

CHAPTER II

OVERVIEW OF THE YAGHNOBI LANGUAGE

1. LANGUAGE FAMILY

Yaghnobi is classified as a Northeastern Iranian language.¹ The criteria for distinguishing the two main genetic divisions of Iranian languages, Eastern and Western, is preservation of the Proto-Iranian voiced stops **b*, **d*, **g*. Western Iranian languages preserve these, but Eastern Iranian has the corresponding fricatives *v*, *ð*, *ɣ*. For example, in Tajik, a Southwestern Iranian language, the word *barodar* ‘brother’ preserves the voiced initial stop, while the Yaghnobi word *viru:t* ‘brother’ has the corresponding initial fricative. The Northeastern subgroup is distinguished by a morphological innovation,² which is the regular plural suffix *-t* (in Yaghnobi), or *-tə* (in Ossete). For example, Yaghnobi has the singular *po:da* ‘foot’ and plural *po:do:t* ‘feet’ (Comrie, 1989).

¹ Comrie (1981) explains that Iranian languages have been traditionally divided into two genetic groups: Western and Eastern. These groups are subdivided into Northwestern, Southwestern, Northeastern, and Southeastern subgroups. The geographical designations are derived from the distribution of forms of Old and Middle Iranian rather than the locations of the modern Iranian languages. See Appendix C for a full listing of languages in the Iranian language family.

² Comrie is not contrasting the Northeastern Iranian regular plural suffix to another particular, regular Southeastern suffix. Innovation and irregularities in morphology are typical of Eastern Iranian languages; a suffix involving the consonant *t* is just the innovation that is characteristic of the two Northeastern Iranian languages. For example, Pashto, a Southeastern Iranian language, has the singular *p̄sa* ‘foot’ and plural *p̄sə:* ‘feet’. (This is in the absolutive case. Pashto also has a variety of other plural suffixes that depend on the case and class of the noun.) Shughni, another Southeastern Iranian language, has the singular *pað* ‘foot’ and the plural *paðe:u* ‘feet’.

In the Middle Persian period (the fourth century BCE to the ninth century CE), the Northeast Iranian language subgroup was composed of Khwarezmian, Alan (Scythian), and Sogdian. Sogdian is well known through a wealth of Manichean, Buddhist, and Christian (Nestorian) texts from the seventh through ninth centuries CE. Khwarezmian has no modern descendants. Ossete is a modern descendant of Alan, and Yaghnobi is believed to be a descendant of Sogdian.

Yaghnobi has been identified as a descendant of Sogdian, in part, by its vocabulary. Table 2.1 shows examples of Sogdian words that have reflexes in Yaghnobi (Mirzozoda, 1998). The Tajik (Southwest Iranian) translation of each word is given for comparison to show that these words are not from a common Middle Persian source.

Table 2.1. Reflexes of Sogdian words in modern Yaghnobi

Sogdian	Yaghnobi	Tajik	English
<i>jatak</i>	<i>jota</i>	<i>gu:ft</i>	‘meat’
<i>afkamp, fkamp</i>	<i>famp</i>	<i>fift</i>	‘ceiling’
<i>tʃi</i>	<i>tʃi</i>	<i>az</i>	‘from’
<i>akut, kut</i>	<i>kut</i>	<i>sag</i>	‘dog’
<i>zaj</i>	<i>zoj</i>	<i>kiftzor</i>	‘cultivated field’

2. TYPOLOGY

In this section the two main traditional categories of typology (morphology and word order) will be described. Many other categories could be addressed, but typological classification is not the focus of this thesis.

2.1. MORPHOLOGY

The morphology of the Yaghnobi language is agglutinating³ and, to a limited degree, polysynthetic.⁴ The words in examples 1 and 2 exhibit both agglutination and polysynthesis.

(1) *intf -af*
 wife -3S
 ‘his wife’

(2) *na- mun -om -ift*
 not- permit -1S -PRS
 ‘I won’t permit him’

Since both of these words have more than one morpheme, and these morphemes are invariant and have fixed boundaries, they exemplify agglutination. These same words also exhibit polymorphism, since each word incorporates multiple morphemes

³ The term agglutinating is being used in the sense defined by Comrie (1981): “In an agglutinating language, a word may consist of more than one morpheme, but the boundaries between morphemes in the word are always clear-cut; moreover, a given morpheme has at least a reasonably invariant shape.”

