ILLUSTRATIONS OF THE IPA

Dari (Afghan Persian)

Adam Baker

Academy of Sciences of Afghanistan &
University of North Dakota
adam.baker@iam-afghanistan.org

Dari, or Afghan Persian (Indo-European, Iranian, Southwestern; ISO 639-3: prs; Lewis, Simons & Fennig 2013), is one of the national languages of Afghanistan, and a language of wider communication in much of the country. It is one of the three major Persian varieties, the others being Farsi (or Iranian Persian) and Tajiki (or Tajik Persian). Afghans refer to Dari either as [dɐˈɾi] or [fʌɾˈsi].

An important phonological analysis of Dari is Henderson (1972). Kieffer's (1985) discussion provides a number of insights into dialectal variation around Afghanistan, and the relationship between Afghan Persian and other dialects. A wealth of data is to be found in Glassman (2000), though the focus of the text is pedagogical rather than linguistic. The phonology of Iranian Persian has previously been illustrated in Majidi & Ternes (1999). More detailed phonological studies of the Iranian and Tajik varieties of Persian are Windfuhr (1997) and Windfuhr & Perry (2009).

This Illustration is of the Dari spoken in Kabul, the capital of Afghanistan. Speakers of Dari exhibit a range of speech styles, depending on the formality of the situation. To capture some of this variation, recordings from two speakers are presented. The isolated words and the text so marked represent the formal speech style. The speaker for these recordings is from Kabul, though he has spent much of his adult life in Faizabad, Badakhshan province, and some time in Iran. In spite of these influences, he retains an identifiably Kabuli accent. The text is also presented in informal style. The speaker for the informal version is a lifelong resident of Kabul, educated, and about forty years of age. While not a trained linguist, he has considerable grammatical knowledge and metalinguistic awareness, arising from his position as a teacher of Dari to foreigners.

As the written form of the language is the basis for the formal style, formal pronunciations tend toward fidelity to the orthographic form. For example, the formal text contains [ˈme-kɐɾd-ɐnd] 'they were doing' where the informal text has [ˈme-kɐd-ɐn]. The [ɾ] and [d] present in the formal pronunciation reflect the written form, می‌کردند. Accordingly, formal Afghan and Iranian Persian are quite similar to one another, though not identical. The phonological analysis cannot merely recapitulate the orthography, however: several of the vowels are not represented orthographically, and orthographic forms often reflect the Arabic rather than the Persian pronunciations of words.

Consonants

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Glottal
Plosive	p b		t d			k g	q	
Nasal	m		n					
Fricative		f	s z	ʃ ʒ		x ɣ		h
Affricate				t͡ʃ d͡ʒ				
Tap or Flap			ɾ					
Approximant	w				j			
Lateral approximant			l					

Initial

pʌ	پا(ی)	foot
bʌd̪	باد	wind (N)
tʌd͡ʒ	تاج	crown
dʌˈxɛl	داخل	inside
kʌˈkʌ	کاکا	father's brother
gʌh	گاه	time
qʌˈzi	قاضی	judge (N)
mʌˈmʌ	ماما	mother's brother
nʌn	نان	bread
fʌɾˈsi	فارسی	Farsi
sʌˈlun	سالون	parlour
zʌt	ذات	essence
ʃʌm	شام	evening
ʒʌˈlɐ	ژاله	hail (N)
xʌˈdɛm	خادم	servant
ɣʌr	غار	hole
hʌˈfɛz	حافظ	protector
t͡ʃʌp	چاپ	print (N)
d͡ʒʌm	جام	bowl
rʌ	راه	road
lʌˈzɛm	لازم	necessary
wɐswɐˈsɐ	وسوسه	temptation
jʌfˈtɐn	یافتن	to find

