Wakhi in New York:
Multilingualism and language contact in a Pamiri diaspora community

Ross Perlin, Habib Borjian, Daniel Kaufman and Husniya Khujamyorova

Abstract

The Pamiri languages, which make up a sizeable group within the Eastern Iranian language family, are spoken today in an areal continuum in the high-mountain region where the national borders of Tajikistan, Afghanistan, Pakistan, and China meet. Within just the last few decades, increasing outmigration from the region and new patterns of global mobility have resulted in a small but growing Pamiri community in New York City. The Endangered Language Alliance (ELA), a non-profit research organization founded in 2010, has been engaged in documenting several Pamiri heritage languages. This paper aims to provide a sociolinguistic description of this still relatively small and young community within the current universal trend of mass migration from remote rural areas to megacities. Two sample texts conclude the paper, analyzing contact effects and language attrition in Wakhi, one Pamiri heritage language in New York City.

1. Pamiri People and Languages

The Pamir Plateau, sometimes known as the "Roof of the World", is surrounded by steep riverine valleys which are home to diverse but related peoples who increasingly self-identify as Pamiri. The Pamiri domain also extends south into the adjoining Hindu Kush mountain range. In all there are some dozen valleys, each with several villages where a distinct Pamiri language or dialect is spoken, and local identities are often strong. Today this ecologically and culturally distinct region is politically divided, being at a juncture of four modern nations: Tajikistan, Afghanistan, Pakistan, and China. Most live in Tajikistan's Badakhshan province, with a sizeable community in the capital city of Dushanbe as well and an increasingly large diaspora in

Russia and beyond. Directly across the Panj River from Badakhshan lies northern Afghanistan, where sizable Pamiri communities also live.

Today's Pamiris most likely represent vestiges of the ancient Iranian-speaking groups that once occupied the entirety of Central Asia, although the exact linguistic relations between modern Pamiri languages and the ancient languages of the region, such as Sogdian, Bactrian and Khotanese, are as of yet unresolved. Relatively little is known about the history of the Pamir Plateau itself, given its remoteness to most of the historians and chroniclers whose records remain, but it has clearly been a well traversed, multilingual and multiethnic region for millennia (see Bahry 2016a for a recent summary). Chinese and other travelers along what is today called the Silk Road mentioned the region in passing, including Marco Polo, who refers to an area called "Vokhan" where "the people worship Mahommet, and they have a peculiar language" (Marco Polo I.32) – an early reference to the Wakhan corridor and the Pamiri language Wakhi which is spoken there. In one early 20th century survey of Badakhshan (Koshkaki 1979), Pamiris are already clearly distinguished from both their Turkic- and Tajik-speaking neighbors not only linguistically but also due to their Ismaili faith and their mountain-dwelling way of life – resulting in the exonym *ghalcha* (Shaw 1876), from the Iranian word *gar* "mountain". Many Pamiris also remain distinct in terms of their physical appearance, presumably owing to low level of contacts with the Turkic peoples who entered the region.

The term "Pamiri languages" refers to a remarkably diverse group of Eastern Iranian languages (Sims-Williams 2010) which are best described not as a subgroup distinguished by a set of shared innovations (Wendtland 2010) but as a kind of Sprachbund (Payne 1989), a group of languages that share areal features resulting from geographical proximity and language contact.

With by far the largest number of speakers, the Shughni group consists of Shughni (with its Ghund, Shāhdara, and Baju dialects), Rushani, Bartangi, Khufi, and Roshorvi, spoken in western Pamir in Tajikistan. Bartangi is sometimes classified as a dialect of Shughni, and one variety closer to the Shughni area is reported to be particularly close to it. Buddruss (1988) mentions two dialects that differ in minor details (Basidi and Siponji) and reports that speakers would consider Bartangi and Roshorvi to be related languages of a single language group, rather than dialects of a single language. There are also Shughni speakers on the Afghanistan side of the Panj

River. Closely related but now under heavier influence from Turkic varieties and now potentially from Chinese is Sarikoli, spoken on the eastern slopes of the Pamir plateau in what is today Tashkurgan Autonomous County in the Chinese province of Xinjiang.

Linguistically quite distinct and more conservative is Wakhi (Lorimer 1958, Grjunberg and Steblin-Kamiskij 1976, Paxalina 1975, Backstrom 1992), which is spoken in a small section of all four countries but is centered on (and named for) the Wakhan corridor which is divided (along the Panj) between Tajikistan and Afghanistan.

Just west of the Wakhan valley, endangered Ishkashimi (Paxalina 1959, Grierson 1920) is spoken on both sides of the Panj, in Tajikistan primarily in the single town of Rin and by an unknown but likely small number of speakers in a neighboring part of Afghanistan, where the closely related Sanglechi is also spoken. Further south in Afghanistan are Yidgha and Munji, spoken along the Afghanistan-Pakistan border (Morgenstierne 1938), for which there is relatively little recent information given ongoing instability in Afghanistan (but see Beck 2012, 2013 for recent survey of Pamiri languages in Afghanistan). Rounding out the group, at the western edge of the Pamir region in Tajikistan, is Yazgulami (Gauthiot 1916, Edel'man 1966), whose speakers are not Ismailis.

Most Pamiris are highly multilingual, speaking at least one Pamiri language natively as well as the national language and a language of wider communication. Those living in Tajikistan speak Tajik as the national language and learn Russian for higher education; in addition, Shughni (with the largest number of speakers) functions as local koiné among all Pamiris of the country and is sometimes referred to by Shughnis as "Pamiri". Bahry (2016b) exemplifies this multilingualism with a hypothetical example that happens to perfectly describe co-author Khujamyorova's experience:

> take the example of someone raised in a Wakhi-speaking family in Murgab, learning oral Kyrgyz from neighbors and classmates, Tajik and Russian in school, and on arriving at Khorough State University in the centre of the Shughni-speaking domain, adding Shughni informally and English formally to this long list of language repertoires for a grand total of six languages. (Bahry 2016b)

Mueller et al (2005) provide a detailed look at language use by domain and language attitudes among Shughni speakers, concluding that Shughni, while not seen as having great instrumental value in the larger society, is maintained with conviction in the areas of life where it dominates today:

> Our first impression was that the Shughni do not think highly of their mother tongue. Shughni was commonly characterised as being useful "to the airport," that is, it is useful in the region, but not outside the GBAP [Gorno-Badakhshan Autonomous Province]. With further conversation, however, a different picture slowly emerged. All people interviewed agreed that children should learn Shughni first. An overwhelming majority of those interviewed in all five locations think Shughni should be the language of the home; in fact, only one person interviewed felt Tajik should be used in the home to give children a better start in life. While Tajik is valued as the national language, the language of interethnic communication and for its historical role in literature and religion, its role is limited to apparently well-defined domains. (Mueller et al 2005: 15-16)

In Afghanistan, the Pamiri communities have knowledge of Dari Persian, the language of education and media. Based on the authors' own observations, Pamiris in Khyber Pakhtunkhwa in Pakistan are exposed to Urdu, English, and sometimes Pashto, while some Wakhi speakers in Hunza also learn Burushaski and/or Shina. The Hunza region, where there is a large population of Wakhis, had long been under the rule of Burushaski-speaking elites, and thus Burushaski, a language isolate, was the prestige language for several surrounding ethnic groups. While many older Wakhi speakers in certain areas are still fluent in Burushaski, younger Wakhis are typically only familiar with a handful of expressions. This generational divide is apparent among Pakistani Wakhis in New York as well. Nonetheless, Burushaski has made its imprint on the Hunza dialects of Wakhi leading to isoglosses that now distinguish it from northern dialeccts, e.g. Hunza Wakhi ṣapik 'bread', from Burushaski, where other dialects have native çɘtʃ. Grammatical differences between these dialect areas, especially with regard to alignment system, may also be due to contact with Burushaski, although this remains to be explored. In areas of Tajikistan with Turkic-speaking populations, like

Murghab, some Pamiris may also be conversant in Kyrgyz, in addition to Tajik, Russian and their Pamiri mother tongues. In China, Wakhi and Sarikoli speakers often learn Uighur as the regional language and Chinese as the national one. It should be noted that the linguistic literature on the Pamiri languages has a far stronger historical focus, as thus there has not been much written about the synchronic sociolinguistic situation in this multilingual area.

The sociopolitical status of Pamiris differs by the country they live in, but in all cases they are to different degrees a marginalized minority. In Tajikistan, the Mountainous Badakhshan Autonomous Province (Gorno-Badakhshan in Russian) was established in 1929, along with the Tajik Soviet Socialist Republic. Today Gorno-Badakhshan survives within the independent Republic of Tajikistan, and there is at least a degree of autonomy on the ground, though tensions with Dushanbe following the Civil War of the 1990s remain.

After the formation of the People's Republic of China in 1949, the Pamiris within China were recognized as "Tajiks", an official nationality (*minzu*), of which there are 56 in the country including the Han Chinese majority. In 1954, Tashkurgan Tajik Autonomous County was established in Xinjiang Province, which survives to this date, but research and information coming out of the county, on China's far northwestern border, have been sparse.

There is no special status for the Pamiri people in Afghanistan or Pakistan. In Afghanistan they are mentioned along with other ethnic groups of the country in the Pashto-language national anthem as Pāmiriān.