⁴ Comrie (1981) defines a polysynthetic language as a type of language in which “it is possible to combine a large number of morphemes, be they lexical or grammatical, into a single word, often corresponding to a whole sentence of English.”

and incorporates both grammatical and lexical morphemes. The polysynthesis is limited, since the only subjects and objects that can be incorporated are pronominal suffixes. Furthermore, these polysynthetic words rarely stand alone as a sentence; the pronominal suffixes usually correspond to nouns elsewhere in the sentence and function as grammatical agreement suffixes, as shown in example 3. In this sentence, *-om* 1S.PRS agrees with *man* ‘I’.

(3) *man amonatdor -im hitfuxs -i tisak na- mun -om -ift*
 I steward -1S no.one -CS enter not- put -1S -PRS
 (LOC)

‘I have been entrusted [with guarding the door]. I won’t let anyone in.’

2.2. WORD ORDER TYPOLOGY

The basic word order is subject, object,⁵ verb (SOV), although there is some flexibility in word order, which will be discussed in the section on grammatical relations. Greenberg (1966) observed certain word order tendencies for SOV languages, which are shown in the left column of Table 2.2. The right column indicates whether or not Yaghnobi has the predicted word order.

⁵ The terms *subject* and *object* are used to mean respectively: most agent like, and most patient like. They do not necessarily correspond to the categories of *subject* and *object* in English.

Table 2.2. Greenberg's word order predictions vs. word order in Yaghnobi

Greenberg's prediction	Compared to Yaghnobi
Adjectives precede nouns.	Yes.
Genitives precede nouns.	Yes.
Relative clauses precede nouns.	No. Relative clauses follow nouns.
Noun phrases precede adpositions.	No, but becoming yes. There are both prepositions (older) and postpositions (newer) in Yaghnobi.
Verbs precede auxiliaries.	Yes.
Standard of comparison precedes comparative adjective.	Yes.

Samples of sentences with each of type of word order are shown in example 4.

These examples are self-explanatory, with the exception of 4d. Yaghnobi does not have fully developed postpositions, but it does have relator nouns that appear to be well along the way to becoming postpositions and function as postpositions. This is discussed further in chapter III in the section on spatial relators.

(4) Examples of word order in Yaghnobi

a. Adjectives precede nouns

ADJ N

man xun -i fov kut a- ven -im

I dream -CS **black dog** PST- see -1SG

(LOC)

'I saw a **black dog** in a dream.'

b. Genitives precede nouns

N -GEN N

safar karim -i tup a- nos

Safar **Karim** -CS **ball** PST- take

(GEN)

'Safar took **Karim's** ball.'

3.1. PHONOLOGY

The phonemes of the Yaghnobi language are shown in tables 1 through 3 which present a summary⁶ of the phonological analysis done by Khromov (1972, 1987) and Vinogradova (2000).⁷ Vinogradova observes that the distinction between long and short vowels is disappearing; all vowels appear as short, except in stressed syllables, where the long vowels listed in Table 2.4 may still occur. Neither Khromov nor Vinogradova described the methodology⁸ used to determine the Yaghnobi phonemes, so it is unknown what minimal pairs or complementary distributions they may have found.

⁶ These tables were adapted from tables compiled by Lubomír Novák who translated the material from Russian to English (Internet: Phonology of Yaghnobi. *The Yaghnobi*. <http://yaghnobi.wordpress.com/2007/07/25/phonology-of-yaghnobi/>)

⁷ These tables have been modified to use the IPA rather than the phonetic transcription system used by Khromov (1972, 1987) and Vinogradova (2000).

⁸ There are a number of issues with the analysis presented in these tables. First, a number of the words presented in the examples are Tajik words (*gird* ‘round’, *xær* ‘donkey’, *fer* ‘lion’). These words may reflect Tajik rather than Yaghnobi phonology. Second, some of the choices for “basic realization” seem rather arbitrary. For example, why was /u/ chosen as the target (underlying form) rather than /ʊ/, or /o/?