Medial

kʌˈpi	کاپی	copy (N)
bʌˈbʌ	بابا	grandfather
hɐˈtʌ	حتی	even
ʌˈdɛl	عادل	righteous person
kʌˈkʌ	کاکا	father's brother
ˈɐgɐr	اگر	if
ʌˈqʌ	آقا	sir
mʌˈmʌ	ماما	mother's brother
bɐˈnʌ	بنا	build
ˈrɐfɐ	رفع	satisfaction
ɐˈsɛl	عسل	honey
qʌˈzi	قاضی	judge (N)
ʌˈʃʊq	عاشق	lover
nɛˈʒʌd̪	نژاد	ethnicity
dʌˈxɛl	داخل	inside
bɐˈɣɐl	بغل	side
mʊtɐˈhɐm	متهم	accused
lot͡ʃ	لوچ	naked
ɐˈd͡ʒɐl	اجل	fate
ɐˈrɐ	اره	saw (N)
sʌˈlun	سالون	parlor
hɐˈwʌ	هوا	air
ˈʌjʌ	آیا	interrogative particle

Final

t͡ʃʌp	چاپ	print (N)
t͡ʃob	چوب	wood
hʌˈlɐt	حالت	situation
bʌd̪	باد	wind (N)
nʌk	ناک	pear
sɐg	سگ	dog
bɛi̯ˈrɐq	بیرق	flag
ʃʌm	شام	evening meal
nʌn	نان	bread
nʌf	ناف	navel
dʌs	داس	sickle
hʌˈfɛz	حافظ	protector
ʌʃ	آش	noodle
loʒ	لوژ	balcony
jɐx	یخ	ice
dʌɣ	داغ	hot
koh	کوه	mountain
nʌt͡ʃʌr	ناچار	needy
tʌd͡ʒ	تاج	crown
ɣʌr	غار	cave, hole
sʌl	سال	year

There is not a great deal of categorical allophonic variation in Dari, though a number of consonant deletions and vowel coalescences occur at morpheme boundaries, as illustrated in the transcribed text below. The stops of Dari are unaspirated or lightly aspirated in all positions, except in careful speech. Voiced stops and affricates are optionally devoiced; in the recordings of isolated words, compare the devoiced allophones in [bʌd̪̊] 'wind (N)', [xub̥]

'good', [tʌd͡ʒ] 'crown', [neˈʒʌd̥] 'ethnicity' with the voiced voiced allophones in [bʊrd] 'he carried', [sud] 'interest', [t͡ʃob] 'wood', and [sɐg] 'dog'. The phoneme /ɾ/ may be trilled to [r] in emphatic speech, as occurs in the words [ˈɐgɐr] 'if', [ˈrɐfɐ] 'satisfaction', [ɣʌr] 'cave, hole', [rʌ] 'road' in the word list above; compare the flap [ɾ] which occurs in the transcribed text. The phoneme /h/ is frequently lost in everyday speech, though less often in formal contexts. Variation in the (formally produced) word list is illustrated above, with the word-final [h] preserved in [gʌh] 'time' and [koh] 'mountain', but dropped in [rʌ] 'road'.

In relation to Iranian Persian, it can be noted that Dari has distinct /q/ and /ɣ/, which are merged in Iranian Persian but similarly distinct in Tajik Persian (Kieffer 1985; Windfuhr & Perry 2009: 427). Dari always has the labiovelar approximant [w] instead of the Iranian or Tajik Persian [v]; [v] is never heard, even in formal speech.