There are no reliable sources for calculating the Pamiri population. Tajikistan's Mountainous Badakhshan Autonomous Region had a population of 227,000 in 2019 ([Agency of Statistics](#)), constituting approximately three percent of the population of Tajikistan (Borjian 2007). Most are Shughni speakers. The total estimated number of Wakhi speakers, split between Tajikistan, Pakistan, Afghanistan, and China, has been variously estimated at 37,570, based on 1990-91 data (Bashir 2009) and 72,000 (Obrtelová 2019). Beck 2013 gives a figure of 25,000 for the Rushani-speaking population, with 18,000 in Tajikistan and 7,000 on the Afghan side, but the figure may well be larger. Sarikoli speakers may number around 25,000, constituting the majority of all Pamiri people living in Tashkurgan (Arlund 2006).Other varieties (e.g. Bartangi, Ishkashimi etc.) are all spoken by smaller populations.
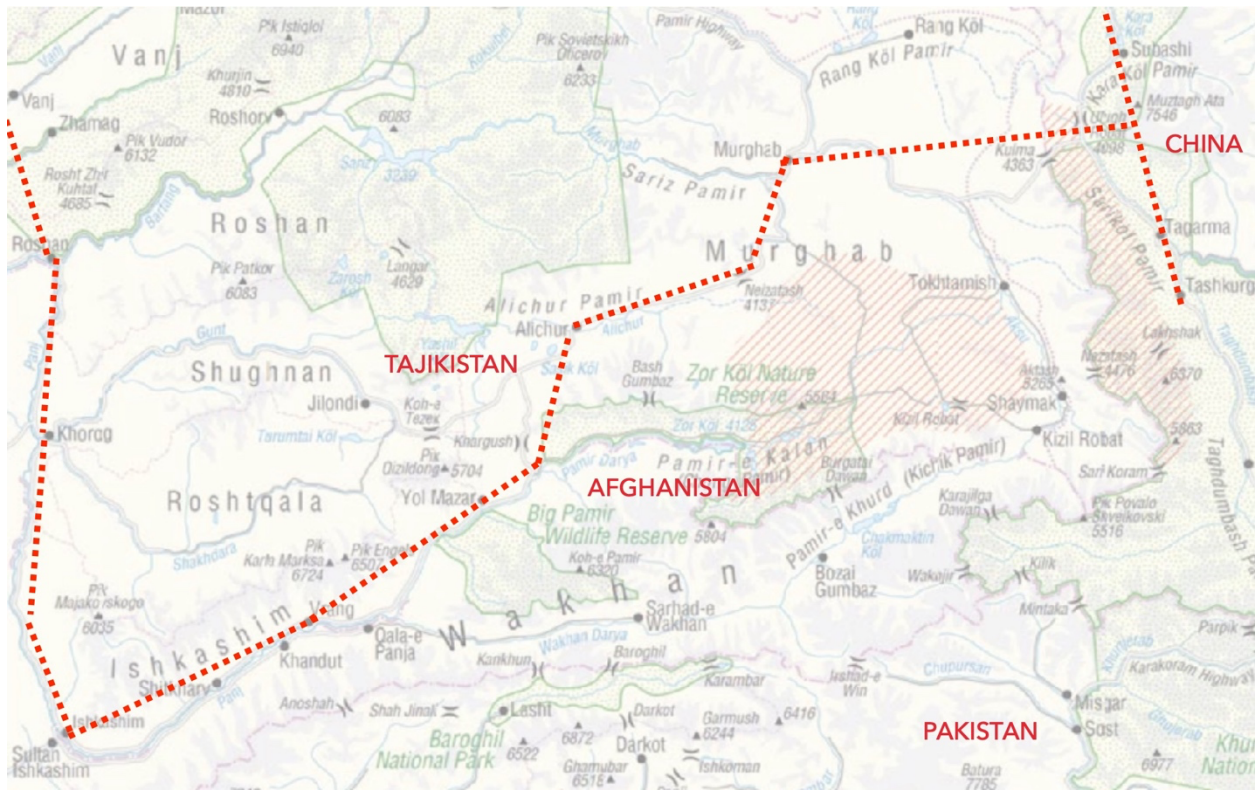
Figure 1. The Pamir region

2. The Pamiri Community of New York City

The past few decades have seen the rise of a small but growing Pamiri community in New York City, one of several new nodes in the United States and elsewhere in a well-networked Pamiri diaspora that now stretches beyond Central Asia and Russia (where there are larger and more established communities). According to community estimates, there may be as many as 300 Pamiris from Tajikistan in New York, of whom the majority are Shughni speakers from in and around Khorog, the capital of Tajikistan's Pamir region. Small numbers also speak Rushani, Bartangi, Wakhi, and Ishkashimi (Perlin et al. 2021) The first Pamiris arrived around 30 years ago, following the dissolution of the Soviet Union and in the wake of Tajikistan's Civil War in the 1990s.

A highly multilingual contingent of families and individuals in New York have also come in recent decades from Giglit-Baltistan, a mountainous and multilingual

region of northern Pakistan, including speakers not only of English, Urdu, and Balti, but also one Khowar-speaking family, a handful of Shina-speaking families, and some speakers of Burushaski. In terms of Pamiri languages, there are also at least half a dozen Wakhi families from Hunza.

There are not known to be any Pamiris in New York from Afghanistan, where poverty is acute, or from China, where mobility and access to information about the outside world are subject to restrictions.

A variety of "push" and "pull" factors account for migration chains from the Pamir region to New York, which almost always takes place via intermediate cities including Khorog, Dushanbe, and in some cases others. Most Pamiris from Tajikistan who are in New York seem to be young individuals, often in their 20s and 30s, coming for a combination of opportunities to study and work. Many of the men become drivers; others take on a variety of professional roles. Education including the graduate level is highly valued for both men and women, as is the norm in Ismaili communities. All Pamiri New Yorkers are understood to be Ismaili, though a few living outside the region may reportedly choose to follow non-Ismaili forms of Islam.

Pamiri multilingualism, including an increasing focus on English, plays a significant role in enabling migration, and Pamiri New Yorkers have maintained and developed ties with other communities to an extraordinary degree, illustrating the intergroup ties, multiple assimilations, and "panethnic" formations which lie behind the megadiversity of a city like early 21st century New York.

Of crucial importance is what might be termed "Ismaili globalization" and the resulting network of institutions, including both the *jamatkhana*s, or religious centers, where worship takes place, and the non-profit Aga Khan Foundation. Since the fall of the Soviet Union, previously isolated Pamiri Ismailis were rapidly welcomed and integrated into the global Ismaili movement led by the Aga Khan, a distinctly Western- and development-oriented Islamic leader. Most Ismailis in New York, as elsewhere, are South Asians, including a significant number from Gujarat and the Gujarati African diaspora who have been in the city for some time and founded numerous businesses. Pamiri New Yorkers, worshipping either at the Queens *jamatkhana* or its smaller Manhattan offshoot, have participated in and drawn from New York's existing Ismaili networks, forming notable bonds of mutual aid.

The area around the Queens *jamatkhana* is one focal point of settlement, but not only for religious reasons, for the Rego Park neighborhood where some Pamiris live and work is also near the center of Russian- and Bukhori-speaking Queens. In the 1970s, *refusenik* Ashkenazi Soviet Jews came to Rego Park to live among American-born Ashkenazi coreligionists, and in recent decades the neighborhood has become a hub for a diverse array of Jewish, Christian, and Muslim Russian speakers, with the largest element today being Bukharian Jews, for whom the area is a world center (see Borjian and Perlin 2015). Pamiris from Tajikistan and Jews from Bukhara, Samarkand, and Tashkent – Bukhori has also been called "Judeo-Tajik" – can readily communicate through Tajik or Russian, and some Pamiris have worked in the area's numerous Bukharian-owned businesses, including restaurants.

The other hub for Pamiri New Yorkers is in south Brooklyn, which is by far the largest Russian language and culture zone in North America. As in Rego Park, but to a much greater degree, a once Soviet Jewish cluster in parts of Brighton Beach, Bensonhurst, and adjoining neighborhoods has grown into a cultural and linguistic "microcosm" of the Soviet Union, with the Russian language and late Soviet culture as the glue. Here, Pamiris find themselves in a hyperdiverse zone that includes Russians, Tajiks, other Central Asians, South Asians (especially Pakistanis), Turks, as well as Latinos and many others. While Russian and Tajik may be important languages for work and daily life in this context, English certainly exerts a significant pull as well, and some Pamiris appear eager to embrace a more American identity.

The net result of these complex, fast-changing developments is as yet unclear, but language shift already appears to be underway among New York Pamiris. Younger Pamiris are living more and more of their lives in English, a language that connects them to American society and the global Ismaili world as well as South Asian communities. Tajik and Russian remain important, and to some extent Shughni as the lingua franca for Pamiri spaces (offline and online), but the smaller mother tongues such as Wakhi, Rushani, Bartangi, and Ishkashimi are clearly threatened. As with other smaller languages in New York that are doubly or triply embedded within larger communities, speakers report using their languages only rarely, often virtually, with relatives who may be scattered around the world. With some New York Pamiris now having children, the challenge of language transmission – even when language attitudes are quite positive – is coming to the fore. With English and Arabic the

languages of the *jamatkhana* and English, Russian, and Tajik most likely the languages of work, there are few defined spheres for the use of Pamiri languages in New York, though the community is well networked and Shughni, at least, may be used at gatherings or in group chats. Literacy, however, is not common in any Pamiri language, despite recent efforts.

The most popular annual event for the Ismaili community of New York is the Nowruz ceremony, observed around the 21st day of March. It draws hundreds of families, mostly South Asian Ismailis, whose observation of Nowruz is entirely confessional. The number of Pamiris attending the festive has increased to include scores of individuals in recent years. Moreover, members of the Pamiri community, within a team representing Tajikistan, participate in the New York Persian Parade every Nowruz.

3.   Fieldwork in New York and the Pamirs

The work reported on here is based at the Endangered Language Alliance (ELA), a nonprofit organization founded in 2010 with a mission to document endangered languages and support linguistic diversity. The network behind ELA consists of linguists, community leaders, language activists, and volunteers who document the speech, stories, and songs of immigrant, refugee, and diaspora communities, both for scholarly purposes and to bring multilingualism to a wider audience. Activities have ranged widely depending on community priorities, limited resources, and past work – they have included numerous public events, "traditional" language documentation, the production of pedagogical materials and the hosting of language classes, language mapping across urban space, and more.