Table 2.3. Short vowels

Phoneme	Allophones	Environment	Examples
/i/	[i]	Basic realization with a wide range of free variation: [i - ɪ - e]	
	[ĩ] (extra short)	In an open syllable preceding a stressed syllable	<i>xĩ'fift</i> 'milk', <i>tĩ'raʃ</i> 'three'
	[i]	Adjacent to fricatives and in closed syllables following palatal or velar stops	<i>fĩ'rok</i> 'tomorrow', <i>vi'rot</i> 'younger brother', <i>kif-</i> 'to plant', <i>gird</i> 'round'
	[e]	In word-final position and preceding a pharyngeal or uvular	<i>'morti</i> 'man', <i>'aβi</i> 'his/her', <i>ix</i> 'this(?)', <i>dih-</i> 'to hit'
/a/	[a]	Basic realization	
	[a:]	Result of a compensatory lengthening in case of loss of /h/ or /ʔ/	<i>ka:du:m</i> 'squash' < <i>kahdu:n</i> , <i>ʒa:m</i> 'to gather' < <i>ʒa:ʔm</i> < <i>ʒam</i>
	[æ]	Adjacent to a uvular	<i>ɣæɾ</i> 'mountain', <i>xæɾ</i> 'donkey'
/u/	[u]	Basic realization with free variation of [u - ʊ - o]	
	[ũ] (ultra short)	In open pre-stressed syllables	<i>su'tur</i> 'sheep', <i>ʃu'mox</i> 'you', <i>puxóy-</i> 'to cut'
	[u]	Adjacent to a fricative	
	[o]	In closed syllables containing a plosive	<i>urk</i> 'wolf', <i>uxf</i> 'six', <i>kut</i> 'dog'
	[y]	Preceding a uvular or pharyngeal. (Dialectal feature?)	

Table 2.4. Long vowels

/i:/	[i]	Basic realization	
	[i:] (lowered)	Between plosives	<i>tik</i> ‘also, again’
/e:/	[e:]	In stressed syllables word medially and finally	
	[e] (raised)	Adjacent to /ʃ, ʒ, m, n/	<i>fer</i> ‘lion’, <i>met</i> ‘day’
	[ɛ:]	In the Eastern dialect (from historical <i>*a^j</i> , usually merges with the historical <i>*e</i>)	<i>men</i> ‘village’, <i>βeʃ</i> ‘grass’
	[aḯ]	In the Western dialect (maintains historical <i>*a^j</i>)	<i>maj̄n</i> ‘village’, <i>βaj̄ʃ</i> ‘grass’
/o:/	[ɔ:]	Basic realization	
	[o:]	In stressed, closed syllables	
	[u:]	Before nasals	<i>nom/nu:m</i> ‘name’, <i>pi'rɔnt-/piru:nt-</i>
/u:/	[u:]	Basic realization	
	[y:]	Variant of historical <i>*u:</i> in stressed position; pronunciation can merge with /i/. Dialectal feature.	<i>xu:r/xy:r</i> (<i>xir</i>) < <i>*xu:r</i> ‘sun’, <i>kabu:d/kaby:d</i> (<i>kabi:d</i>) < <i>*kabu:d</i> ‘blue’

Table 2.5. Consonants

/p/	[p]	
	[p ^h]	In word-initial or word-final position
/b/	[b]	
/t/	[t]	
	[t ^h]	In word-initial or word-final position

/d/	[d]	
/tʃ/	[tʃ]	
/ɕ /	[ɕ]	
/k/	[k]	
	[k ^h]	In word-initial or word-final position
	[c]	Articulation of /k/adjacent to a front vowel
	[c ^h]	In word-initial or word-final position when adjacent to a front vowel
/g/	[g]	
	[ʝ]	Adjacent to a front vowel
/q/	[q]	
/m/	[m]	
/n/	[n]	
	[ŋ]	Preceding a velar
/f/	[f]	
/v/	[v]	
/s/	[s]	
/z/	[z]	
/ʃ/	[ʃ]	Yaghnobi pronunciation of /ʃ/ is more palatal than in Tajik
/ʒ/	[ʒ]	Yaghnobi pronunciation of /ʒ/ is more palatal than in Tajik
/ɣ /	[ɣ]	