Vowels
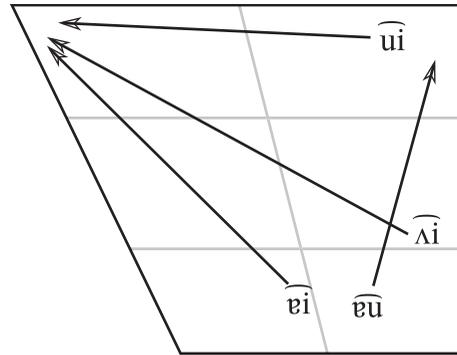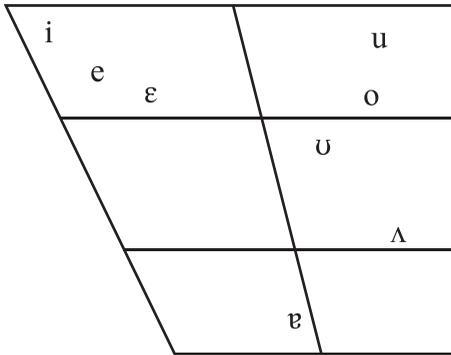
i	sim	سیم	wire	ʊ	bʊz	بز	goat
	lik	لیک	leak (N)		lʊtf	لطف	kindness
	mim	میم	meem (letter)		pʊl	پل	bridge
e	seb	سیب	apple		pʊt	پت	hidden
	bel	بیل	shovel	u	sud	سود	interest (financial)
	mex	میج	nail		nur	نور	light (N)
ɛ	sɛn	سن	age (N)		pul	پول	money
	gɛl	گل	clay	ʌ͡i	pʌ͡i	پای	foot/leg
	kɛʃt	کشت	cultivation	ɐ͡i	pɐ͡i	پی	tendon
ɐ	sɐd	صد	hundred	ɐ͡u	qɐ͡um	قوم	tribe
ʌ	bʌd̥	باد	wind (N)	u͡i	tu͡i	توی	wedding
o	bolt	بولت	bolt		ru͡i	روی	face
	boɣ	بوغ	trumpet				

Dari has a larger inventory of vowels than Iranian Persian (see the six-vowel system indicated by Majidi & Ternes 1999, and also more 'traditional' grammars like Thackston 1993 and Mace 2003). The additional contrasts – reported previously by Henderson (1972) and Glassman (2000) – are between the front round vowels [e] and [ɛ], and the back round vowels [o] and [ʊ]. These contrasts are similar to vowel contrasts present in Middle and Early Modern Persian (Windfuhr 1997) though a more precise statement of the relation awaits further etymological study.

The distinction between the back rounded vowels is illustrated by the words [xub̥] 'good', [xʊd] 'self', and [t͡ʃob] 'wood'. The phonemes /e/ and /ɛ/ are distinguished by the similar pair

Table 1 Vowel duration measurements of monophthongal vowels produced in isolation in the present paper.

Vowel	Duration (ms)		
	25% Quantile	Median	75% Quantile
e	293	302	312
i	272	282	296
ʌ	184	282	317
u	247	263	278
o	222	246	281
ɐ	158	222	269
ɛ	181	204	241
ʊ	169	177	190

[bel] 'shovel' and [gɛl] 'mud'.[1] Several examples of each of these vowels are given above. As discussed below, the difference between [e] and [ɛ], and [o] and [ʊ], is reinforced with a length contrast.

The vowels [ʌ͡i], [ɐ͡i], [ɐ͡u], and [u͡i] are analyzed here as diphthongs, but could alternately be analyzed as a monophthong–glide sequence. There is evidence in favor of both analyses. Although the maximal syllable template of Dari is CVCC, diphthongs do not occur in syllables with two coda consonants. This could be taken as an indication that one consonant position is being taken up by a glide. On the other hand, syllables with two consonants are not so frequent as to rule out the possibility of an accidental gap in the lexicon. Conversely, glides do not occur freely in coda position: a monophthongal analysis would then have to stipulate a list of permissible vowel–consonant sequences (i.e. [ʌj], [ɐj], [ɐw], [uj]), which in a sense merely recapitulates the diphthongal analysis. Therefore, though the phonological facts do not force one analysis over another, the diphthongal analysis is adopted here.

In the diagram above, the monophthongs are shown in their approximate locations in F2-F1 space. Two measurements are given for the diphthongs, corresponding to steady states towards the beginning and ends of the diphthongs.

Traditional analyses of Persian have distinguished between long vowels (as transcribed here, [i ʌ u]) and short vowels ([ɛ ɐ o]). Dari does not make a length distinction independent of vowel quality, though there are differences in the lengths of vowels. For purposes of this Illustration, duration measurements were made for the monophthongal vowels of the words produced in isolation. The results are shown in Table 1.