ELA has had a particular focus on Pamiri languages with Wakhi-speaker Husniya Khujamyorova, originally from Murghab (Tajikistan), spearheading much of the fieldwork in New York, as well as in Tajikistan and China. Results to date include a large corpus of recordings, made both in New York City and the Pamirs, almost all of which have been transcribed and translated by Husniya Khujamyorova and glossed by Daniel Barry and Daniel Kaufman. The majority of this corpus has also been interlinearized and is presented publicly through Kratylos (www.kratylos.org), an online program that allows users to search and browse interlinearized glossed text and

lexicons.[1] The ELA corpus includes original fieldwork as well as glossed versions of Wakhi texts (and a small number of other Pamiri language texts) published by Russian linguists but now generally inaccessible to Pamiri communities (e.g. Grjunberg & Steblin-Kamenskij 1976; Paxalina 1975; Zarubin 1960; Sokolova 1959; Zarubin 1937; Paxalina 1959). New recordings made by ELA researchers in New York and the Pamirs are both "traditional texts" (folktales, personal narratives) as well as material of a more overtly sociolinguistic nature. ELA's transcribed and translated video documentation is disseminated through YouTube and other popular platforms in order to better reach the widest possible audience, in addition to linguists and other specialists.

Although Wakhi remains a relatively vital language, still learned by community members of all ages, use of the language is declining in areas where the Wakhi are a minority. Traditionally the language has not been written, with Persian/Tajik serving as a "high culture" written standard in the region. In recent years, activists and researchers working to preserve and write the language have developed a variety of Wakhi orthographies using Arabic, Cyrillic, and Roman letters. The diversity of writing systems reflects the four different countries where Wakhi speakers live, with the Cyrillic-based system used only in Tajikistan and the Urdu-based system only being used in Pakistan, for instance. ELA researchers have generally employed an IPA-based orthography that is not too dissimilar from one system promulgated by community members.

In the summer of 2016, Husniya Khujamyorova traveled throughout Tajikistan's Badakhshan region to record older native speakers telling traditional stories, folk tales, oral histories, and songs (*ruboi, bilbilik,* and *lalajik*), especially in Wakhi but also in Shughni and Bartangi. The resulting recordings, edited with original photography and made available through social media, have now been seen tens of thousands of times, especially by Pamiri people all over the world. Subjects range from traditional foods to Sufi songs to the story of the revered 11th century Persian poet and philosopher Nasir Khusraw, who is credited with spreading Ismailism in Badakhshan.

In the summer of 2018, with support from the National Geographic Society, Khujamyorova, filmmaker Nicole Galpern, and ELA Co-Director Ross Perlin spent much

---

[1] Kratylos was created under National Science Foundation DEL grant 1500753 to PIs Raphael Finkel and Daniel Kaufman.

of the summer in the Pamir region, filming recordings and interviews with approximately 70 speakers of 11 languages and dialects, totaling some 35 hours of video. Particularly significant are over 10 hours of material in the highly endangered and underdocumented Ishkashimi language as spoken in the village of Rin. A significant portion of these recordings are also being translated, transcribed, and disseminated online.

Driving over 1,000 miles of the Afghan-Tajik border, the extraordinary multilingualism and vitality of the region is clear. Murghab (Borjian 2014) was a Soviet melting pot and is today an emerging new Silk Road center. Across the Chinese border, the research team witnessed the tense security situation in Xinjiang, aimed primarily at Uighurs but also affecting the "Tajik" (i.e. Pamiri) population. A few recordings of Wakhi and Sarikoli were made on the Chinese side, with help from local collaborators who wished to remain anonymous out of fear. Given the dangers of working in Afghanistan, with Taliban forces near, the team arranged data-sharing with a local linguist who provided valuable samples of Afghan Wakhi.

In the following section, we focus on one Pamiri language in particular, Wakhi, as a heritage language in New York City, using two texts to analyze contact effects and language attrition.

4. Wakhi as a heritage language

The Pamiri diaspora is not particularly large, with a small concentration of immigrants in New York, mostly from Pakistan but with a few from Tajikistan, and a larger concentration in Moscow, likely all from Tajikistan. Wakhi is a language with a relatively small base of speakers compared to other regional languages that nonetheless spans many countries, having spread from the Wakhan corridor in Afghanistan north and south to what is today Tajikistan and Pakistan, respectively. More recently, within the last 200 years, further migrations took place from Pakistan into the Xinjiang region of China, where there are a small number of Wakhi speakers today. As can be imagined, the Wakhi spoken in each community has undergone contact effects due to multilingualism. This process has been progressively shaping Wakhi over many centuries, giving rise to a rather sharp dialect division between Tajik Wakhi and Pakistani Wakhi, which can both be broken up into small dialect areas. In Tajikistan, nearly all Wakhi speakers are multilingual, generally speaking Shughni, the local

lingua franca, as well as Tajik and Russian. Kyrgyz and other Turkic languages have also made a considerable impact on Tajik Wakhi. In Pakistan, many older Wakhi speakers also speak Burushaski and Shina in addition to Urdu. Younger speakers speak Urdu and, to some degree, English in addition to Wakhi. In China, Wakhi speakers also speak Chinese although linguistic influence from Chinese still seems limited at least from our brief fieldwork. The very different language profiles of these various communities have accelerated dialect differentiation to the point where mutual intelligibility between speakers from the Tajik and Pakistani communities is strained. Even before the advent of contact effects from these national languages, Wakhi already possessed a layered lexicon with vast amounts of Persian and Perso-Arabic borrowings in addition to older contact effects, which, for instance, likely gave rise to Wakhi's retroflex consonants, setting it apart from all other Pamiri languages.

Here, we are only examining the most recent layer of contact effects among two members of the New York community originally from Pakistan. The younger speaker immigrated at the young age of 6 and the older speaker immigrated as an adult and has lived in New York for roughly 17 years as of this writing. The younger speaker, as could be surmised, is a heritage speaker in the canonical sense (see Sedighi 2018 for various definitions of this term as it relates to Persian). She has an excellent command of the language but her speech is easily identified as that of someone who largely grew up in the United States. As we will discuss, the older speaker also employs English words, but her speech is otherwise not easily distinguishable from speakers who stayed in the Wakhi homeland. Because English is also an official language of Pakistan, it is not always clear, especially with younger speakers in New York, if English loan words have been introduced in New York or are merely reinforced here after having been introduced in Pakistan. The younger speaker, Rahila, pronounces English words using American English phonology, while the older speaker, Jamila Bibi, employs the phonology of Pakistani English.

In the following, we first examine aspects of grammar and lexicon followed by narrative content across the two speakers introduced above. The entire first text is provided in an interlinearized format in §5.1 and a portion of the longer second text is presented in §5.2 with a transcription and translation. The original media can be viewed in their entirety by following the links provided in §5.

4.1. Characteristics of heritage Wakhi

Discourse connectives and fillers are one of the most borrowable word classes (Matras 1998 *et seq*) and predictably we find different types of borrowed discourse connectives in the speech of the two speakers. Jamila makes very frequent use of *atʃa*, an Urdu connective, meaning 'good'. Rahila, on the other hand, possibly identifies this as Urdu and actively avoids it. Rahila's hesitation markers are English-derived, as is her frequent use of *like*, as in (1) (excerpted from [9] in §5.1).

(1)   çik   wor   *like program*-v-i   tain-əm   sək   *internet*
       Wakhi language like program-OBL.PL-OBL see-1SG.NPST in   internet
       'I watch, Wakhi language, like, programs on the internet.'

In contrast to Jamila, Rahila, who grew up surrounded by English since age six, also makes use of the English connective *so*, as in (2) (excerpted from [25] in §5.1), as well as full phrasal switches such as *you know*, as in (3) (excerpted from [11] in §5.1).

(2)   *so* tu=ʂ   ɣaftʃ qismi təj   ki...
       so 2SG.NOM=PROG   many kind EXT.PST COMP
       'So there are a lot ways you...'

(3)   tu=ki   koʃiʃ   tsar-∅   tsma ça-k-ər
       2SG=COND   struggle do-2SG.NPST search do-INF-DAT

       like   *resource*-iʃt=əv   kam woz kitob, *you know*
       like   resource-PL.NOM=3PL   little and book you know
       'If you try to search, like resources and books are few, you know.'

Note that in both (1) and (3) above, the use of English *program* and *resource* are completely integrated into Wakhi morphology, taking the oblique plural marker in (1) and the nominative plural marker in (3). The US English pronunciation of both of these words suggests that they represent a heritage feature rather than more widespread English borrowings.

Certain Urdu adverbs can be seen such as *xaskar* 'especially', and *kabi-kabi* 'sometimes' but it is not clear how widespread such usage is in Pakistani Wakhi. The fact that words like *xaskar* 'especially', *muxtalif* 'various', among many others, derive from Persian (often ultimately from Arabic) highlights the complexity of sorting out differences between Pakistani and Tajik dialect areas. Persian has permeated the region from varied sources and across a long period of time. Although both Pakistani and Tajik varieties of Wakhi are replete with Persian loans, some of these loans arrived before the breakup of the dialects, while other loans came through larger lingua francas, namely, Tajik and Urdu. To some extent the immediate provenance of a Persian loan can be diagnosed by sound change; Tajik loans reflect the sound change *ā > o while Persian loans that came via Urdu reflect *ā > a, as do early Persian loans directly into Wakhi. One example is *wadʒi* 'obligatory', which enters Pakistani Wakhi through Urdu, versus *woʒib* which enters elsewhere through Tajik. In addition to the basic geographical distribution, certain semantic developments unique to Tajik or Urdu also betray the origin of a particular loan. There is far more to be said about this, but for present purposes, it should be noted that there exists great variation in the meaning, pronunciation and distribution of Persian words in the two Wakhi dialect areas.

