comparable to Yaghnobi *man éta-īm* and *man éta ōyim*, respectively. The Yaghnobi forms are formally the same as in Tajik, suggesting that the development of this grammatical feature was triggered by contacts with Tajik (Hromov 1972: 116–117); in contrast, Sogdian past tenses are formed with the help of the verb *δ'r-* /*δār-* 'to have' (see Gharib 1965). Under Tajik influence, Yaghnobi verbal morphology has come to look more like that of Western Iranian languages and differs considerably from the original Eastern Iranian type as in Sogdian (Hromov 1972: 116).

Another minor influence on Yaghnobi can be seen in the introduction of the Tajik verbal imperfective prefix to express duration/imperfectivity. The two languages express imperfectivity in different ways: in Yaghnobi, the imperfective is formed by the addition of the suffix *-išt* after the personal ending,²⁵ while in Classical Persian the imperfective was formed by adding the prefix (*ha*)*mē-* (Tajik *mē-*, Farsi *mī-*) before the verbal stem. Thus 'I am doing' is *man kun-ōm-išt* in Yaghnobi and *man*

24 Yaghnobi has no verb 'to have', but makes use of a phrase with the possessor in the oblique followed by the possessed object and copula *ást(i)* (*ōy* for past tense). Thus 'I have a horse' is expressed as *man ī asp ást(i)*; 'Hasan has hundred goats and twenty sheep' as *Hasáni sad vūzi-at bīst s'tūri ást(i)*; 'you (sg.) had five cows' as *taŋ panč yōwi ōy* and so on. The use of the copula for the verb 'to have' is similar to Russian *у меня (есть) лошадь* 'I have a horse' or Latin *mihi equus est* 'id.'; in this case it is an archaism in Yaghnobi, found in Sogdian as well, e.g. *mn' spy'sty /maná aspí (ə)stí/* 'I have a horse', *tw' ync'sty /tawá emj'ast(i)/* 'You (sg.) have a wife'. This feature is certainly not due to influence from the Uzbek/Turkic expression of possession: Uzbek *át-im bâr*, 'I have a horse', literally 'horse-my is' (cf. Turkish *at-im var*, Kyrgyz *at-īm bar*), cf. Northern Tajik *ásp-am ay* (< *ast*) instead of standard Tajik *man asp dórám* (Doerfer 1967: 56). This type of construction can be analysed as 'horse'+enclitic pronoun + 3 sg. copula; in Yaghnobi it rarely appears as *ásp-im ást(i)*.

25 The ending *-išt* undergoes changes in the endings of the 3rd persons, e.g. 3 sg. *-čī* < *-čīt* < **-tšīt* < *-t-išt* (cf. Klimčickij 1938: 99–100), 3rd pl. *-ōšt* < *-ō(y)išt* < *-ōr-išt* (cf. Klimčickij 1940, 100). For the ending of the 3 sg. *-čī* is the most commonly used form, but other forms have been recorded in some villages, e.g. Nomitkon *-čīt* (Klimčickij 1938: 99–100), Qūl *-čīš*, *-tši* (Andrejev — Livšic — Pisarčik 1957, 236). Hromov recorded the 3 sg. ending *-tišt* among speakers of the western dialect (Hromov 1972: 97), whereas Junker noted both *-tišt* (Junker 1930, 108) and *-či* (Junker 1914, 22). I have not heard the ending *-tišt* among speakers of the western dialect during my stays in Yaghnob in summer 2008 and spring 2009. The 3rd pl. ending *-ōšt* is common in all the dialects; only in the speech of Marghlimayn was the variant *-ōyšt* recorded (Klimčickij 1940, 100). Some hundred years ago Junker recorded the archaic form *-ōrišt* in the speech of the Yaghnobis (Junker 1930, 107).



mé-kun-am in Tajik. Due to the bilingualism of the Yaghnobis and their frequent use of Tajik in everyday communication, however, the Tajik prefix *mé-* may also occasionally be used to express the imperfective in Yaghnobi. In that case, a verb has two imperfective markers: *man mē-kun-ōm-išt*. The use of the imperfective prefix *mē-* is not common among Yaghnobi speakers and is generally considered an error. Its usage is also limited to areas where Yaghnobis speak Tajik more often than their own mother tongue and is mostly found among speakers of the younger generation. I personally have never observed this feature among the Yaghnobis living in the Yaghnob valley itself.

2.2.3. THE NOUN PHRASE

Turning to nominal morphology, one prominent feature borrowed from Persian is the *izafet* construction. The *izafet* construction appears in Yaghnobi within whole noun phrases taken from Persian: Yagh. *júft-i gōu*, Pers. *júft-i gāv* ‘pair of oxen (for ploughing)’; Yagh. *sōatō-yi čōrti*, Taj. *dar soathō-yi čor* ‘at four o’clock’ (Hromov 1972: 114);²⁶ Yagh. *Bōbō-yi Ōdām*, Pers. *Bābā-yi Ādām* ‘Grandfather Adam’. Possibly favouring this grammatical borrowing is that the *izafet* ending *-i* is formally similar to the Yaghnobi oblique case ending *-i*. For example, Yagh. *s^utūr-i yōta* ‘sheep’s meat’ is a construction with an oblique (sheep-OBL meat) and not an *izafet* construction (†sheep+*izafet* meat); cf. Pers. *gōšt-i gōsfānd* (meat+*izafet* sheep) and note the reversed word order in the Yaghnobi oblique phrase in contrast to the Persian *izafet*. Many Yaghnobis do not distinguish the *izafet* and the oblique in their own linguistic analysis: as the endings are both *-i*, i.e. phonetically similar, the interpretation of this phenomenon by native speakers can be influenced by their knowledge of Tajik. In my own analysis of Yaghnobi texts, it seems that the usage of the *izafet* construction is limited only to expressions adopted from Tajik; cf. the similar situation with the use of the *izafet* in Persian phrases in Pamir languages and Kurdish or in Uzbek, Qashqai, Ottoman Turkish and Central Asian Arabic. Another interesting feature concerns the phonological treatment of the *izafet* and the oblique case endings in Yaghnobi: in the older language, the oblique was realised as *-i* after a vowel but the *izafet* following a vowel was always realised as *-yi*, whereas nowadays both suffixes have merged and are realised as *-yi* by many speakers. A similar development both in Yaghnobi and in Zeravshan Tajik can be observed in nouns ending in *-a* followed by the *izafet* or the oblique: in this case *-a* is often replaced with *-i*, e.g. Zeravshan Tajik *xūnī man < xonā-yi man* ‘my house’ (cf. colloquial Teherani Persian *xūnē mān* instead of *xānē-ye mān*). In Yaghnobi, *-a* can also be replaced by oblique *-i*, so that a phrase such as ‘son’s house’ can be realised in four different ways: *žūtayi kat*, *žūti kat* (*-a* replaced by oblique *-i*), and the apocopated *žūtaj kat* (W), *žūtē kat* (E).

