A Phonetic Description of Persian

Fangning Shao

Ling 111: Introduction to Phonetics

Fall 2016

Professor Dan Silverman

1. Introduction

Persian is one of the Western Iranian languages within the Indo-Iranian branch of the Indo-European language family. It used to be the language of the Persian Empire and was widely spoken in the ancient days. It ranges from the east part of India, north part of Russia, the shores of the Persian Gulf to Egypt and the west shore of Mediterranean. Before 18 centuries, even the kings of India used Persian language. Then the British banned Persian after occupying the county. In over a thousand of years' history, Persian has been changed into a modern form now. There are approximately 52 million Persian speakers worldwide according to Wikipedia. Most of the modern Persian speakers lives in Iran, Afghanistan, Tajikistan and parts of Uzbekistan.

Although the name of the language has been maintained as Persian or Parsi or its Arabic form Farsi (because in Arabic they do not have the letter P), the language has undergone great changes and can be categorized into the following groups: Old Persian, Middle Persian, Classical Persian and Modern Persian.

According to my consultant's introduce, nowadays Persian have at least 3 variations: standard (Iranian) Persian, dari Persian (mostly spoken in Afghanistan) and tajik Persian (mostly spoken in Tajikistan and Uzbekistan). Even in Iran, there are quite a few people speak the latter 2 variations instead of Iranian Persian(see map of languages in Iran: Figure 1 below).
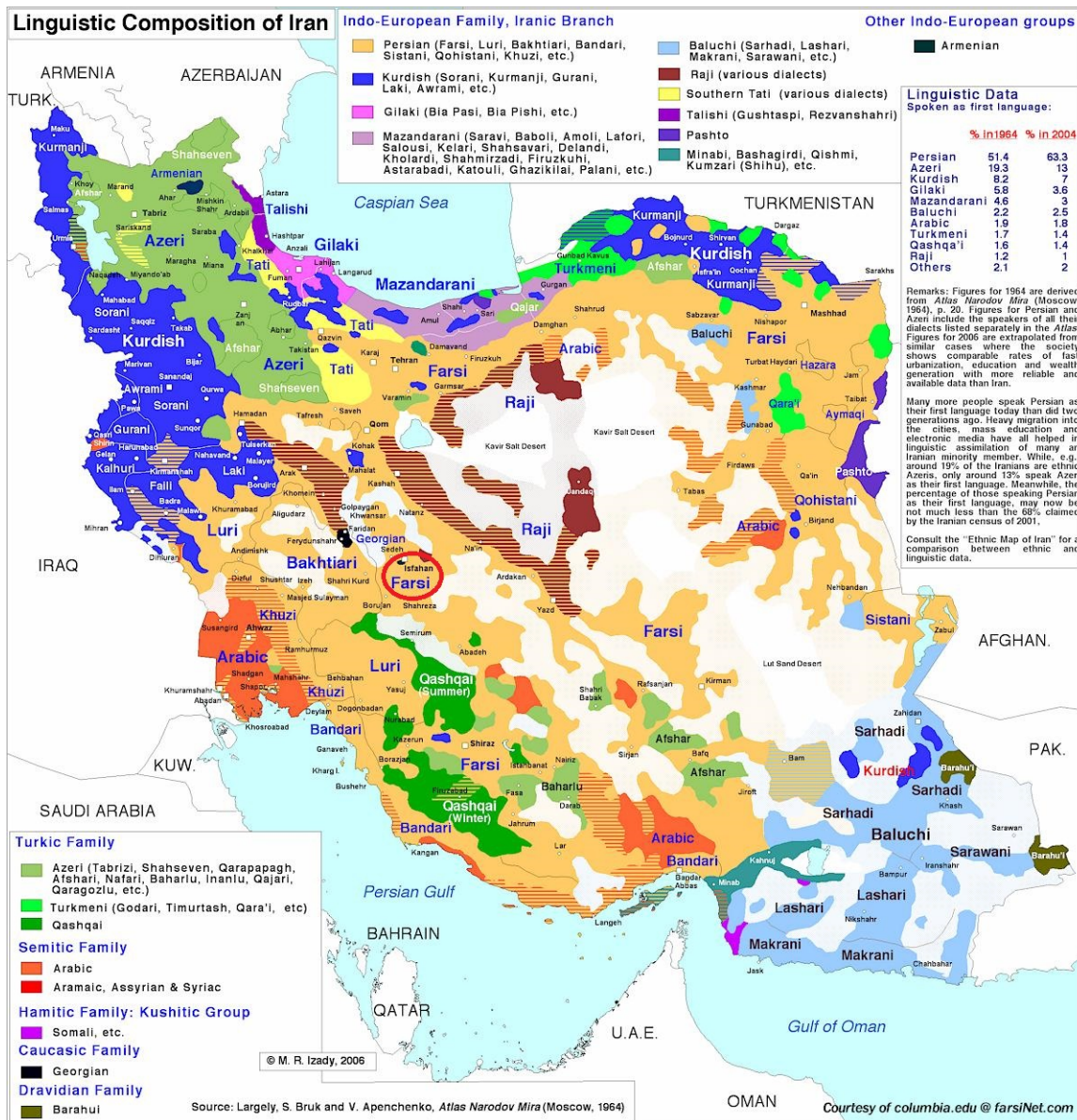
Linguistic Composition of Iran