There is also an interesting lexical split between the northern dialect area, which underwent contact with Turkic languages, and the southern area which did not. This can be seen in lexical splits such as the Persian borrowing *muʃkil* 'difficult' (via Urdu) in Pakistan Wakhi versus *qin* 'difficult', which entered Tajik Wakhi through a Central Asian Turkic language. Similarly, Burushaski had no influence on the Wakhi spoken outside of Pakistan and created lexical splits such as aforementioned *ʂapik* 'bread' in Pakistani Wakhi versus native *çətʃ* 'bread' elsewhere in addition to the Turkic-origin *qemotʃ* in Tajik dialects.

As in Persian, the widespread use of light verbs allows for freer borrowing of verbs than is generally attested crosslinguistically. This can be seen in Rahila's use of *move* in the opening of her narrative, shown in (4). Similar examples abound in her narrative and are discussed in the following subsection.

(4)   wuz = əm   tə   rə-m   nju jork   *move*   vi-tu
        1SG.NOM = 1SG DAT ALL-PROX New York move become-PLPF

'I had moved here to New York.'

Beyond heavier use of English, there are clear "heritage effects" in the speech of Rahila that we do not find in speakers who migrated at an older age. For instance, there are tense/aspect mismatches (cf. Moore & Sadegholvad 2013 for Persian). To take one example, in (5) (excerpted from [3] in §5.1), the intended meaning appears to be 'Growing up, I have been speaking Wakhi with my parents.' However the past tense on the verb *çat* 'spoke/said' carries the (in this case incorrect) meaning that the speaker no longer speaks Wakhi with her parents. Whether or not this is due to interference from English or language attrition is impossible to say at this point.

(5)　to　　lup＝əm tsə　　vi-t＝a,　　　　u＝ş　　　　çik　　wor
　　　 until　big＝1SG when become-PST＝PRT 1SG.NOM＝PROG Wakhi language

　　　 də　　çɨ＝nan　　　　　　tat-və-n　　　　　ça-t-a　　　　woz
　　　 LOC　REFL.GEN＝mother　father-OBL.PL-ABL　say-PST-PST　and
　　　 'Until I grew up, I was speaking with my parents in Wakhi.'

A clearer case of potential English interference is found in (6) (excerpted from [8] in §5.1).

(6)　j-a-v-ən　　　　　　　 qti,　　to　*communication* muʃkil　　wos-t
　　　 DEM-MED-OBL.PL-ABL　 together until communication difficult　 become-3SG.NPST
　　　 'Communication with them becomes difficult.'

Here, the postpositional phrase *javən qti* 'with them' is produced with the ablative case on the noun phrase complement to *qti* 'with/together'. This postposition generally governs the locative *d-* case so that we expect *d-a-v-ən* (LOC-MED-OBL.PL-ABL) 'with them'.

Wakhi distinctions that have no clear structural analog in English also appear to get lost. For instance, a number of verbs can either be inflected directly by agreement or take a light verb, with attendant differences in meaning. One such verb is *tain* 'see', which has a telic/punctual reading when inflected directly but an atelic reading when

occurring as the complement of a light verb, i.e. *tain tsaren* (see do) 'to watch'. In (7), repeated from (1), we see that the directly inflected *tainəm* is used with an atelic intention.

(7) çik   wor   *like program*-v-i   tain-əm   sək   *internet*
     Wakhi language like program-obl.pl-obl see-1SG.NPST in   internet
     'I watch, Wakhi language, like, programs on the internet.'

Similarly, in the beginning of the same utterance, shown in (8), Rahila says *kṣijəm* 'I hear' intending 'I listen', which is generally expressed with a light verb, albeit with a different root.

(8) lekin u=ṣ   koʃiʃ   tsar-əm   ki   çik   wor
     but   1SG.NOM=PROG struggle do-1SG.NPST COMP Wakhi language
     bajd-v-i   kṣɨj-əm
     song-OBL.PL-OBL   hear-1SG.NPST
     'However, I try to listen to listen to Wakhi songs.'

There may also be simplifications of the tense/aspect/mood system evident in heritage Wakhi. One possible instance is shown in (9) (excerpted from [16] in §5.1), where the past existential *təj* is used.

(9) tu=t=ki   digar mɨlk   təj=a
     2SG.NOM=2SG=COND   other country exist.PST=QM
     'Whether you are in your own country or a different one...'

In conditional contexts, Wakhi typically employs a suppletive subjunctive form of the existential (*tsəj*), seen throughout Jamila's narrative in §5.2 but generally absent in Rahila's text in §5.1. As the English subjunctive is well on its way to obsolescence it could be surmised that the Wakhi subjunctive would meet a similar fate among heritage speakers, and this is precisely what we seem to find.

As a final grammatical commentary on §5.1, we find occasional examples of mismatches in the use of derivational morphology. In line (13) in §5.1, we find *dzəqlaj-*

ak (small-NMLZ) with the intended meaning 'childhood' while this would be normally expressed as *dzəqlaj-iǰ*, using another widespread nominalizer *-iǰ*. The formation *dzəqlaj-ak* also exists but with the meaning 'small child'.

4.2   Talking about Wakhi in Diaspora

Rahila speaks about the struggle of maintaining her language in New York with only a small handful of people to speak with. As she notes, there are very few books and resources, although she struggles to obtain what she can. As common among speakers of other threatened languages, she feels not only a personal responsibility to keep her heritage alive but also a responsibility to the language community as a whole to prevent the language from falling into dormancy. Rahila highlights the difference between trying to maintain a world language in diaspora, for which resources are plentiful and easily found on the internet and in libraries, versus maintaining a minority language, which, in her view, can only be preserved properly through the implementation of a home language policy. As someone who sings traditional Wakhi music, she emphasizes the potential of arts in language maintenance. Coming from an area where Wakhi, Shina, Burushaski, Urdu and English are all actively spoken, she marshals a multilingual perspective against the false dichotomy of having to choose between the heritage language and a dominant language. The dominant language, she points out, will always be learned outside, but the home should be reserved for the heritage language. This is ultimately tied to the remembrance of history and the ancestors.

One of the most striking aspects of this narrative is that the metalanguage used for discussing language preservation – indeed, more generally, for discussing culture *in the abstract* – is borrowed, almost without exception. In Rahila's narrative, we find that the core act of language maintenance is expressed by the light verb expression *preserve çak*. The methods for preservation are similarly expressed by English borrowings, such as *research, information, interest, arts,* and *study*. The word *language* itself (alternating with the native terms *wor* and *zik*) also appears in this regard. The terms denoting abstract domains relating to language are also recruited from English, for example, *identity* (alternating with Persian *ʃinaxt*), *culture, history* and *story,* which is used alongside the Urdu equivalent *kahani*. The ideology of cultural preservation in the Wakhi context is relatively recent and most likely shaped by external influence. The

Wakhi-speaking area in Pakistan was, after all, relatively isolated from global assimilatory forces until the opening of the Karakoram Highway in 1979. It is also natural for a heritage speaker to employ English terms for abstract concepts that would have only been introduced late in schooling.

      As to Jamila, even more striking is her repeated use of English *pure*, in (10) and (11), as well as elsewhere, to emphasize the distinction between the unadulterated language of her parents and home village and the mixed language of today.

(10)   wuz = əm   tse   kun   dʒaj-n = əm   ki   təja   *pure* waxij = əm
        1SG.NOM = 1SG from where place-ABL = 1SG COMP EXT.PST pure Wakhi = 1SG
        'I am from a place such that I am pure Wakhi...'

(11)   spo   tat-nan-iʃ   xoli *pure* spo   wor   ça-t
        1PL.GEN parents-NOM.PL only pure 1PL.GEN language   speak-PST
        '...our parents spoke a pure form of our language.'

      Despite the relatively infrequent use of English borrowings in Jamila's narrative, both the words *pure* and *mix* are used consistently employed as metalinguistic terms when discussing the state of the language in relation to others, as seen in (12).

(12)   kam~pam = i tsəj   d-a-n = i   urdu *mix* vi-tk.
        bit-DIM = 3SG   EXT.SBJN LOC-MED-ABL = 3SG Urdu mix become-PERF
        'It has become a little bit mixed with Urdu.'

      When discussing the linguistic diversity of New York, Jamila uses English *language* and *different* in her first pass and later in the same utterance amends these with the Urdu Perso-Arabic *muxtalif* 'various' and the Wakhi *zik*, as seen in (13).

(13)   trəm = i tsəj   *language* ɣa *different* = i   muxtalif = i tʃiz = i   trəm, zik.
        here = 3SG EXT.SBJN language very different = 3SG various = 3SG thing = 3SG here language
        'Here there are many different, various languages.'

Despite the Hunza valley being a long-standing hub of multilingualism, where Iranic and Dardic languages mingle with the language isolate Burushaski, all of which are now layered with Urdu, the terms for talking *about* diversity and difference overwhelmingly hail from English in both narratives. This seems to reflect a reality in which various types of diversity are lived out in the local languages of Hunza but diversity, as an abstract notion, is put forth explicitly as a value (and/or spectacle) by western ideologies associated with the English language. Similarly, the notion of a "pure" versus a "mixed" language is encountered most saliently from an outsider's perspective. A child growing up in a Wakhi village, of course, is not aware that some lexical items are recent "intrusions" from Urdu, English or surrounding languages. The further one is removed from village life, the more it becomes clear that certain aspects of Wakhi speech are most likely not historically native to Wakhi. Thus, the language of purity, mixture and diversity enter the picture most clearly when the speaker is in an English or Urdu speaking environment. The terms employed thus represent the speaker shifting her stance to regard Wakhi from the broader comparative perspective of an outsider looking in.