Another example of Tajik influence on Yaghnobi grammar is the use of the suffix *-akí* in *nomina agentis* derived from Yaghnobi (or Persian, Uzbek etc.) nouns. Thus Yagh. *waiš* (W), *wēš* (E) ‘grass’ plus the suffix *-akí* forms the noun *waišakí* (W), *wēšakí* (E) ‘a person carrying grass’ (Hromov 1972: 93, 116); similarly *aspakí* is a ‘horse-rider’

²⁶ This phrase is a calque of Rus. *часá в четы́ре*.

(*asp* ‘horse’), *paltarakí* ‘a load-carrying person’ (cf. Taj. *paltár* ‘a horse used for carrying load’) and so on.

Under the influence of Tajik, an unstressed enclitic *-ro* may be used in Yaghnobi to mark the direct object. In the Tajik literary language there is just one variant *-ro*, but in the colloquial language this is realised as *-a* following a consonant or *-ra* / *-ya* following a vowel, e.g. literary *man Hasán-ro méšinosàm, lekín Alí-ro námešinoxtà* ‘I know Hasan but I didn’t know Ali’ becomes *man Hasán-a méšinosàm, néki Alí-ra ~ Alí-ya námešinoxtà* in colloquial Tajik. In Yaghnobi the enclitic *-ro* occurs rarely, and is realised as *-a* or *-ya* in the two examples recorded: *nómíš-a apúrs* ‘she asked his name’; *tík póda-ya mayda^hák [v^hrót] hay akún* ‘and again the younger [brother] drove the flock’ (cf. Andrejev — Livšic — Pisarčik 1957: 223). The usage of the enclitic *-ro* in Yaghnobi is also limited due to the function of the oblique case, which can represent the direct object as well; thus the Tajik sentence *Ĵamšéd Širín-ro (~ Širín-a) nayz mébinàd* ‘Jamshed likes Shirin’ is normally translated as *Ĵamšéd Širín-i nayz wénči* in Yaghnobi, and *Ĵamšéd Širín-a nayz wénči* rarely appears. The two examples given above would have been *nóm-i-š apúrs* and *tík póda-ĭ mayda^hák hay akún* in “proper” Yaghnobi.

2.2.4. TURKIC INFLUENCE ON YAGHNOBI

The influence of Turkic grammar on Yaghnobi is slight. One feature I have recorded is the use of the unstressed interrogative particle *-mi* among some Yaghnobis. This particle is typical for forming questions in Turkic languages (Uzbek *-mi*, Old Turkic **-mi*; e.g. Uzb. *siz inglizča gaplašasiz-mi?* ‘do you speak English?’), but it is also often used in the northern dialects of Tajik, e.g. *šumó anglisī gap mézaneton-mī?* ‘do you speak English?’. In Yaghnobi the particle *-mi* is used mainly in the Zafarobod region, where the Yaghnobis live in intense language contact with Tajiks and Uzbeks. Many Yaghnobi speakers speak Tajik and Russian, but they also know Uzbek to some extent, and some local Tajiks and Uzbeks also know a little bit of Yaghnobi (mainly children). The use of *-mi* in Yaghnobi is the same as in Tajik or Uzbek, i.e. *-mi* is never stressed and is always cliticised to the end of the whole phrase: *š^hmóx anglisī gap déhtišt-mi?* ‘do you speak English?’.

The other feature borrowed via Tajik from Uzbek is the indefinite pronominal phrase with the Uzbek interrogative pronoun *kim* ‘what’, e.g. *kim-kád* ‘whenever’ (Taj. *kim-káy*), *kim-čó* ‘whatever’ (Taj. *kim-čī*) and so on.

2.3. LEXICAL BORROWINGS

The main evidence for language contact between Yaghnobi and the neighbouring languages comes from the lexicon. As mentioned above, many words have been borrowed from Tajik, Arabic, Uzbek and Russian. At present, no accurate analysis of the origin of the Yaghnobi lexicon is available. As an estimate based on etymological analysis for the Yaghnobi — Czech dictionary (Novák 2010), I would say that some 34% are of Tajik origin, 8% of Arabic origin, 3% from Uzbek or another Turkic language and approximately 3% from Russian (including many international expressions in-



roduced through Russian). About 6% of lexemes are Yaghnobi-Tajik or Yaghnobi-Arab compounds, 19% are compound verbs and only 27% seems to be native Yaghnobi, i.e. of Eastern Iranian origin. There are also borrowings from other languages, e.g. the ancient Greek word *δραχμή* via Persian *dir(h)ám* has become Yagh. *d'rám*. Many old borrowings from Greek and later from Sanskrit, Hindi/Urdu, Chinese and Mongol were introduced into Yaghnobi in their Persian form. Similarly, several German and English words have been introduced into Yaghnobi via Russian: *buyáltir* 'clerk' ← German *Buchhalter* via Russ. *бухгалтер*; *š'iláng* 'water hose' ← German *Schlange* 'snake' via Russ. *шланг* 'hose'; or Yagh. *kamp'yúter* ← English *computer* via Russ. *компьютер*.

