

SAMPLE PHONOLOGY I TERM PAPER  
 THE PHONOLOGY OF PASHTO  
 MICHAEL M. T. HENDERSON  
 SPRING 1998

**Introduction.** Pashto is a Western Iranian language, descended from Avestan. It has some sounds and syntactic features which are characteristic of Indic languages, and in the nineteenth century was thought to be a mixture of Semitic and Indic. Darmesteter (1888) showed that it was indeed Iranian, and that its ‘Indic’ features were borrowed. It is spoken today by approximately 6,000,000 people in Afghanistan and Pakistan, with four distinct varieties (Henderson 1983). In this paper I will describe the Kandahar variety, which has the greatest prestige and a literary tradition.

**Consonants.** The Pashto consonant system includes retroflex sounds, not found in any other Iranian language. Its consonants, all produced with pulmonic egressive airstream, are as follows:

*Obstruents.* /p b t d ts dz s z ṭ ḍ ʃ ʒ ʃ̣ ʒ̣ ʃ̣̌ ǯ̣ k g x ɣ/. The voiceless stops are lightly aspirated in all positions. The four affricates function as single sounds rather than as sequences. The dorsal fricatives /x ɣ/ are uvular [χ ʁ] before vowels. The retroflex fricatives /ʃ̣ ʒ̣/ contrast in all positions with the postalveolar fricatives /ʃ ʒ/.<sup>1</sup> There are no labial fricatives. There are several ‘elegant phonemes’ (Penzl 1955:34) which appear in Arabic loanwords: [ʔ] representing both the glottal stop and the voiced pharyngeal fricative of Arabic, [h], [q] and [f], the last occurring in loans from Persian and English as well as Arabic.

Examples:

/plar/	p <sup>h</sup> lɛr	‘father’
/biaa/	bjɛ:	‘again’
/tur/	t <sup>h</sup> ur	‘black’
/deg/	dɛ:g	‘pot’
/tsalór/	tsɛ <sup>h</sup> lɔ:r	‘four’
/pindzə/	p <sup>h</sup> in <sup>h</sup> dzə	‘five’
/saatəm/	sɛ: <sup>h</sup> t <sup>h</sup> əm	‘I keep’
/zə/	zə	‘I’
/tɛl/	tɛ:l	‘push’
/dɛr/	dɛ:r	‘very’
/ʃɪʃəm/	ʃɪ <sup>h</sup> ʃəm	‘I drink’
/dʒaan/	dʒɛ:n	‘dear’
/ʃəl/	ʃəl	‘twenty’
/zəbə/	<sup>h</sup> zəbə	‘tongue’
/káləna/	<sup>h</sup> kələnə	‘years’
/gúri/	<sup>h</sup> guri	‘you see’
/xar/	χar	‘donkey’
/ɣwáarəm/	<sup>h</sup> ɣwɛ:rəm	‘I want’

*Approximants.* /w l r ɾ/. /r/ is an alveolar tap [r], contrasting with the retroflex flap [ɾ]. /l/ is dental.<sup>2</sup>

<sup>1</sup> In the Quetta dialect, the retroflex and postalveolar fricatives have merged, being postalveolar everywhere. In the northeast Wardak variety, they are replaced with palatal fricatives [ç ʝ] (in some varieties [ç ɟ], and in Peshawar they have merged with the uvular fricatives.

<sup>2</sup> The semivowel [j] can be derived from vowels. [w] cannot, however. See below.

Examples:

/wɾɔɾ/	wɾɔɾ	‘brother’
/loɪ/	loɪ	‘big’
/ráuəɫ/	ˈrəwəɫ	‘bring’
/ɣwáaɾəm/	ˈɣwə:ɾəm/	‘I want’

The case of [w] is interesting. It cannot be derived from /u/ without trickery, since it contrasts in such near-minimal pairs as [wɾadz] ‘day’ and [ˈɾjadz] ‘cloud’. It forms clusters with other consonants, unlike [j], including stops, fricatives, and approximants, as in [χwɾɪ] ‘eats’. It appears to function as something between an approximant and an obstruent, replacing /f/ in some Persian and Arabic borrowings, and it is tempting to derive it from /v/.<sup>3</sup>

*Nasals.* /m n ŋ/. /n/ assimilates to the place of a following obstruent. /ŋ/ occurs initially only in the name of the letter which represents it, /ŋun/.

Examples:

/maa/	mə:	‘I (oblique)’
/na/	nə	‘not’
/rang/	rəŋg	‘paint, color’
/ŋun/	ŋun	(letter name)

**Vowels.** /i e a ə u o/. Pashto is unusual among Iranian languages in contrasting low vowels with the central vowel /ə/. In the Kandahar variety, there is a contrast between central [ɐ] and [ɛ:], so there is effectively a three-way contrast among central vowels.<sup>4</sup>

Examples:

/taasi/	ˈtʰɛ:si	‘you (pl.)’
/me/	me:	‘my’
/har/	hɛr	‘every’
/zə/	zə	‘I’
/staa/	stʰɛ:	‘your (pl.)’
/lur/	lur	‘daughter’
/ior/	jo:r	‘sister-in-law’

Vowels are slightly nasalized before nasals, and slightly longer in stressed syllables than in unstressed. Nasalization is not contrastive; length is, but only in the low vowels.