Just as in Rahila's narrative, Jamila consistently uses the English word *culture* to refer to Wakhi traditions from an abstract outsider's perspective, as in (14).

(14) agar sak jow koʃiʃ tsar-ən, spo *culture* ɣa xuʃruj.
     if 1PL 3SG.OBL try do-1PL.NPST 1PL.GEN culture very beautiful
     'If we try to keep it, our culture is beautiful.'

Again, we see that while "culture" is produced and experienced in Wakhi without an emic distinction between culture and life itself, it is presented as an imagined abstraction in English. Jamila immediately goes on to mention typical folkloristic aspects of Wakhi *culture* such as clothing and food. While these are minor aspects of the Pamiri traditional complex, they are the features most salient to an outsider's gaze, and are thus most readily indexed by the English term *culture*.

As both of these video recordings had been posted publicly on YouTube, they have accrued thousands of views and dozens of comments both from within the community and from outsiders with an interest in the languages of the region. As is common, code mixing is immediately picked up on by anonymous critics, as seen in a

comment on Jamila's narrative reproduced in (15a). Others, commenting on the same recording, both complement the language, as in (15b) 'Wakhi language is beautiful!', or deride the (minor!) code-mixing, as in (15c), "Auntie, pure is English, not our language".

(15)a. Wow! What an irony. She talking about speaking "pure wakhi" by speaking in "unpure Wakhi" 😁
  b. xušroy xikwor!
  c. ye voch pure ee english spo wor ee nast

In the commentary on Rahila's recording, afficionados comment on the novel sound of Wakhi from a comparative perspective, as in (16a), while another hurries to attribute the "mixed" impression to the fluency of the speaker.

(16)a. Sounds like old english mix with Persian, Urdu, Pashto, Greek, more Turkish and when Harry Potter speak to the snake in Chamber of secret. Lol 😂
  b. KJ Vids this not how the language actually sounds. It's mixed because the speaker is not fluent.

From outside the community, we find commentary in English with an overriding interest in the "true" affiliation of the language. Community members (as indicated by their use of Wakhi), are generally supportive of any attempt to promote the language on the internet but are also concerned about dialect differences, borrowings, code-mixing and levels of fluency, as they perceive these traits in the speakers. For many Wakhi speakers, YouTube videos provide the first opportunity to hear how their language is spoken across national borders. Predictably, we find many cross-border accusations of language adulteration via contact with Urdu, Russian, Tajik, etc.

5. Texts

5.1   Life as a Wakhi American

The following text was recorded by Husniya Khujamyorova in Queens, New York on March 9, 2018 through a grant from the Brooklyn Arts Council. The

transcription and translation was also done by Khujamyorova while the morphological analysis was carried out Daniel Kaufman. The entire video can be viewed at https://youtu.be/Al7Ga2NfUMs.

(1) zɨ   nung=i   Rahila
     1SG.GEN name=3SG Rahila
     My name is Rahila.

(2) woz   wuz=əm   ʃad sol nan   tsə tu   wuz=əm   tə
     and   1SG.NOM=1SG six year mother when EXIST.PST   1SG.NOM=1SG DAT

     rə-m   nju jork   *move* vi-tu   gilgit-ə-n
     ALL-PROX   New York move become-PLPF Gilgit-OBL-ABL

     zɨ   tat   nan-iʃt=əv   hunza-n,   hunza pakistan
     1SG.GEN father mother-PL.NOM=3PL Hunza-ABL Hunza Pakistan
     'When I was six years old, [we] moved to New York from Gilgit [Pakistan]. My parents are from Hunza, Pakistan.'

(3) to   lup=əm tsə   vi-t=a,   u=ʂ   çik   wor
     until big=1SG when become-PST=PRT 1SG.NOM=PROG Wakhi language

     də   çɨ=nan   tat-və-n   ça-t-a   woz
     LOC   REFL.GEN=mother   father-OBL.PL-ABL   say-PST-PST   and
     'Until I grew up, I was speaking with my parents in Wakhi.'

(4) də çɨ=çɨj   vrɨt-və-n=bə   woz çə   jandi t-rə-m=ən
     LOC REFL.GEN=sister brother-OBL.PL-ABL=also   and then then DAT-ALL-PROX=1PL

     tsə   *muv* vi-t-a   jaani=m   də çij   vrɨt-və-n=em
     when move become-PST-PST meaning=1SG LOC sister brother-OBL.PL-ABL=1SG

     ingliʃ   çən-ak   vi-t=a
     English say-INF   become-PST=PRT
     'With my siblings, when we moved here, I spoke to my sisters and brothers in English.'

(5)   çə də çɨ=nan   tat-n=əm=ş   çik   wor   çan-əm
        and LOC REFL.GEN=mother father-ABL=1SG=PROG Wakhi language say-1SG.NPST
        'But with my parents I speak in Wakhi.'

(6)   to   təmikir muʃkil=ş   wos-t   *kabi~kabi* tʃiz-r   ki
        until ??   difficult=PROG become-3SG.NPST sometimes what-DAT COMP

        zɨ   çik   wor=i   γaftʃ baf=bə   nast   o
        1SG.GEN Wakhi language=3SG very good=also NEG.COP PRT
        'It has been challenging sometimes – my Wakhi is not really good.'

(7)   ji sar tʃiz=ş   wuz çan-ak   nə=bas   wəzj-əm .
        one head thing=PROG   1SG.NOM say-INF   NEG=capable come-1SG.NPST
        'Certain things I cannot say.'

(8)   j-a-v-ən   qti,   to *communication* muʃkil   wos-t
        DEM-MED-OBL.PL-ABL together until communication difficult become-3SG.NPST
        'Communication with them becomes difficult.'

(9)   lekin u=ş   koʃiʃ tsar-əm   ki   çik   wor
        but 1SG.NOM=PROG struggle do-1SG.NPST COMP Wakhi language

        bajd-v-i   kşɨj-əm   woz,
        song-OBL.PL-OBL   hear-1SG.NPST   and

        çik   wor   *like program*-v-i   tain-əm   sək *internet* çə
        Wakhi language like program-OBL.PL-OBL see-1SG.NPST in   internet and

        çə jo jod   tsar-əm   jo   diʃ-əm   tʃiz-ər   ki
        and or memory do-1SG.NPST or   know-1SG.NPST what-DAT   COMP

        çɨ=zəbon   diʃ-n-i   γaftʃ zururi ma-r   sðɨj-d.
        REFL.GEN=language know-GER-OBL very necessary 1SG.OBL-DAT seem-3SG.NPST

'However, I try to listen to listen to Wakhi songs, and watch programs in Wakhi on the internet, and I learn and understand. Because it seems to me very important to know your native language.'

(10) woz   j-a-w   mer-n-e   nə=letsre-n   dɨstan=i   baf
      and   DEM-MED-PRO die-GER-OBL   NEG=let-GER   reason=3SG   good

'And it is good to not let it go extinct.'

(11) to   kabi~kabi muʃkil=ʂ   wos-t   ki   çik   wor   jəçk
      until   sometimes difficult=PROG become-3SG.NPST COMP Wakhi language learn

      wots-n=ən   woz diʃ-n=ən   tʃiz-r   ki   tu=ki
      become-GER=1PL and know-GER=1PL thing-DAT COMP 2SG.NOM=COND

      koʃiʃ   tsar-∅   tsma   ça-k-ər
      struggle do-2SG.NPST   search   do-INF-DAT

      like   resource-iʃt=əv   kam woz kitob, *you know*
      like   resource-PL.NOM=3PL   little and book you know

'Sometimes it is hard for us to learn Wakhi and to understand it, because when you try to learn the language, if you try to search, like resources and books are few, you know.'

(12) digar *language*, j-a   digar tʃiz jəçk wots-n-e   dzoq tsar-∅,
      other language DEM-MED other thing learn become-GER-OBL   desire do-2SG.NPST

      rətʂ-∅   tə library jo d-a   dukon çə   kitob dɨrz-∅   çə
      go-2SG.NPST LOC library or LOC-MED store and   book take-2SG.NPST and

      j-o-w   barar jəçk wots   jo   sək   *internet*
      DEM-MED-PRO about learn become-2SG.NPST   or   in   internet

      tain tsar-∅   çə *research* tsar-∅   ji   dam-ə-r
      see do-2SG.NPST and research do-2SG.NPST one breath-OBL-DAT

      to   tu   got-∅
      until 2SG.NOM   find-2SG.NPST

'With other languages, when you are willing to learn, you go to the library or store, get the book and learn the language. Also you can search on the internet and right away you can find everything.'

(13)
to   wuz=əm   dzəqlaj-ak-ən tʃiz ki di-ʐ   dast wəz-d-əj,
until 1SG.NOM=1SG small-INF-ABL what COMP LOC-1SG.GEN hand come-PST-PST

j-o-w=iʂ   dʒoj-əm   çik   wor   barər   muʃkilu
DEM-MED-PRO=PROG read-1SG.NPST Wakhi language about   difficult

ɣaftʃ tʃiz=i   nast,   ɣaftʃ *information*=i   nast
very thing=3SG   NEG.COP   very   information=3SG   NEG.COP

çik   wor   barar
Wakhi language   about
'Since my childhood, whenever I found something, I used to read it for the purpose of learning Wakhi. There is not enough information about the Wakhi language and it has been challenging.'

(14)
to   miʃkil=iʂ   wos-t   j-o   barar diʃ-n=ən
until difficult=PROG become-3SG.NPST DEM-MED.PRO about   know-GER=1PL

woz j-o   barar   dʒoj-n=ən
and DEM-MED.PRO   about   read-GER=1PL
'It is difficult for us to learn it and read it.'