Through the study of borrowings in Yaghnobi, we can see that nearly all loanwords were introduced through Tajik. An analysis of the 139 words of Turkic origin contained in the dictionary of Andrejev — Livšic — Pisarčik (1957) reveals that 123 also appear in literary Tajik, a further seven words can be found in neighbouring Tajik dialects and only nine Uzbek words are not found in Tajik dictionaries (Doerfer 1990). An interesting fact is that the proportion of genuine Yaghnobi words in the lexicon of approximately 7600 units gathered by the author in the years 2007–2009 is very small (Novák 2010). Native vocabulary comprises approximately 27% of the lexicon words, while compound verbs and Yaghnobi-Tajik/Arabic compounds make up 25%. The latter are often calques from Tajik formed from Tajik/Arabic words together with Yaghnobi words or suffixes. The question is whether these compound verbs and calques can be regarded as Yaghnobi words, but in many cases those new compounds are unintelligible to Tajiks. Other parts of the lexicon can likewise be recognised as loans (up to 48%).

For comparison, some 46,5% of the vocabulary of modern literary Farsi is of Arabic origin (Perry 2002). The lexicon of Uzbek is also strongly influenced by Tajik and Arabic: the ratio of Turkic, Persian/Arabic and Russian words in a modern Uzbek text translated from Russian is 56:31:13, and the extent of Persian/Arabic vocabulary in Uzbek ranges from 25 to 45%, whereas in Chaghatay/Old Uzbek the ratio of Persian/Arabic to Turkic lexicon was approximately 2:3 (Doerfer 1990). On the other hand, Kyrgyz and Kazakh do not have so many borrowings as Uzbek, and modern Uyghur likewise has relatively fewer Persianisms or Arabisms. In the Pamir languages and in Pashto there are also many Persian, Arabic and Turkic loans, but not to the same extent as in Yaghnobi.

Many borrowings have been adapted to Yaghnobi phonology. The main changes are the insertion of a svarabhakti vowel in word-initial consonant clusters and the treatment of word-medial/final clusters *-CjV(-). Other adaptations to native pronunciation may be observed, e.g. Persian *zambór* 'wasp' was rendered as *zambúr*, with Tajik and Uzbek *ú* (Taj. *ú* [ø:] < *ō, Uzb. *ú* [ø ~ ø] < *ō, *ö) merging with native Yaghnobi *ū* (due also to the historical development *ō > *ū* in Yaghnobi and in the Zeravshan dialects of Tajik).

As an example of Yaghnobi-Tajik language contact, we can identify three phases in the process of introduction of Tajik words: 1) in the first phase, voiced stops were adopted as their unvoiced equivalents, cf. *kūr* ← Pers. *gōr*; 2) in the probable second stage, Persian *δ* underwent the Yaghnobi change *δ > *d*, cf. *g^udárak* ← Pers. *gudāštán*; and 3) in the last period, Tajik words were adopted in close resemblance to the origi-



nal form. To explain why I distinguish between phases 2 and 3, note that the earlier borrowings of Persian verbs were fully adapted to the Yaghnobi grammatical system, so verb stems ending in *-ar-* change to *-ōr-* when the ending **-θ(-)* (i.e. *-t(-) ~ -s(-)*), *-t(-)* or *-či*) is added; but in later borrowings this does not happen. Hence we find Yagh. *váarak* ‘to carry’, **vórtišt ~ *vórsišt* ‘you (pl.) carry’; *g^udáarak* ‘to pass by’, *g^udórci* ‘he passes by’; but *páarak* ‘to fly’, *párci* ‘he flies’ (not *ʃpórci*), as the verb *páarak* is a relatively new borrowing from Pers. *parrídán* : *parr-* (Hromov 1962: 13).

An interesting phenomenon may be seen in the case of Yaghnobi numerals. The native numerals from ‘1’ to ‘10’ have been preserved, but the full range of Tajik numerals was also borrowed (with Arabic and Russian loans for the number ‘zero’ and, via Tajik, Russian names for ‘million’, etc.). The use of Yaghnobi numerals is limited to counting things up to 10 units, whereas Tajik numerals are used when counting weights and lengths, when speaking about time and when counting in numbers greater than 10. With native numerals from ‘2’ up to ‘10’, the oblique case is used with the counted subject (*ī vuz* ‘one goat’, *du yōw-i* ‘two cows’, *t^ufōr mórti-i* ‘four men’, *avd s^utūr-i* ‘seven sheep’); this also happens when Tajik numerals are used (*šiš rúz-i* ‘six days’, *hašt sōát-i* ‘eight o’clock’, *pōnzdá^h yáyk-i* ‘fifteen girls’, *bíst-u se kⁱtōb-i* ‘twenty-three books’, *pinjō^h-u yaq táxm-i* ‘fifty-one eggs’, *sad sōl-i* ‘a hundred years’).

In some cases we can document the process of borrowing new words into Yaghnobi via Tajik over the last hundred years, whereby archaic words are no longer used and have been replaced by new words from (colloquial) Tajik, e.g. *man* ‘apple’ > *sēb*; *kⁱmēr* ‘red’ > *surx*; *zērta* ‘yellow’ > *zard*; *šōu* ‘black’ > *sⁱyō^h*; *rōut* ‘river’ > *dayrō ~ dar(i) yō*; *dōn* ‘seed’ > *dōna*; *yaγd* ‘wide’ > *pahm*; *rītistar* ‘before’ > *pēštár* (Bogoljubov 1966: 359); *ipōrá* or *yalbalá* ‘much, many’ > *bis(i)yōr* (Klimčickij 1940b); *ētk* ‘bridge’ (documented in texts but nowadays rarely used even in the Yaghnob valley) > *kūprúk* (cf. coll. Taj. *kūprúk* < Chaghatay *kōprúk*, Uzb. *kūprik*) or *most* (← Russ. *мост*);²⁷ *v^urúk* / *v^urýk* ‘eyebrow’ > *qōš* (cf. coll. Taj. *qōš* < Uzb. *qāš*); *šī* ‘upper’ > *bōlō*; and often *mēs* (E) : *mēt* (W) ‘day’ > *rūz*. Some words that also existed in Sogdian were later changed to “fit” the corresponding Tajik forms: Sogd. *θyš-*/θaxš- / ‘to forgive’* > *baxš-*; Sogd. *θⁱy^{*} / θáy / ‘garden’* > *bōy*; Sogd. *mⁱy^{*} / máx / ‘month’* > *mōh* (Bogoljubov 1966: 359); Sogd. *θym^{*} / θím / ‘fear’* > *bim*. Many original Yaghnobi prepositions, postpositions and conjunctions have also been recently replaced by Tajik ones (Hromov 1972: 115–6).