Source: Largely, S. Bruk and V. Apenchenko, *Atlas Narodov Mira* (Moscow, 1964)

Courtesy of columbia.edu @ farsiNet.com

Indo-European Family, Iranic Branch

- Persian (Farsi, Luri, Bakhtiari, Bandari, Sistani, Qohistani, Khuzi, etc.)
- Kurdish (Sorani, Kurmanji, Gurani, Laki, Awrami, etc.)
- Gilaki (Bia Pasi, Bia Pishi, etc.)
- Mazandarani (Saravi, Baboli, Amoli, Lafori, Salousi, Kelari, Shahsavari, Delandi, Kholardi, Shahmirzadi, Firuzkuhi, Astarabadi, Katouli, Ghazikilai, Palani, etc.)
- Baluchi (Sarhadi, Lashari, Makrani, Sarawani, etc.)
- Raji (various dialects)
- Southern Tati (various dialects)
- Talishi (Gushtaspi, Rezvanshahri)
- Pashto
- Minabi, Bashagirdi, Qishmi, Kumzari (Shihu), etc.

Other Indo-European groups

- Armenian

Linguistic Data
Spoken as first language:

	% in 1964	% in 2004
Persian	51.4	63.3
Azeri	19.3	13
Kurdish	8.2	7
Gilaki	5.8	3.6
Mazandarani	4.6	3
Baluchi	2.2	2.5
Arabic	1.9	1.8
Turkmeni	1.7	1.4
Qashqa'i	1.6	1.4
Raji	1.2	1
Others	2.1	2

Remarks: Figures for 1964 are derived from *Atlas Narodov Mira* (Moscow, 1964), p. 20. Figures for Persian and Azeri include the speakers of all their dialects listed separately in the *Atlas*. Figures for 2006 are extrapolated from similar cases where the society shows comparable rates of fast urbanization, education and wealth generation with more reliable and available data than Iran.

Many more people speak Persian as their first language today than did two generations ago. Heavy migration into the cities, mass education and electronic media have all helped in linguistic assimilation of many an Iranian minority member. While, e.g., around 19% of the Iranians are ethnic Azeris, only around 13% speak Azeri as their first language. Meanwhile, the percentage of those speaking Persian as their first language, may now be not much less than the 68% claimed by the Iranian census of 2001.

Consult the "Ethnic Map of Iran" for a comparison between ethnic and linguistic data.

Turkic Family
- Azeri (Tabrizi, Shahseven, Qarapapagh, Afshari, Nafari, Baharlu, Inanlu, Qajari, Qaragozlu, etc.)
- Turkmeni (Godari, Timurtash, Qara'i, etc)
- Qashqai

Semitic Family
- Arabic
- Aramaic, Assyrian & Syriac

Hamitic Family: Kushitic Group
- Somali, etc.

Caucasic Family
- Georgian

Dravidian Family
- Barahui

Figure 1: map of languages in Iran

After the Muslim conquest of Persia, it took approximately 150 years before Persians adopted the Arabic script in place of the older alphabet. Modern Iranian Persian and dari Persian are written in Arabic script and contains a considerable amount of Arabic vocabulary. On the other hand, tajik Persian is written in modified Cyrillic script. Furthermore, modern farsi contains also part of vocabulary from English, French and German. These words are actually very recognizable. As an Iranian Persian speaker, my consultant states that she can understand most of these 2 other variations as long as they

speak solely and clearly. To her, these variations are more like the ancient Persian(old persian). They keep more old expressions and lexical items than Iranian Persian.