(15)
to   spo   zik   zinda   lətsər-n=i distan ɣaftʃ muxtalif tʃiz
until 1PL.GEN tongue alive   let-ABL=3SG reason very   various thing

tu=ʂ   ç-ak   bas   wəzəj-∅
2SG.NOM=PROG   make-INF   capable come-2SG.NPST
'In order to keep our language alive, there are many things that you can do.'

(16)
tu=t=ki   digar miɫk   təj=a   tu=t   də
2SG.NOM=2SG=COND other country exist.PST=QM   2SG.NOM=2SG LOC

çɨ=miɫk=ət=bə=ki   təj   çik   wor   d-a

REFL.GEN=country=and=also=COND exist.PST Wakhi language LOC-MED

çɨ=zəman-v-ən   qəsa tsar-∅   har   zedor,   tʃiz-r   ki
REFL.GEN=child-OBL.PL-ABL talk do-2SG.NPST every necessary what-DAT COMP

de kum mɨlk=ʂ   ki   hal-∅,   digar zəbon=pə   j-a-w
LOC where country=PROG COMP stay-2SG.NPST other language=FUT   DEM-MED-PRO

jəçk   wos-t
learn become-3SG.NPST
'Whether you are in your own country or a different one, you should always speak Wakhi to your children, so that in whatever country you stay in, the [child] will learn the language.'

(17) lekin çik   wor   jəçk wots-n=əp   trəbar   nə=ç-ak
      but   Wakhi language learn become-ABL=FUT outdoor   NEG=make-INF

      bas   wizi-t   to
      capable   come-3SG.NPST until
      'But [the child] won't be able to learn Wakhi outside [the home].'

(18) tə   xun   çik   wor   qəsa çak=i   <u>wadʒi</u>
      DAT   house Wakhi language   talk make=3SG   necessary
      'At home, it is necessary to speak in Wakhi.'

(19) woz   tu=ʂ=ki   dʒoj-∅,   koʃiʃ, *like* tin-ən
      and   2SG.NOM=PROG=COND   read-2SG.NPST struggle like 2SG.GEN-ABL

      ki   *interest* təj,   çik-v-i   barar dʒoj-∅,
      COMP interest exist.PST   Wakhi-OBL.PL-OBL   about read-2SG.NPST

      çik-v-i   barar *study* tsar-∅,
      Wakhi-OBL.PL-OBL   about study do-2SG.NPST

      woz   ilm   <u>hasel</u> tsar-∅   woz
      and   knowledge result do-2SG.NPST   and

de *education field* waxi barar *research* tsar-∅
LOC education field Wakhi about research do-2SG.NPST

'If you are willing to read, and have an interest in reading about Wakhis, you need to learn about Wakhis, and to seek knowledge, through education and by doing research, Wakhi learning.'

(20) tʃiz-r ki təj = əv *organization*-iʃt = əv təj woz *grants*
what-DAT COMP EXT.PST = 3PL organization-PL.NOM = 3PL EXT.PST and grants

ta-r mila wots-ən bas wizi-t ki tu
2SG.OBL-DAT meet become-1/3PL.NPST capable come-3SG.NPST COMP 2SG.NOM

jə-m zəbon barar dʒoj-∅ woz jəçk wots-∅
DEM-PROX language about read-2SG.NPST and learn become-2SG.NPST

'There are also organizations and grants that can be found in order to read about and learn the language.'

(21) woz *arts* jiw woz lup tʃiz, tu = ṣ çɨ = zəbon
and arts one and big thing 2SG.NOM = PROG REFL.GEN = language

zindal tsr-en bas wizej-∅
alive do-GER capable come-2SG.NPST

'Also, arts is a big thing, you (can) keep your language alive.'

(22) ʃairi-və niviʃ-∅ de çik wor, tini = ki
poet-OBL.PL write-2SG.NPST LOC Wakhi language 2SG.GEN = COND

ʃairi ʃoq təj, ki nej tini = ki *photography* ʃoq təj,
poet passion EXT.PST COMP NEG 2SG.GEN = COND photography passion EXT.PST

'Write poems in Wakhi if you have a poetry-desire. If not, if you have photography-desire...'

(23) j-o-w-n = iṣ zəbon tum *preserve* ç-ak
DEM-MED-PRO-ABL = PROG language much preserve make-INF

nə = bas wizi-t
NEG = capable come-3SG.NPST

'With that, it might not be possible to preserve the language much with that,'

(24) lekin *video*-v-ni＝ki   tin   ʃoq   təj,
     but   video-OBL-ABL＝COMP   2SG.GEN   passion EXT.PST

video＝ş   *banaj*   ç-ak   bas   wizj-∅＝a
video＝PROG produce   make-INF   capable   come-2SG.NPST＝PRT
'But if you have a passion for video, you can produce video.'

(25) *so* tu＝ş   ɣaftʃ qismi təj   ki   tu＝ş
     so 2SG.NOM＝PROG   many kind   EXT.PST COMP 2SG.NOM＝PROG

     çɨ＝zəbon   *preserve* ç-ak   bas   wəzja
     REFL.GEN＝language preserve do-INF   capable   can
     'So there are different ways you can preserve your language.'

(26) to   jan   dzəng
     until   yes   like.that
     'So yes, it's like that.'

(27) ma-r   sðujd ki   çik   wor   zinda lətsər-n＝i   ɣaftʃ zəruri
     1SG.OBL-DAT   seems COMP   Wakhi language alive let-GER＝3SG very necessary

     tʃiz-r   ki   kum çik   ki də kum   miłk＝iş   ki
     what-DAT COMP where Wakhi COMP LOC where country＝PROG   COMP

     çik   hal-d,   j-o＝ş   ham çik   wor   çan-d   ham
     Wakhi stay-3SG.NPST DEM-MED.PRO＝PROG both Wakhi language say-3SG.NPST both

     a-d-a   miłk   ki kum zəbon   ki   təj＝a   çan-d
     EMPH-LOC-MED country COMP where language COMP EXT.PST＝PRT   say-3SG.NPST
     'I believe it is essential to keep the Wakhi language alive because any Wakhi-
     when a Wakhi person lives in any country, s/he speaks both Wakhi and speaks
     the language that exists there in that country.'

(28) to   pots-n-i,   çik-v-ni   bu ʃinaxt *identity*, ham çik-e-n
     until   1PL.GEN-ABL-OBL Wakhi-OBL-ABL two identity identity both Wakhi-OBL-ABL

ham də   kum   mɨlk＝iş   ki   hal-ən   ha-j-o-w
both LOC   where country＝PROG   COMP stay-1/3PL.NPST LOC EMPH-DEM-MED-PRO

zejl＝ş   wots-∅
assimilate＝PROG   become-2SG.NPST
'We, Wakhis have two identities: we are Wakhi and we are of whatever country
we are staying in.'

(29) to＝ki   çik   nə ça-t,   çik   wor＝ət＝ki   nə ça-t
      until＝COND Wakhi NEG say-PST Wakhi language＝2SG＝COND NEG say-PST

woz ne diʃ-t,
and NEG know-PST

tu＝p   j-a   digar *culture* vəzer   wots-∅.
2SG.NOM＝FUT DEM-MED other culture assimilate become-2SG.NPST
'If you didn't speak Wakhi and don't know it, you will assimilate to the other
culture.'

(30) j-o＝i   ʃak   nast   j-o＝i   baf ki
      DEM-MED.PRO＝3SG bad   NEG.COP DEM-MED.PRO＝3SG good COMP
      tu＝t＝şə   kum-r   ki   hal-∅   ha-j-a-w＝bə
      2SG.NOM＝2SG＝PROG   where-DAT COMP stay-2SG.NPST EMPH-DEM-MED-PRO＝also

zejl   tşaw-∅   lekin jə-m-i＝bə   baf ki
assimilate   go-2SG.NPST but   DEM-PROX-OBL＝also good COMP

tu   rɨmɨş məj   ki   ti...   ti   *history*
2SG.NOM forget NEG.SBJN   COMP 2SG.GEN 2SG.GEN history

tʃiz woz   ti   xalg-v-ə   zəbon＝i   tʃiz
what and   2SG.GEN person-OBL.PL-OBL language＝3SG   what
'It's not something bad that whenever you live somewhere, you get into it, but it
is also good not to forget [your language], your history, and what is your
people's language,'

(31) woz   j-a-v   tʃiz~tʃiz   də   çɨ=zəndagi   win-ətk
      and   DEM-MED-OBL.PL   thing~thing   LOC   REFL.GEN=life   see-PRF
      'and the things they have seen in their life.'

(32) jə-m   zəbon=ʂ   ha-j-a   kɨ tʃiz   a-d-a
      DEM-PROX language=PROG EMPH-DEM-MED all thing EMPH-LOC-MED

      zəbon=i   ha-j-a   kɨ tʃiz   ti   history wos-t,
      language=3SG EMPH-DEM-MED all what 2SG.GEN history become-3SG.NPST

      ti   story   ti   kahani   ti   pup-iʃt=əv
      2SG.GEN story   2SG.GEN   story   2SG.GEN   grandfather-PL.NOM=3PL

      tʃiz çə-tk   a-j-o-w   də ti   zəbon
      what make-PRF EMPH-DEM-MED-PRO LOC 2SG.GEN language

      zəruri   ki   tu   çɨ=zbon   me-rɨmɨʂ   to   ki
      necessary   COMP 2SG.NOM REFL.GEN=language NEG.SBJN-forget until COMP
      tu   çɨ=history   me-rɨmɨʂ
      2SG.NOM   REFL.GEN=history   NEG.SBJN-forget
      'So everything is in the language. Your history, what your grandfathers have
      done, for that language is important. You should not forget your language, so
      that you will not forget your history.'