To document each borrowing into Yaghnobi is a task for future studies. At present, a more accurate study of the structure of the borrowed lexicon is the most urgent task. With increasing collections of reliable data, much more work can now be done.²⁸

²⁷ There is a native Persian word for ‘bridge’: Taj. *pul* (Pers. *pul*; Farsi *pol*), but this is not used very often in the colloquial language, perhaps because of confusion with the homophonous word *pul* ‘money’ < Pers. *pūl* (Farsi *pūl*). Instead, the word *kūprúk* of Uzbek origin is normally used; the younger generation also often uses the Russian word *most*. In Yaghnob only *kūprúk* is used (*ētk* in a very limited number of cases), while *most* can be used in the Lower Varzob area and in Dushanbe.

²⁸ I have tried to include some basic etymological information on Yaghnobi words in my Yaghnobi-Czech Dictionary (Novák 2010), but a full etymological study of Yaghnobi remains a task for the future.



3. YAGHNOBI AS A CONTACT LANGUAGE

In the previous paragraphs I have tried to outline some basic features shared by Yaghnobi with Tajik. Several core domains of Tajik influence on Yaghnobi may be identified, although the task is complicated by several factors. Above all, it is difficult to examine borrowings from other language(s) in the absence of an up-to-date Tajik etymological dictionary. For Tajik etymological studies, the *Farhang-i zabon-i tojikī* (Šukurov — Kapranov — Hošim — Maʼsumī 1969) can be helpful for basic information about words of Arabic and Turkic origin.

It may be supposed that Yaghnobi was originally spoken not only in the Yaghnob valley itself but, as mentioned above in §1, also in some other areas of north-western Tajikistan and perhaps also in the adjacent areas of Uzbekistan. Some Tajik place names appear to be of non-Persian Iranian origin, and were later adapted to the phonology of Persian. Unfortunately, we cannot tell whether those place names are of early Yaghnobi or Sogdian origin since those languages were close to one another.

To take one example, the Tajik name of the river *Varzób* should be of pre-Persian, probably Sogdian, origin. Its form */βəʾz-āp/ 'high water' shows the Sogdian development of Iranian *br̥z- (Av. *bərəz-* : *barəz-*; cf. OPers. *br̥d*° in the personal name *Bṛdiya-*) to Sogd. *βrz-* */βəʾzi/ in contrast to Persian *burz* or *bul[ānd]* (Horn 1988: 46; Gharib 2004: 111), while Sogdian *p */āp/ (Yagh. *ōp*) has been replaced by Persian *āb* (Taj. *ob*). Similarly, we can detect Yaghnobi/Sogdian elements in names such as *Anzób* (Iranian *anzū-āpa- 'narrow water'; cf. Sogd. 'nz'βH), *γarm* (Sogd. *γrm* 'warm'; cf. Pers. *garm*), *Iskodár* (Sogd. 'sk'tr), *Urmetán* (Sogd. *'wr-myδn), *Farmetán* (Sogd. *prnmyδn*), *Falyár* (Sogd. *pryrh*), *Madm* (Sogd. *mδmh*), *Kum* (Sogd. *kwm*), *Dary* (Sogd. *δrγH*), *Varz(-i Minor)* (Sogd. *βrz-*; today *Aynī*), *Rarz* (Sogd. *rzrh*), *Poxút* (Sogd. *p'γwt*), *Fatmév* (Sogd. *θtmyβH*, 'θtmyβH), *Falmoút* (Sogd. 'θm'twt, Yagh. *Fatmō'ūt*), *Xušekát* (Sogd. 'γsyknδH, *γsyknδH*, 'γsykt), *Panjaként* (Sogd. *pn̄cyknδH*, *pn̄cknδH*, Yagh. *Panjikát*) and many others (cf. Bogoljubov — Smirnova 1963: 101–108).

The comparison of the lexicon of Yaghnobi and Tajik reveals other similarities. The Tajik word *nayz* 'pretty, nice' is considered to be of Sogdian origin. According to *Farhang-i zabon-i tojikī*, this word first appears in Firdausi's *Shahnameh* and is regarded as native Persian/Tajik (Šukurov — Kapranov — Hošim — Maʼsumī 1969: 843); for a Sogdian etymology see Gharib 2004: 238; Andrejev — Livšic — Pisarcik 1957: 289. Several other words have the same form in Yaghnobi and Tajik, but they differ from the Persian form. To consider the famous case of Taj. and Yagh. *asp*, Sogd. 'sp- */aspí/ vs. Pers. *asb* 'horse', there is a question whether the Tajik form might show influence of the earlier non-Persian pronunciation (cf. also Av. *aspa-*, Khwarezmian 'sp */asp/, Bactrian *ασπο* */aspɔ/, Yazghulami *asp*, Munji and Yidgha *yasp*). Note also that many archaic words of Sogdian origin can be found in the Persian dictionary *Luyāt-i Furs* by Abūmañšūr 'Alī ibn-i Aḥmad Asadī-yi Tūsī. Some of the words recorded in this lexicon are still found in Yaghnobi, while others are used in the Tajik dialects of Mastchoh and Falghar (Hromov 1962: 83–4).

Finally, the similarities in the phonology of Zeravshan Tajik and Yaghnobi can be interpreted in two ways: early Yaghnobi could have influenced the adjacent Tajik dialects or *vice versa*. The striking similarities in the development of Central Tajik and



Yaghnobi **ō* > *ū* and Yagh. **ū* > *ÿ* and Zeravshan Tajik *u* > *u*, as well as Tajik and Yaghnobi **ā* > *ō*, may point to language contacts in earlier periods (cf. Ido 2009).