---

<sup>3</sup> This would result in the strange phenomenon of a language with a /v/ but no /f/, so I won’t pursue this solution further.

<sup>4</sup> In the Wardak variety, the contrast is between rounded [ɔ] and front [a]. In the Peshawar variety [ɐ] and [ə] have merged to [ɛ], contrasting with [ɛ:]. In this description, I’ll use the phoneme symbol /a/ for ease of exposition.

**Features.** The following distinctive features serve to characterize the consonant phonemes of Pashto, omitting the elegant phonemes:<sup>5</sup>

	p			t		s			t			ʃ	k	x
+voice	b	m	w	d	n	z	l	r	ɖ	ʈ	ɳ	ʒ	g	ɣ
Son	-	+	+	-	+	-	+	+	-	+	+	-	-	-
Cons	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Nasal	-	+	-	-	+	-	-	-	-	-	+	-	-	-
Lateral	-	-	-	-	-	-	+	-	-	-	-	-	-	-
Cont	-	-	+	-	-	+	+	+	-	+	-	+	-	+
PLACE	L	L	L	C	C	C	C	C	C	C	C	C	D	D
Anterior				+	+	+	+	+	-	-	-	-		

Assignment of features to the vowels is as follows:

	i	e	a	ə	o	u
High	+	-	-	-	-	+
Low	-	-	+	-	-	-
Back	-	-			+	+
Round	-	-	-	-	+	+

This analysis forces the analysis of [ɐ:] as geminate /aa/, for which there is no synchronic evidence (though it has some diachronic support). It also begs the question of the value of [back] in central vowels, which have traditionally been analyzed as nonround back vowels by generative phonologists. /e/ and /o/ are redundantly long, at least in stressed syllables; /i/ and /u/ are short, but can appear long as elegant phonemes in Persian/Arabic loans.

**Syllable Structure.** Unusual for an Iranian language in yet another respect, Pashto allows quiet complex CC onsets such /st/, /sp/, /sk/, along with /ts/, /ps/, /ʃp/, /zm/, and /ndʒ/ for a few examples. The only constraints on these onsets appear to be that obstruents must agree in voicing, and that one must be a stop and one a fricative or affricate.<sup>6</sup> The few CCC onsets all involve either an initial nasal, as in /ndr/, and/or an affricate, as in /ndʒ/, or a /w/ which can be analyzed as labialization, as in /ʃxw/ or /nɣw/. CC onsets with approximants as one of the members do not have the restrictions of the two-obstruent clusters: /pr/, /pl/, /lɣ/, /lm/, /dr/, /wr/, /gr/, and /zr/ all occur. Onsets are not required, nor are codas.

The rhymes of Pashto are more restricted than the onsets. Two vowels can fill a nucleus, but they can be identical only in the case of the low vowel /a/; otherwise one must be /i/ or /u/, with the syllabic peak on the other: /aa/, /ai/, /au/, /əu/ and /əi/ are the permitted VV nuclei.<sup>7</sup> In codas, CC clusters are permitted, but with less freedom than in onsets. Among those found are /lx/, /ks/, /ʃ/, /ɖʒ/ (which can follow /n/), /sp/, /st/, /sk/, /rd/, /nd/, /rk/, and /rg/.

The canonical shape of Pashto syllables is therefore (C)(C)(C)V(V)(C)(C), with the restrictions listed above.

**Prosodic Features.** Stress is contrastive in some instances: /kʃənəm/ ‘I will be sitting’ contrasts with /kʃénəm/ ‘I will sit’ (Penzl 1955:35). But some affixes such as /aan/ ‘animate plural’ and /una/ ‘inanimate plural’ always take the primary word stress, and others, such as the first person singular /-əm/ appear to take primary stress when attached to stems with short vowels, as in

<sup>5</sup> The affricates [ts dz ʃ ɖʒ] are analyzed as CC sequences which occupy one timing slot.

<sup>6</sup> In varieties where /z/ is represented by a stop, even this constraint is relaxed, allowing forms such as /gmundz/ ‘comb’, also heard as /mgundz/.

<sup>7</sup> Orthographically, diphthongs such as /aai/ and /aau/ appear, but they never contrast with /ai/ and /au/.

/kawə́m/ ‘I do’ vs. /wáíə́m/ ‘I call’. Pitch is used to differentiate questions from statements, which are otherwise grammatically identical: rising intonation indicates a question, falling a statement.

**Phonological Processes.** Pashto has a number of alternations between present and past verb stems, and a rather complex lexical phonology with alternations such as /trix/ ~ /tərxə/ ‘bitter’ (m. ~ f.) and /skor/ ~ /skaarə/ ‘charcoal’ (sg. ~ pl.). Its post-lexical phonology seems to consist mainly of vocalic alternations whereby long vowels get shortened when unstressed, as in /plaar/ ~ /plarúna/ ‘father ~ fathers’.

#### REFERENCES

- Darmesteter, James. 1888. *Chants populaires des afghans*. Paris: Klincksieck.
- Henderson, Michael. 1983. Four varieties of Pashto. *Journal of the American Oriental Society* 103.595-8.
- Penzl, Herbert. 1955. *A grammar of Pashto*. Washington, DC: American Council of Learned Societies.