5.2. A Wakhi Grandmother, From Pakistan to Brooklyn, New York

The following text is excerpted from a recording of Jamila Bibi, who tells the
story of her family and her language in Wakhi. The recording was made in Brooklyn on
April 24, 2017 by Cameddu Biancarelli working with Husniya Khujamyorova as the
interviewer, and Husniya also provides the transcription and translation. The entire
video can be viewed at https://youtu.be/O5Xbfxm06LU.

(1)   wuz=əm tse kun dʒaj-n=əm ki təja pure waxij=əm, waxi=ʂ çan-əm.
       'I am from a place such that I am pure Wakhi, and I speak Wakhi.'

(2)   wuz=ɨʂ aləm tse godʒal, ɣulkən zɨ dijor.
       'My hometown is Gulkin village in Gojal.'

(3)   jor wuz＝əm tuj＝əm ɣulkən vitk
        'I also got married in Gulkin.'

(4)   zɨ tat dijori tse murxuni.
        'My parents were from Murkhuni [village].'

(5)   spo ziki waxi＝ş sak kɨ çan-ən.
        '[Among ourselves] we speak only our language, Wakhi.'

(6)   godʒal xalg-iʃ tsumən be ki təj tsəj
        'All the people living in Gojal do,

        dzaq kam-pam＝i tsəj dani urdu *mix* vitk. …
        but it is a little bit mixed with Urdu.

(7)   e sakən, tsə trə mis waxt tsəj spo tat-nan-iʃ xoli *pure* spo wor çat
        'With us, back then, our parents only spoke a pure form of our language.'

(8)   spo tat-nan-e tsəj, *atʃa* sak＝iş be bet
        'Our parents did, and we do as well.'

(9)   wuz＝əm tsə zjadater＝em tsə uz＝əm tə hel al-ətk, ʃpɨn-iɟ＝əm çə-tk,
        'Most of the time I stayed in the summer pasture and did herding.'

(10)   *atʃa* ðas pandz sol-r＝ev tsəj maʐ＝ev tuj çətk. …
        'I was fifteen when I got married.'

(11)   ðas pandz soli xatm vitu diʐ jod təj
        'I remember I was fifteen.'

(12)   kam-pami diʐ jod təj tuj tum, maslam maʐ＝əv ɣa dʒald tuj çə-tu
        'I remember a little bit about my wedding, I got married very early.'

(13) *atʃa* z̪ini pandz zman. wuz=m maz̪ ta dʒaj jav çə-tu ɣulkin
‘I have five children, I gave birth to them in Gulkin.’

(14) *atʃa* niv=em, z̪i̡ tuj-en bad uz=əm karatʃi wəs-tu, re ʃer, *city*
‘After my wedding, I went to Karachi, to the city.’

(15) ram urdu jəçk vi-tu. dzaq-paq=əm wuz *school* rəj̡d çə, urdu=m dra jəçk vit=a.
‘I learned Urdu there. I went to school there a bit and learned Urdu.’

(16) wa pas=əm woz rəj̡d di̡ç dijor çe, tam wist sol ald-ət,
‘After that I went back to my village and stayed there for twenty years,

pure waxi=m çat, tʃiz ki z̪i̡ zik təj.
and spoke pure Wakhi, whatever my language was.’

(17) niv=əm woz wez-g amrika,
‘Then I came here to America,

tremi mar ðas pandz–ðas ʃad sol=i̡ş mar tş-it.
and I have been living here for 15, 16 years.’

(18) çe trəmi tsəj *language* ɣa *different*=i muxtalif=i tʃizi trəm, zik.
‘Here there are many different, various languages.’

(19) trəm=i̡ş tsəj har qism zik=i̡ş çan-ən,
There are many types of language spoken here,

çe *engliʃi, spaniʃi, tʃajnizi.*
like English, Spanish and Chinese.’

(20) sak=i̡ş tsəj trəm spo wor zjadatar jod tsar-ən, spo çi̡ zik.
‘We mostly use our language here, our native language.’

(21) de çi̡ xalg-və-n=ş ki *mila* wots-ən spo wor zijada, spo çi̡ waxi=ş çan-ən.

'Whenever we meet our people, mostly we talk our language, Wakhi.'

(22) dɨç zman-v-ən *xaskar* dəm çɨ xun = əm tsəj, dəm çɨ xun = ɨʂ wuz = ɨʂ waxi çan-əm.
'With my children, especially in my house, I speak Wakhi with my children.'

(23) wu = ʂ çɨ zman-və-r bə çan-əm ki, spo zik me rɨmʂit.
'I also tell my children not to forget our language.'

(24) *atʃa* woz zɨ nɨpɨs-iʃ, zɨn bu nɨpɨs uʂ d-av-ən bə çɨ wor çan-əm. …
'Even my grandchildren – I have two grandchildren and I speak my language with them.'

(25) *atʃa* zɨ zman-iʃ be pura po wor çan-ən. waxi = ʂ çan-ən
'All my children speak our language. They speak Wakhi.'

(26) woz wuz çat bə waxi
'And I myself too (speak) Wakhi.'

(27) zɨ tsumər ki zɨ qom-iʃ drəm tej, godʒal nək ki tsumər tej, pandz xun-ən sak drəm,
'and whatever relatives I have here, as there are five families from Gojal here [in New York],

sak = ɨʂ ki tan de ləman-ən mila wots-ən waxi = ʂ çan-ən.
whenever we get together, we talk in Wakhi.'

(28) sak = ɨʂ *get together* de ləmanɨ = ʂ wots-ən, neja ji tʃiz = ɨʂ gon-ən,
'We get together, you know, one thing we do,

e, spo wor = əʂ masalan jar çan-ən
we speak our language,

xɨʃ-iɟ = ɨʂ *banaj* tsar-ən drəm, də po xun-əv.
we make happiness here, in our home.'

(29) za-v-ər pidaj-iʃ rewor＝iʂ ki wost, pidov ki, tsə vi-tk aja rewor＝iʂ *banaj* tsar-ən
      'On the day when a child is born, we celebrate the day of their birth.'

(30) gon＝iʂ çə, aja rwor＝eʂ tsəj pura spo wor
      'That day, it is entirely in our language.'

(31) spo wor bajd, spo wor nəmendʒ, *atʃa* spo wor *muzik,*
      'Our songs, dance, music,

      bajd＝əʂ çan-ən, də tʃiz spo wor
      and singing songs in our language.'

(32) sak-r＝iʂ masalan trəm＝ən tsəj ɣaftʃ tʃiz sakər wost, *taklif* sakər wost,
      'For us, for example, we're here, we have many difficulties,

      masalan engliʃ çnak.
      for example, speaking English.'

(33) sak＝iʂ koʃiʃ tsar-ən dzaq＝iʂ engliʃ çan-ən,
      'We are trying to speak in English a little

      woz nəj woz＝iʂ çɨ wor zijada-tər saki jod tsar-ən
      and mostly speak our language.'

(34) dɨ çɨ xun çɨ wor.
      'At our home, our language.'

(35) woz spo *culture* tʃiz＝i ki tsəj, sak＝iʂ də çɨ xun jo jod tsar-ən,
      'Also, for our culture, whatever it is, we recall it in our house,

      jow-e＝ʂ istemol tsar-ən.
      we practice it here.'

(36) zɨ nɨpɨs-iʃt tsəj *pure* masalan, u＝ʂ d-av-ən koʃiʃ tsar-əm,
      'My grandchildren, pure, for example, I try [to speak] with them,

wuz be bet, ja nan be bet

I, as well as the mother,

jaw pup be bet kɨ tanɨʂ spo wor çanən.

the grandfather, too, are all speaking our language.'

(37) wu = ʂ çɨ zman-ve-r kiçt-er çan-əm ki, masalan spo wor qsa daven tsarit.

'I tell all my children to speak with them [their children] in our language.'

(38) jətiʃ çanən, zɨ lup nɨpɨs: "tati, nani" azɨng "gizi, nezdi"

'They say, my big grandson says, "dad, mom", like that, "stand, sit"

azəng. spo wor = əʂ jaw çand.

like that. He says it in our language.'

(39) Ajan zɨ nɨpɨs jiw nung = i Alhan,

'My grandsons' names are Ayan and Alhan,

atʃa zɨ ðəjd zman-iʃt = əv, Amina-n.

they are my daughter Amina's children.'

(40) bas baqi, spo wor tsəj spo wor ɣʃruj zik.

'Our language, our language is a very beautiful language.'

(41) agar sak jow koʃiʃ tsarən, spo *culture* ɣa xuʃruj.

'If we try to keep it, our culture is beautiful.'

(42) spo *culture* e tsəj skid = i, atʃa spots-n = i tsəj ptek = iʂ pɨrw-en te spo dijor,

'Our culture, we have skullcaps and scarves we used in our village.'

(43) skid = iʂ di-n, ptək = iʂ pɨrw-ən.

'They put on the cap and then the scarf on top of it.'

(44) spo xurok tsəj *xaskar*. spo xurok ɣaftʃ maslan dʒeda qismi tse ʃer-ən.

'And especially our food. Our food, for example, is very separate from city food.'

(45)   spotsni tsəj spo xurok tsəj: bati, molida＝i, gral＝i, semn＝i.
       'Some examples of our food are *bat, molida, gral, semn.*

(46)   e dzɨng-dzɨng tʃiz-v＝ɨʂ sak go-n.
       'These are some of the things we make.'

(47)   *atʃa* dzɨng spo tʃiz zik-iʃt, spo wor masalan tʃiz. …
       'Our languages are like that, our languages, for example…'

…

(48)   trəm amrika adzəng tʃiz ki, kɨ tan＝ɨʂ tsəj çɨ wor＝ɨʂ jod tsar-ən, istemol tsar-ən.
       'Here in America, everyone speaks their own languages.'