In the Yaghnob valley there are now some Tajik-speaking villages whose inhabitants no longer use the Yaghnobi language, but whose speech is quite different from neighbouring dialects of Tajik. The Tajik dialect of Yaghnob belongs to the Zeravshan dialect group of Tajik, to which it is relatively close: they share similarities in phonology, but the inhabitants of the Yaghnob valley use Yaghnobi-based lexicon in part, which is unintelligible to those outside the valley. In the following examples, words of Yaghnobi origin are underlined: *tīs*, *aŕsoná-ra ȳūš dór* ‘enter, listen to the fairy-tale’ (Taj. *daró(ȳ)*, *afsoná-ro ȳūš dor*; Yagh. *tīs*, *ōfsónaiȳ ȳūš dór*); a *Nómitkon úrū avésom mégūt* ‘we set off that side from Nomitkon, he says’ (Taj. *az Nomitkón on rú(ȳ) furóma-dem*, *mégúyad*; Yagh. *či Nómítkōn áwi nēmaiȳ avésōm*, *wóči*). This dialect is also used by Yaghnobi speakers when they communicate among themselves or with strangers. The main difference between this Yaghnob-Tajik and the other Zeravshan dialects is stress, which often shifts to the penultimate syllable if this is historically long, e.g. *xúna* / *xóna* vs. Taj. *xoná*, Zerav. *xúná* ‘house’. Yaghnob-Tajik shares this feature with Yaghnobi, where stress also falls on a historical long vowel in the root of a verb, e.g. *čorúq medúzad* (Taj. *čorúq médúzad*) ‘he sews shoes’; *gūspánd-a mebinand* (Taj. *gūsfánd-ro mebinand*) ‘they see the sheep’. The 1st person plural ending is not *-em* as in Tajik but *-im*, which corresponds to Yaghnobi *-im*. Also, the 1st and 3rd person singular enclitics *-im* and *-iš* (same as in Yaghnobi) are used instead of Tajik *-am* and *-aš*. Further study of the Yaghnob-Tajik dialect is urgently needed, as until today only a short outline with texts and a dictionary has been compiled by Hromov in his dissertation (Hromov 1969: 305–23, 327–85, see also Hromov 1972: 118–9).

4. DOMAINS OF CONTACT

The Yaghnobi language area can be divided into seven main groups:

a) The Yaghnob valley itself. Today the Yaghnobi-speaking population lives in 12 villages. Four other villages in the valley are inhabited by Tajik-speaking Yaghnobis, while 14 villages have been abandoned since the 1970s. Following the forced migration of the Yaghnobis in 1970–71, some of the people returned back to their homeland after 1989.

Before the beginning of the 20th century, we can suppose that the Yaghnob valley was populated entirely by Yaghnobi speakers. Later on, some villages of the area became Tajik-speaking and their inhabitants stopped using Yaghnobi completely. At present, the situation in the valley itself is quite complicated: when the people were forced to leave their homes and move to Zafarobod, the valley was deserted. From the late 1980s onwards, some of the Yaghnobis returned to their homes in the valley. Today, Yaghnobi is spoken in everyday conversation; the Tajik language may also be used. Tajik proficiency is higher among men in comparison to women and children, but almost everyone in the valley speaks both Tajik and Yaghnobi. The domain of



Tajik is limited: it is mainly used when travelling outside of the valley or when speaking with a stranger. Contact with Tajik is quite limited for these speakers, and people who do not travel out of the valley do not have much contact with Tajik speakers. The other source for Tajik language contact is television, owned by some families in Yaghnob (Paul et al. 2005: 79–82). Knowledge of Russian is very low among the younger inhabitants of the valley and among women, as there are no schools in Yaghnob and education is provided mainly by women.

b) Upper Varzob, i.e. the valleys of Ziddeh and Takob. In this area there are four Yaghnobi-speaking villages. As in the Yaghnob valley, the Yaghnobi speakers in the Upper Varzob are also quite isolated from the outside world. This fact helps to preserve the language in its full form. Yaghnobi is the dominant language in these villages, and its domain appears to be similar to that of the language in the Yaghnob valley. On the other hand, proficiency in Tajik is higher than in Yaghnob due to the location of Yaghnobi speakers in a Tajik-language area, where possibilities for contact are not so limited. In the Upper Varzob area there are some schools, which also leads to a better knowledge of Tajik among the Yaghnobis (see Table 5.; Paul et al. 2005: 82–6). The inhabitants of Kûkteppa have more intensive contacts with the Yaghnob valley, as there is a path through a mountain pass, through which a traveller can reach the village of Qûl or Gharmen after a day's walking.

Tajik domains	Education (school and university) Religious ceremonies Speeches and formal occasions Television and radio broadcasts Written literature Trade
Yaghnobi domains	Conversation with friends and family at home Conversation with colleagues within the local area Oral literature (e.g. spoken poems and stories) Instruction to local workmen, children and so on

TABLE 5. Domains of language use for Zumand, Upper Varzob (after Paul et al. 2005: 86)

c) Zafarobod, the area to which the whole Yaghnobi population from the Yaghnob valley was moved in 1970–71. The Yaghnobi community in this area appears to be the largest in Tajikistan. This contributes to the linguistic viability of the population, but there are also many Tajik and Uzbek speakers who play their part in language contact. Here Yaghnobi is used mainly at home or when talking with friends, and Tajik when communicating with a stranger. In this area there are many schools with education in Tajik, and proficiency in Tajik is higher than in the areas of the Upper Varzob and Yaghnob. On the other hand, the Yaghnobis who work at home do not have as much contact with Tajik as those who work outside of their homes (see Table 6; Paul et al. 2005: 87–9).

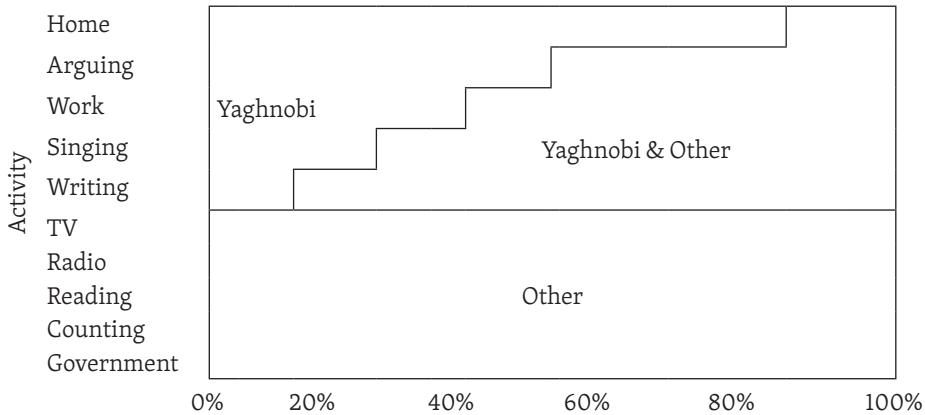


TABLE 6. Domains of language use in the Zafarobod area (after Paul et al. 2005: 89).