An ANOVA shows a general effect of vowel identity on vowel duration: $F(7,135) = 3.78$, $p < .001$. The difference in length between traditional long vowels of [i ʌ u] and the short vowels [ɛ ɐ o] is also significant: $F(1,141) = 19.20$ ($p < .001$). Finally, the difference is again significant if the Dari long vowel [e] and short vowel [ʊ] are included in the test: $F(1,131) = 13.35$, $p < .001$. In summary, then, there are length distinctions in Dari, which are redundant with vowel quality. These correspond to the traditional analysis of Persian short and long vowels.

Syllable structure follows a maximal syllable template of CVCC. The onset is optional; the obligatory epenthetic glottal stop reported by Windfuhr & Perry (2009) is not found in Dari (compare the words [ɐˈrɐ] 'saw (N)' and [ɐˈdʒɐl] 'fate'). Windfuhr & Perry (2009) note that Arabic loanwords are a source of many complex codas. On nouns, stress is typically syllable-final, including any suffixes and clitics. On verbs, stress falls on the first prefix, or on the last syllable of the root if there is no prefix.

---

[1] Henderson (1972) suggests the similar pair [sɛɾ] 'riddle' and [ʃeɾ] 'poetry', though the former of these words was unknown to my informants.

Transcription of recorded passage

The story 'The North Wind and the Sun' was adapted by the language consultant from Majidi & Ternes's (1999) Farsi version, to reflect Afghan word choice and usage.

Formal speech style

jɛk ɾoz bʌd hɛmˈɾʌjɛ ʌf ˈtʌb ‖ ˈkɐti jɛk dɛˈgɐɾ dɛˈwʌ ˈmekɐɾdɛnd ke kʊˈdʌmɛʃ qɐwiˈtɐɾ ɐs ‖ dɐɾ in wɐqt jɛk mʊsʌˈfɐɾ ɾɐˈsid | ke jɛk pɐˈtuje dɐˈbɛl dɐɾ ˈdɐwɾɛ xʊd peˈt͡ʃiˈdɐ bud ‖ ʌnˈhʌ gʊftɐnd ke | hɐɾ kʊˈdʌmɛ mʌ ke ɐˈwɐl ˈbɛtɐwʌnim mʊsʌˈfɐɾ ɾʌ meˈd͡ʒˈbuɾ kʊˈnim | pɐˈtujɛʃ ɾʌ ɐz ˈd͡ʒʌnɛʃ duɾ ˈkʊnɐd | mɐˈlum ˈmɛʃɛwɐd ke ˈzuɾɛʃ zijʌd ɐst ‖ bʌd tʌ tɛwʌˈnɛst wɐˈzid ‖ ˈleken ˈhɐɾt͡ʃe ke | beʃˈtɐɾ wɐˈzid | mʊsʌˈfɐɾ pɐˈtujɛʃ ɾʌ dɐɾ ˈdɐwɾɛ ˈxʊdɛʃ d͡ʒɐmɛ kɐɾd ‖ ˈbɛlʌxɐɾɐ | bʌd xɐsˈtɐ ʃud wɐ mʊnsɐˈɾɐf ʃud ‖ bʌd ɐz ʌn | ʌf ˈtʌb tʌˈbid wɐ hɐˈwʌ ˈinqɐdɐɾ gɐɾm ʃud ke | ˈfɐwɾɐn mʊsʌˈfɐɾ pɐˈtujɛʃ ɾʌ ɐz ˈd͡ʒʌnɛʃ duɾ kɐɾd ‖ bʌd ɐz ʌn | bʌd med͡ʒˈbuɾ ʃud ke ɛqˈɾʌɾ kʊˈnɐd ke ɐf ˈtʌb ˈzuɾɛʃ zijʌd ɐst fɐ