(49)   tʃajniz＝eʂ be çɨ zik, tʃiz tsar-t, istmol tsar-t.
       'The Chinese are also using their language.'

(50)   kum zik be tsəj, raʃjan da zik be ki, tsəj,
       'Any language, Russians in that language,

       har xalg çɨ, speniʃ çɨ zik,
       all people, the Spanish and their language.'

(51)   ingliʃ tsəj ɣa zuri nast, tumər ki çɨ zik-ər zuri təj.
       'English is not that much harder than our native language.'

(52)   sak çɨ zik çak-en dɨç zman-v-ən.
       'We speak our language with our children.'

(53)   po zik＝ki romoʂ-t＝əv ajane ɣa ʃak wots-n＝a.
       'It would be really bad if our language is forgotten.'

(54)   po zik＝i ɣa lazmi.
       'Our language is very useful.'

(55)   te dijor, te spo dijor=ki rətʂ-n=a d-av-ən qsa çak-ər tsəj spo ziki. …
        'In the village, if we go to the village, we must speak our language with them.'

6. Conclusion

We have presented an overview of the Pamiri community in New York and a summary of the overall Pamiri language situation. We have also made a first attempt at analyzing a Pamiri language as a heritage language in diaspora, contrasting a younger "generation 1.5" speaker with an older speaker who migrated later in life. We explored some structural aspects of language contact as well as possible motivations for code mixing with a brief foray into public attitudes towards mixing.

With regard to future directions, the researchers at ELA are continually working to edit, transcribe, translate, and analyze as many of the recordings as possible, integrating these into our corpora and making the results available both on popular platforms such as YouTube as well as linguist-oriented platforms such as Kratylos. Another planned focus is an illustrated series of children's books in five Pamiri languages, which Khujamyorova is editing, and further documentation work on Ishkashimi, with local partners, when the opportunity allows.

The complex language contact situation in the Pamir region has been mentioned by all linguists working in the region and has been studied using sociolinguistic surveys. A persistent gap in our knowledge is how Pamiri languages are changing due to contact and the various inter-variety accommodations that are being made both in rural areas and regional centers like Khorog, as well as in diaspora. We hope to have taken an early step here towards the linguistic analysis of contact in the latter context.

Raphael Finkel and Daniel Kaufman with Research Assistants Husniya Khujamyorova and Daniel Barry.

References:

Arlund, Pamela S. 2006. *An acoustic, historical, and developmental analysis of Tajik diphthongs.* Ph.D. dissertation. Arlington, TX. University of Texas at Arlington.

Backstrom, Peter C. 1992. Wakhi. In P. C. Backstrom & Carla J. Radloff (eds.), *Languages of Northern Areas*, 57-74. Islamabad: National Institute of Pakistan Studies, Quaid-i-Azam University and Summer Institute of Linguistics.

Bahry, Stephen A. 2016. Language Ecology: Understanding Central Asian Multilingualism. In *Language Change in Central Asia*, Elise S. Ahn and Juldyz Smgulova (eds.). Boston and Berlin: Walter de Gruyter.

Bahry, Stephen A. 2016. Societal multilingualism and personal plurilingualism in Pamir Tajikistan's complex language ecology. In *Language Change in Central Asia*, Elise S. Ahn and Juldyz Smgulova (eds.). Boston and Berlin: Walter de Gruyter.

Bashir, E. 2009. "Wakhi." In *The Iranian Languages*, ed. Gerhart Windfuhr, 825-862. Routledge.

Beck, Simone. 2012. The effect of accessibility on language vitality: The Ishkashimi and the Sanglechi language varieties in Afghanistan. *Linguistic Discovery* 10 (2): 157-233. http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/1/xmlpage/1/archive.

Beck, Simone. 2013. A sociolinguistic assessment of the Roshani speech variety in Afghanistan. *Language Documentation and Conservation* 7: 235–301. http://scholarspace.manoa.hawaii.edu/bitstream/handle/10125/4573/beck.pdf?sequenc e=1.

Borjian, Habib. 2007. Tajikistan. *Encyclopaedia Islamica* XIV, Tehran, 247-60.

Borjian, Habib. 2014. Morḡāb. *Encyclopaedia Iranica*, Online Edition. Retrieved on February 28, 2014.

Borjian, Habib & Ross Perlin. 2015. Bukhori in New York. In Matteo De Chiara and Evelin Grassi, eds., *Iranian Languages and Literatures of Central Asia: From the Eighteenth Century to the Present* (Studia Iranica. Cahier 57), 15-27. Paris: Association pour l'avancement des études iraniennes.

Buddruss, G. 1988. Bartangī, *Encyclopaedia Iranica* II.8, pp. 827-830. Republished in
    The Circle of Ancient Iranian Studies. CAIS. At https://www.cais-
    soas.com/CAIS/Languages/bartangi.htm. Retrieved June 15, 2020.

Edel'man, Džoy I. 1966. *Jazguljamskij Jazyk*. Moscow: Akademia Nauk SSSR.

Edel'man, Džoy I. 1980. History of the consonant systems of the North-Pamir
    languages. Indo-Iranian Journal 22: 287–310.

Edelman, Džoy I. & Dodykhudoeva, Laila R. 2009a. Shughni. *The Iranian Languages*. ed.
    by Gernot Windfuhr. New York: Routledge.

Edelman, Džoy I. & Dodykhudoeva, Laila R. 2009b. The Pamir Languages. *The Iranian
    Languages*. ed. by Gernot Windfuhr. New York: Routledge.

ELA: Endangered Language Alliance. Available at
    http://elalliance.org/languages/iranic/.

Gauthiot, Robert. 1916. Notes sur le Yazgoulami, dialecte iranien des confins du Pamir.
    Journal Asiatique 11. 239-270.

Grierson, George A. 1920. *Ishkashmi, Zebaki, and Yazghulami*. London: Royal Asiatic
    Society.

Grjunberg, Aleksandr L. & Ivan M. Steblin-Kamenskij. 1976. *Vaxanskij jazyk*. Moscow:
    Academy of Science.

Hjuler, A. 1912. *The languages spoken in the Western Pamir (Shughnan and Vakhan)*.
    Copenhagen. Nordisk Forlag.

Koshkaki, Mawlawi Borhân-al-Din. 1979. *Qataghan et Badakhshan, description du pays
    d'après l'inspection d'un ministre afghan en 1922*. Trans. Marguerte Reut. Paris:
    Editions du Centre National de la Recherche Scientifique.

Lorimer, D. L. R. 1958. *The Wakhi Language*. London: University of London.

Marco Polo. 1903. *The Book of Ser Marco Polo, the Venetian, Concerning the Kingdoms
    and Marvels of the East*, tr. Henry Yule, 3rd ed., London.

Matras, Yaron. 1998. Utterance modifers and universals of grammatical borrowing.
    Linguistics 36, 281-331.

Moore, J. & Sadegholvad, E. 2013. The Linguistic Markers of Persian Heritage
    Language Speakers: Evidence from the Classroom. *Heritage Language Journal*
    10(1), 83-107.

Morgenstierne, Georg. 1938. *Iranian Pamir languages*. (Indo-Iranian frontier languages,
    2.) Oslo: Aschehoug & Co.

Mueller, Katja, Abbess, Elizabeth, Paul Daniel, Thiessen, Calvin, Thiessen, Gabriela. 2005. Language in community-oriented and contact-oriented domains: the case of the Shughni of Tajikistan. In Clifton, John (ed.). 2005. *Studies in Languages of Tajikistan*. Dushanbe, Tajikistan and St. Petersburg, Russia: National State University of Tajikistan and North Eurasia Group, SIL International.

Obrtelová, J., 2019. *From Oral to Written: A Text-linguistic Study of Wakhi Narratives* (Doctoral dissertation, Acta Universitatis Upsaliensis).

Paxalina, Tatjana N. 1959. *Iškašimškij jazyk: očerk fonetiki i grammatiki, teksty i slovar'*. Moscow Academy of Science.

Paxalina, Tatjana N. 1975. *Vaxanskij jazyk*. (Jazyki Vostočnogo Gindukusha.) Moscow: Akademia Nauk SSSR.

Payne, John. 1989. Pāmir Languages. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 417-444. Wiesbaden: Reichert.

Perlin, Ross, Daniel Kaufman, Jason Lampel, Maya Daurio, Mark Turin, Sienna Craig, eds., 2021. Languages of New York City (digital version), map. New York: Endangered Language Alliance.

Sedighi, Anousha. 2018. Persian as a Heritage Language. In *The Oxford Handbook of Persian Linguistics*, ed. by Anousha Sedighi and Pouneh Shabani-Jadidi. Oxford: Oxford University Press.

Shaw, R. B. 1876. On the Ghalchah Languages (Wakhí and Sarikolí). *Journal of the Asiatic Society of Bengal* 45. 139-278.

Sims-Williams, Nicholas. 2010. Eastern Iranian languages. *Encyclopaedia Iranica*, Online Edition. Retrieved on November 28, 2019.

Sokolova, Valentina Stepanovna. 1959. *Rušanskie i Xufskie teksty i slovar'*. Moscow: Academy of Science.

Wendtland, Antije. 2009. The position of the Pamir languages within East Iranian, *Orientalia Suecana* 58.

Windfuhr, G., ed. 2009. *The Iranian Languages*. London and New York: Routledge.

Zarubin, Ivan I. 1937. *Bartangskie i rušanskie teksty i slovar'*. Moscow: Academy of Science.

Zarubin, Ivan I. 1960. *Šugnanskie teksty i slovar'*. Moscow: Academy of Science.