d) Lower Varzob: there are some Yagnobi speakers living in various villages of the Lower Varzob. The village Dughoba-yi Bolo may be regarded as one of the largest Yagnobi-populated localities in the area, together with Varzob-GĔS on the opposite side of the river Varzob. The population in Dughoba-yi Bolo and Varzob-GĔS is of Tajik and Yagnobi origin. Among the Yagnobis, their mother tongue is used in everyday life with the family or when talking to Yagnobi guests. On the streets, Tajik is the main language of conversation, but Yagnobi can be used whenever the audience would understand it. There is a school in Varzob-GĔS for pupils from both villages, but the Yagnobi language is not taught there. Due to the proximity of Dushanbe, the situation of Yagnobi speakers differs somewhat from the situation described above. Proficiency in and usage and accessibility of Tajik is even greater than in Zafarobod (Paul et al. 2005: 89–90).

e) The city of Dushanbe: nowadays many Yagnobis live in the capital of Tajikistan. According to the data available, the Yagnobi community in Dushanbe is located mainly in the north-eastern part of the city. The community is quite homogenous, but Yagnobi is limited mainly to family conversation, the Tajik language being the main language of communication. In Dushanbe there are also many other ethnic groups such as Russians, Uzbeks, Pamiris, and Roma. Russians use their language on an everyday basis (they often do not understand or speak any Tajik), while other ethnic groups do not often use their language outside of their homes.

f) The Hisor and Kofarnihon valleys: when we look at a map representing the areas populated by Yagnobis (Figure 1), an area in south Tajikistan seems to be largely populated by people of Yagnobi origin. This is because of the relatively large spread of Yagnobis in the towns and villages in the districts of Rūdākī, Hisor, Vaḥdat, Shaḥrinav and Yovon. However, the Yagnobi communities in those districts are quite small, and they are a minority among the Tajik population. Their use of Yagnobi is limited: the language is spoken mainly at home, and its use on the street is rare. In



the southern areas populated by Yaghnobis (i.e. the Hisor and Kofarnihon valleys and the city of Dushanbe), Tajik influence can be of greater extent than in areas a) to d).

g) Ghonchī: in this district there are some Yaghnobis living in eight villages, who count themselves as descendants of people from the Yaghnob valley; unfortunately, they do not use the language any more. An interesting feature is that in Ghonchī some of the Yaghnobi-populated villages have Yaghnobi names. This can be observed only in areas settled by Yaghnobis in an older period: apart from the Yaghnob valley itself, Yaghnobi place names are otherwise recorded only in the Upper Varzob area, whereas other places populated by Yaghnobis have Tajik or Uzbek names. According to Buzurgmehr, the Ghonchī region may have been settled by Yaghnobis in the 16th-17th centuries (Buzurgmehr 2005: 121).

5. CONCLUSION

Throughout its history, the Yaghnobi language has been influenced by neighbouring languages, the main influence coming from Tajik. The Yaghnobis live their lives in quite intensive contact with Tajik; almost every Yaghnobi also speaks Tajik, though proficiency in Tajik differs from region to region, depending on the need to use Tajik in everyday life. Yaghnobi shares some features in its historical development with the Central Tajik dialectal group in particular. This Yaghnobi-Tajik contact has been so intense that it has affected not only the vocabulary of Yaghnobi but also some aspects of its grammar, as discussed above in §2. On the other hand, Yaghnobi has retained its own linguistic structure, and borrowings that do not fit its structure were adapted phonologically and morphologically, at least until recently.

The Yaghnobi lexicon is composed of some 27% native vocabulary, while the remainder are loanwords from other languages, mainly from Tajik, Arabic and Uzbek. However, these numbers need to be treated with caution, as my analysis was done on material from various sources and no precise etymological information was available for some of the languages involved, including Tajik. It is also important to note that not every word of foreign origin is used frequently. Tajik and especially Arabic words are connected with legal, religious and educational domains, and their usage in daily life is less frequent. When I was learning the language as it is spoken in the Lower Varzob and in the Yaghnob valleys, I came to understand that the usage of loanwords is not as high as it appears in the available lexica. Studies of lexical entries encompass all the words used by Yaghnobis in various (often limited) occasions (cf. Novák 2010), so a study of the frequency of the lexemes within spoken corpora is needed to draw more accurate conclusions.

As mentioned above, Tajik is dominant in many areas in which the use of Yaghnobi is very limited. For example, Yaghnobis often write in Tajik as Yaghnobi does not have an established written form, and many Yaghnobis did not have any opportunity to learn their native language at school. There are also few books in Yaghnobi; furthermore, radio broadcasting in Yaghnobi is rare, and there are no television programs in Yaghnobi. This situation is not conducive to the long-term survival of

the language. There are some attempts to print books about and in Yaghnobi, but this movement is still in its early stages. Yet the outlook appears to have improved in recent years for the preservation of the language. Most importantly, the language seems to have great prestige among its speakers. This is also supported by official recognition of Yaghnobi and the Pamir languages in the Tajik constitution as a part of the cultural heritage of Tajikistan.

To summarise, the Yaghnobi language is influenced by Tajik in many respects, but it still retains its own distinct features that make it so different from Tajik, so that monolingual Tajiks often do not understand even if a Yaghnobi is using Tajik vocabulary within a Yaghnobi sentence. The core of the Yaghnobi lexicon and grammar, such as the pronominal and verbal system, remain distinct from that of Tajik, and there are no signs that this situation will change in the immediate future.