Informal speech style

jɛk ɾoz bʌd ɛmˈɾʌjɛ ɐf ˈtɛw ˈkɐti jɛkidɛˈgɐ dʌˈwʌ ˈmekɐdɐn ke kʊˈdʌmɛʃʌn qɐwiˈtɐɾ ɐs ‖ dɐ i wɐxt jɛk mʊsʌˈfɐɾ ɾɐˈsid | ke jɛk pɐˈtuje dɐˈbɛl dɐ ˈdɐwɾɛ xʊd pet͡ʃʌnˈdɐ bud ‖ uˈnʌ ˈgʊftɛn ke | hɐɾkʊˈdʌme mʌ ke ɐˈwɐl ˈbɛtʌnim mʊsʌˈfɐɾɐ med͡ʒˈbuɾ kʊˈnim ke | pɐˈtujɛʃɐ ɐz ˈd͡ʒʌnɛʃ duɾ ˈkʊnɐ | mʌˈlum ˈmɛʃe ke ˈzuɾɛʃ zijʌdˈtɐɾ ɐs ‖ bʌd tʌ ke tʌˈnɛst | wɐˈzid ‖ ˈlʌkɐn | ˈɐɾt͡ʃe ke beʃˈtɐɾ wɐˈzid | mʊsʌˈfɐɾ pɐˈtujɛʃɐ dɐ ˈdɐwɾɛ ˈxʊdɛʃ d͡ʒɐm kɐd ‖ ˈbɛlʌxɐɾɐ | bʌd xɐsˈtɐ ʃud o mʊnsɐˈɾɐf ʃud ‖ bʌd ɐz u | ɐf ˈtɛw tʌˈbid o hɐˈwʌ ˈɛqɐdɐɾ gɐɾm ʃud ke ˈfɐwɾɐn mʊsʌˈfɐɾ pɐˈtujɛʃɐ ɐz ˈd͡ʒʌnɛʃ duɾ kɐd ‖ bʌd ɐz u | bʌd med͡ʒˈbuɾ ʃud ke ɛqˈɾʌɾ kʊˈnɐ ke ɐf ˈtɛw ˈzuɾɛʃ zijʌdˈtɐɾ ɐs

Orthographic version

یک روز باد همراه آفتاب کتی یکدیگر دعوا می‌کردند که کدامشن قویتر است. در این وقت یک مسافر رسید، که یک پتوی دبل در دور خود پیچیده بود. آن‌ها گفتند که، هرکدام ما که اوّل بتوانیم مسافر را مجبور کنیم پتویش را از جانش دور کند، معلوم می‌شود که زورش زیادتر است. باد تا توانست وزید. لیکن هرچه که بیشتر وزید، مسافر پتویش را در دور خودش جمع کرد. بالاخده، باد خسته شد و منصرف شد. بعد از آن آفتاب تابید، و هوا اینقدر گرم شد که فوراً مسافر پتویش را از جانش دور کرد. بعد از آن باد مجبور شد که اقرار کند که آفتاب زورش زیادتر است.

Acknowledgements

References

Glassman, Eugene. 2000. *Conversational Dari: An introductory language course in Dari (Afghan Persian)*, 5th edn. Peshawar: The Interlit Foundation.

Henderson, Michael M. T. 1972. *Dari (Kabuli Persian) phonology*. Ph.D. thesis, The University of Wisconsin.

Kieffer, Charles M. 1985. Afghanistan v. languages. *Encyclopaedia Iranica* (online edition). http://www.iranica.com/articles/afghanistan-v-languages.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2013. *Ethnologue: Languages of the world*, 17th edn. Dallas, TX: SIL International. Online version: http://www.ethnologue.com.

Mace, John. 2003. *Persian grammar: For reference and revision*. New York: RoutledgeCurzon.

Majidi, Mohammad-Reza & Elmar Ternes. 1999. Persian. In IPA (ed.), *Handbook of the International Phonetic Association*, 124–125. Cambridge: Cambridge University Press.

Thackston, Wheeler M. 1993. *An introduction to Persian*, revised 3rd edn. Bethesda, MD: IBEX Publishers.

Windfuhr, Gernot. 1997. Persian phonology. In Alan S. Kaye & Peter T. Daniels (eds.), *Phonologies of Asia and Africa*, 2nd edn., vol. 2, 675–689. London: Routledge.

Windfuhr, Gernot & John R. Perry. 2009. Persian and Tajik. In Gernot Windfuhr (ed.), *The Iranian languages* (Routledge Language Family Series), 416–544. London: Routledge.