/x/	[x]	
/xʷ/	[xʷ]	In Yaghnobi there can be no consonant clusters word initially, so /xʷ/ is considered to be single labialized vowel, not a cluster of /x/ + /w/.
/h/	[h]	
	[ħ]	Pronunciation of /h/ as [ħ] is a hyper-correct pronunciation of Arabic loans by Mullahs, ordinary people usually pronounce [h].
/ʕ/	[ʕ]	Realization of <i>ʕayn</i> as a pharyngeal fricative in words borrowed from Arabic often disappears or is pronounced as [ʔ] (a glottal stop) or [.] (a pause) in some words.
/β/	[β]	According to Khromov (1972) Yaghnobi /w/ is realized as a bilabial approximant; on the other hand, Vinogradova (2000) claims that it is realized more like a bilabial fricative.
	[β]	
	[w̥]	Following a vowel at the end of a syllable
/j/	[j]	
/r/	[r]	
/l/	[l]	

3.2. ORTHOGRAPHY

The transcriptions in this thesis were originally done using a modified version of the Cyrillic alphabet. This modified version is the same alphabet that is used for writing Tajik. In order to make the work in this thesis more widely accessible, all of the

transcriptions have been converted from Tajik Cyrillic to IPA. Table 2.6 shows the correlation of Tajik Cyrillic to IPA used to make this conversion. This system represents the phonological system used by Saifiddin Mirzozoda⁹, the primary Yaghnobi consultant on this project, who based his determination of phonemes on his own analysis and intuition as a Yaghnobi speaker as well as on the work of Khromov (1972) and Vinogradova (2002) which was summarized in the first part of this section.

There are a few differences between Mirzozoda's determination of phonemes and Khromov's: 1) Mirzozoda only recognizes two long vowels: /i:/ and /u:/, while Khromov¹⁰ recognizes three: /i:/, /a:/, and /u:/. 2) Mirzozoda does not distinguish /χ^w/ as a separate phoneme. There are also some issues with the Tajik Cyrillic transcription system used by Morzozoda. 1) He transcribes /ju/, /ja/, and /jo/ using single Cyrillic characters. This is a byproduct of using Cyrillic orthography¹¹ and does not indicate that these are single phonemes. 4) One shortcoming of the Tajik Cyrillic alphabet is that it does not distinguish /v/ and /w/. Mirzozoda does, however, recognize these as separate phonemes. 5) Mirzozoda uses a separate character, *й*, to represent glided [i], but this does not mean that he considers it to be a distinct phoneme. The decision to represent the Tajik Cyrillic character *й* in IPA as a non-syllabic vowel *j* (a glide), rather than *ɨ* (an extra short vowel) was the author's.

⁹ His name appears elsewhere as Mirzoyev, since he has published using both his official Tajik surname as well as his preferred Yaghnobi surname. Recently the Tajik government has ceased using russified patronymics for official documents, so Mirzozoda will become his official surname.

¹¹ In Mirzozoda's earlier works he transcribed [ja] as *йа* (*ja*), [ju] as *йу* (*ju*), and [jo] as *йо* (*jo*). It is not clear whether he changed his transcription system for phonological reasons or to make the writing system easier to learn for students familiar with Russian and Tajik.

Table 2.6. Orthography

Tajik Cyrillic	IPA	Tajik Cyrillic	IPA
А а	a	Н н	N
Б б	b	О о	O
В в	v	П п	P
В в	β	Р р	R
Г г	g	С с	S
Ғ ғ	к	Т т	T
Д д	d	У у	U
Е е	e	Ў ў	u:
Ё ё	jo	Ф ф	F
Ж ж	з	Х х	X
З з	z	Ҳ ҳ	H
И и	i	Ҷ ҷ	tʃ
Ӣ ӣ	i:	Ҷ ҷ	dʒ
Ӯ ӯ	i̇	Ш ш	ʃ
К к	k	Ю ю	Ju
Қ қ	q	Я я	Ja
Л л	l	Ъ	?
М м	m		