LANGUAGE ABBREVIATIONS:

Arab.	Arabic
Av.	Avestan
Iran.	Iranian
OPers.	Old Persian
Pers.	Persian
Russ.	Russian
Sogd.	Sogdian
Taj.	Tajik
Uzb.	Uzbek
Yagh.	Yaghnobi
Zerav.	Zeravshan dialects of Tajik
(E)	eastern dialect of Yaghnobi
(W)	western dialect of Yaghnobi

REFERENCES

- Andrejev — Livšic — Pisarčik (1957) М. С. Андреев — В. А. Лившиц — А. К. Писарчик: *Словарь*, in: М. С. Андреев — Е. М. Пещерева (eds.), *Ягнобские тексты с приложением ягнобско-русского словаря составленного М. С. Андреевым, В. А. Лившицем и А. К. Писарчик*. Москва — Ленинград: Издательство Академии Наук СССР, 215–391.
- Andrejev — Peščereva (1957) М. С. Андреев — Е. М. Пещерева: *Ягнобские тексты с приложением ягнобско-русского словаря составленного М. С. Андреевым, В. А. Лившицем и А. К. Писарчик*, Москва — Ленинград: Издательство Академии Наук СССР, 215–391.
- van der Ark, R. — P. Menecier — J. Nerbonne — F. Manni (2007) Preliminary Identification of Language Groups and Loan Words in Central Asia, in: P. Osenova et al. (eds.), *Proceedings of the RANLP Workshop on Computational Phonology Workshop at the conference Recent Advances in Natural Language Phonology*, Borovetz: RANPL, 13–20.





- Bashiri, I. (1997) *The Yaghnobis*. [online; <http://www.angelfire.com/rnb/bashiri/Tajling/yaghnob.html>], cited 08. 11. 2008]
- Benveniste, E. (1955) Un lexique du Yagnobi. *Journal Asiatique* CCXLIII/2, 139–162.
- Bielmeier, R. (1989) Yaghnōbī, in: R. Schmidt (ed.), *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert, 480–488.
- (2006) Yaghnobi, in: E. Yarshater (ed.): *Encyclopædia Iranica*. [online; http://www.iranica.com/newsite/articles/unicode/ot_grp10/ot_yaghnobi_20060303.html], cited 18. 10. 2007]
- Bogoljubov (1956) М. Н. Боголюбов: *Ягнобский (новосогдийский) язык. Исследование и материалы*. Автореферат на соискание учёной степени доктора филологических наук. Ленинград: Ленинградский университет.
- (1966) Ягнобский язык, in: В. В. Виноградов (ed.): *Языки народов СССР. Том первый: Индоевропейские языки*. Москва: Наука, 342–361.
- Bogoljubov — Smirnova (1963) М. Н. Боголюбов — О. И. Смирнова: *Хозяйственные документы. Согдийские документы с горы Муз. Чтение. Перевод. Комментарий. Выпуск III*. Москва: Издательство восточной литературы.
- Borovkov (1959) А. К. Боровков: *Краткий очерк грамматики узбекского языка*. In: А. К. Боровков — С. Ф. Акабиров — З. М. Магруфов — А. Т. Ходжахонов: *Узбекско-русский словарь. Ўзбекча-русча лугат*. Москва: Государственное издательство иностранных и национальных словарей, 677–727.
- Vuzurgmehr (2005) Б. Вузургмеҳр (Бурхониддин Тураев): *Ягнобиёни муқими Душанбешаҳр ва музофоти он*, in: Ю. Шодипур — А. Абдуллоев: *Душанбе дар масири таърих (Маҷмуаи мақолаҳо)*. Душанбе: Студент, 117–128.
- Doerfer, G. (1967) *Türkische Lehnwörter im Tadschikischen*. Wiesbaden: Harrassowitz.
- (1990) Central Asia. xiv. Turkish-Iranian Language Contacts, in: E. Yarshater (ed.): *Encyclopædia Iranica* [online; URL: <http://www.iranica.com/newsite/articles/unicode/v5f2/v5f2a017.html>], cited 25. 08. 2009]
- Gauthiot, R. — E. Benveniste (1929) *Essai de grammaire sogdienne. Deuxième partie: Morphologie, syntaxe et glossaire*. [Mission Pelliot en Asie centrale: Série petit in-Octavo, 3], Paris: Geuthner 1929.
- Geiger, W. (1898–1901) Über das Yaghnōbī, in: W. Geiger — E. Kuhn (eds.): *Grundriss der iranischen Philologie*, Erster Band, 2. Abteilung. Straßburg: Trübner, 334–344.
- Gharib, B. (1965) *Analysis of the Verbal System in the Sogdian Language*. Diss. University of Pennsylvania.
- (2004) *Sogdian Dictionary (Sogdian — Persian — English)*. Tehran: Farhangian Publications. (بدر الزمان قریب: فرهنگ سغدی (سغدی — فارسی — انگلیسی). تهران: فرهنگان، ۱۳۸۳)
- Horn, P. (1988) *Grundriss der neupersischen Etymologie. Sammlung indogermanischer Wörterbücher*. Hildesheim — Zürich — New York: Georg Olms Verlag.
- Hromov (1962) А. Л. Хромов: *Говоры таджиков Матчинского района. Гуишҳои тоҷикии райони Мастчоҳ*. Труды, т. CVII. Душанбе: Издательство Академии наук таджикской ССР.
- (1969) *Историко-лингвистическое исследование Ягноба и Верхнего Зеравшана*. Диссертация на соискание учёной степени кандидата филологических наук. Душанбе 1969.
- (1972) *Ягнобский язык*. Москва: Наука.
- (1976) Ягнобские архаизмы в топонимических названиях, in: В. С. Расторгуева (ed.): *Иранское языкознание*. Москва: Наука, 160–167.
- (1987) Ягнобский язык. In: В. С. Расторгуева (ed.): *Основные иранского языкознания. Новоиранские языки II. — Восточная группа*. Москва: Наука, 644–701.
- Ido, S. (2009) An analysis of the formation of the Tajik vowel system, in: A. Saxena — Å. Viberg (eds.): *Multilingualism. Proceedings of the 23rd Scandinavian conference of linguistics*. [Acta universitatis Upsaliensis: Studia Linguistica Upsaliensia 8.], Uppsala: Uppsala University, 65–74.

- Jefimov — Rastorgujeva — Šarova (1982) Валентин Александрович Ефимов — Вера Сергеевна Расторгуева — Е. Н. Шарова: Персидский, таджикский, дари. In: Вера Сергеевна Расторгуева (ed.): *Основы иранского языкознания. Новоиранские языки I. — Западная группа, прикаспийские языки.* Москва: Наука 1982, pp. 5–230.
- Junker, F. J. (1914) *Drei Erzählungen auf Yaghnābī.* Heidelberg: Carl Winter.
- . (1930) *Arische Forschungen. Yaghnōbī-Studien I. Die Sprachgeographische Gliederung des Yaghnōb-Tales.* Leipzig: Hirzel.
- Kerimova (1997) А. А. Керимова: Таджикский язык, in: *Языки мира. Иранские языки I. — Юго-западные иранские языки.* Москва: ИНДРИК, 102–133.
- Klimčickij (1935a) С. И. Климчицкий: Название Согдианы в топонимике Таджикистана, in: *Записки института востоковедения академии наук VI.*, 9–13.
- . (1935b) Ягнобско-согдийские соответствия, in: *Записки института востоковедения академии наук VI.*, 15–25.
- . (1940a) Ягнобская сказка, in: *Академия наук СССР — Труды Таджикистанской базы, т. IX — 1938 — История — язык — литература. Akademijaji Fanho SSSR: Asarhoji bazaji Točikiston, čildi IX — Tarix — zabon — adabijot.* Москва — Ленинград: Издательство Академии наук СССР, 94–103.
- . (1940b) Ягнобцы и их язык, in: *Академия наук СССР — Труды Таджикистанской базы, т. IX — 1938 — История — язык — литература. Akademijaji Fanho SSSR: Asarhoji bazaji Točikiston, čildi IX — Tarix — zabon — adabijot.* Москва — Ленинград: Издательство Академии наук СССР, 137–139.
- Livšic (2003) В. А. Лившиц: Согдийские документы из замка Чильхуджра, in: *Scripta Gregoriana. Сборник в честь семидесятилетия академика Г. М. Бонгард-Левина.* Москва: Восточная литература 2003, 77–88.
- Livšic — Hromov (1981) В. А. Лившиц — А. Л. Хромов: Согдийский язык, in: В. С. Расторгуева (ed.): *Основы иранского языкознания. Среднеиранские языки I.* Москва: Наука, 347–514.
- Loy, T. (2005) *Jaghnoḅ 1970. Erinnerungen an eine Zwangsumsiedlung in der Tadschikischen SSR.* Wiesbaden: Reichert.
- Mallickij (1924) Н. Г. Маллицкий: Ягнобцы, in: *Известия Туркестанского отдела Географического общества (Ташкент), Том XVII.* 174–178.
- Milovskij (1991) А. Миловский: Последние из шестнадцатой сатрапии. *Вокруг света* 1/ январь 1991; [online: <http://www.vokrugsveta.ru/vs/article/2279/> cited 28. 09. 2009, 11:28]
- Mirzozoda (2007) С. Мирзоев: *Суннатҳои маҳалли мардуми Ягноб.* Душанбе: Деваштич.
- . (2008) *Фарҳанги ягнобӣ-тоҷикӣ.* Душанбе: Деваштич.
- Mirzozoda — Alavī (2008) С. Мирзозода — Б. Алави: *Дастури забони ягнобӣ. Ягнобӣ зивои дастур.* Душанбе: Деваштич.
- Mirzozoda — Qāsemī (1995) С. Мирзозода — М. Қосимӣ: *Фарҳанги забони ягнобӣ.* Душанбе: Пайванд (مسعود — میرزا زاده سيف‌الدين ميرزا زاده). (قاسمی: فرهنگ زبان یغناپی. دوشنبه: پیوند، ۱۳۷۴).
- Novák, L. (2010) *Jaghnōbsko-český slovník s přehledem jaghnōbské gramatiky. Ягнобӣ-чехӣ луғат ягнобӣ зивои дастури феҳрастони.* Praha: Filozofická fakulta Univerzity Karlovy v Praze.
- Panfilov (1991) О. Панфилов: Ягноб акунси вайрон. *Фарҳанг* 7/1991, 45–50.
- Paul, D. — Abbes, E. — Müller, K. — Tiessen, C. (2005) *The Ethnolinguistic Vitality of Yaghnobi*, in: John M. Clifton (ed.): *Studies in languages of Tajikistan.* Dushanbe — St. Petersburg: SIL, 65–105.
- Perry, J. R. (2002) *Arabic Elements in Persian*, In: E. Yarshater (ed.): *Encyclopædia Iranica*, [online: http://www.iranica.com/newsite/articles/unicode/sup/Arab_Elements.html, cited 25. 08. 2009]
- . (2005) *A Tajik Persian Reference Grammar.* Leiden: Brill (Handbuch der Orientalistik 11.).
- Rastorgujeva (1952) В. С. Расторгуева: *Очерки по таджикской диалектологии.* Вып. 1 — *Варзобский говор таджикского языка.* Москва: Издательство Академии наук СССР.
- . (1964) *Опыт сравнительного изучения таджикских говоров.* Москва: Наука.



- Rzehak, L. (1999) *Tadschikische Studiengrammatik*. Wiesbaden: Reichert.
- Sims-Williams, N. (1989) Sogdian, in: R. Schmitt (ed.): *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert, 174–192.
- Sokolova (1953) В. С. Соколова: Ягнобский язык, in: В. С. Соколова: *Очерки по фонетике иранских языков*. Выпуск II. *Осетинский, ягнобский и памирские языки*. Москва — Ленинград: Издательство Академии наук СССР, 59–79.
- Steingass, F. J. (1892) *A Comprehensive Persian-English Dictionary including the Arabic words and phrases to be met with in Persian literature*. London: Routledge & Kegan Paul. [reprinted New Delhi: Manohar, 2006]
- Šukurov — Kapranov — Hošim — Mašsumī (1969) М. Ш. Шукуров — В. А. Капранов — Р. Хошим — Н. А. Маъсумӣ: *Фарҳанги забони тоҷикӣ (аз асри X то ибтидои асри XX)*. Ҷилди IА — О. Ҷилди II П — Ҷ. فرهنگ زبان تاجیکی. *Словарь таджикского языка (X — начало XX века)*. Москва: Советская энциклопедия.
- Vinogradova (1999) С. П. Виноградова: Ягнобский язык, in: *Языки мира. Иранские языки III. — Восточноиранские языки*. Москва: ИНДРИК, 290–310.
- Yorzoda (2007) Т. Ёрзода: *Фолклори Ягноб (Маҷмӯаи илмӣ-оммавӣ)*. *Ҷауднои folklor*. Душанбе: Империл-Групп.