Figure 2: map of Iran

2. About my consultant

My consultant is an Iranian female in her late 20s. She came from Isfahan (also known as Esfahan), Iran. Isfahan is the third biggest city in Iran and locates in the middle-west part of the country (shown in the red circle in Figure 1 and 2). Most of people in Isfahan speak Iranian Persian. All the school in Iran use standard Persian as working language. The Persian in Isfahan is not a dialect but it has some accent. My consultant uses a standard form and does not have an accent. All of her family members were in Iran and speak standard Persian. She accepted all the education in Persian language, too. She also went to her college in Isfahan in which she took 3 years of literature classes. Although she is not a professional writer, my constant always loves writing. Some of her works

have been published in some Iranian medias. after college, she worked in Iran as a software engineer for 4 years. Three years before, she moved to the states and started to work in Texas. One year before, she got married with an Iranian American and moved to California. For all the years in us, she kept using Persian in her daily life. She talks with her husband in Persian, too. And also, she still keeps writing some poems in Persian. Therefore, she is a native speaker and actively uses this language every day.

3. The vowels in Persian

3.1. Syllables in Persian

As is commonly known, the core of a syllable is usually a vowel. If there is no vowel in a syllable, then it must have a syllabic liquid or nasal. Therefore, one syllable consists of one obligatory vowel and optionally surrounded consonants. Persian syllables have only three patterns: CV, CVC, and CVCC [Windfuhr 1979]. Therefore, a Persian syllable cannot start with vowels. It can even start with a glottal stop [?], but never a vowel. Furthermore, according to my observation, the initiation of a syllable can only be one consonant. On the other hand, the ending of a syllable can be one or two but no more than two consonants.

3.2. Vowels in Persian
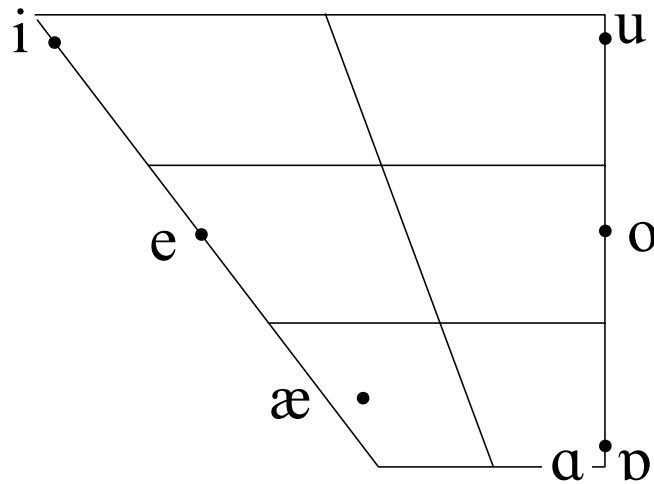
Figure 3. Vowels in Persian

From the literatures I reviewed, most hold the opinion that Persian has six distinct vowels. These are three front vowels [i] [e] [æ] and three back vowels [u] [o] [ɒ]. However, from my recording of my consultant, I think that among the recorded words there exists a seventh vowel [ɑ]. This might be a personal accent of hers, since most of the literatures stated that standard Persian has only six vowels.

Although old Persian has distinction between long and short vowels, nowadays modern Persian does not have any distinction in length of vowels in formal speech.

During my studies of related literatures, I found an interesting example of Persian vowels from [Hall 2007], which I paste below in Figure 4. These different vowel combinations give words different meanings, and reveal the distinctions between vowels. Unfortunately I did not get the chance to ask my consultant to record these words below.

Root Form

كرم   k_r_m

Five words derived from the Root Form

كرم	/kɒrɒm/	'benevolence'
كرم	/kerem/	'cream'
كرم	/kɒrm/	'vine'
كرم	/korom/	'chrome'
كرم	/kerm/	'worm'

Figure 4. Different vowel combinations with same root

In our word list, we can find all these seven vowels:
- [i]: 11 (who, [kʰi]), 12 (what, [tʃʰi]), 18 (many, [heili])
- [e]: 22 (one, [jek]), 24 (three, [se]), 54 (fruit, [mive])
- [æ]: 10 (there, [ændʒɑ]), 43 (father, [pædɑl]), 161 (fog, [mæh])
- [u]: 3 (he, [u]), 63 (meat, [guʃt]), 71 (hair, [mu])
- [o]: 2 (you (singular), [tʰo]), 5 (you (plural), [ʃomã]), 23 (two, [d̪o])
- [ɒ]: 35 (thin, [nɒzok]), 42 (mother, [mɒdɑl]), 45 (fish, [mɒhi])
- [ɑ]: 9 (here, [ind̠ʒɑ]), 10 (there, [ændʒɑ]), 39 (child, [bɑtʃʰe])

According to other papers, Persian has two diphthongs: [ei] and [ou]. But in our word list I can only find one dipthong [ei] in 14 (when, [kʰei]) and 18 (many, [heili]).

3.3. [ɑ] and [ɒ]

[ɑ] and [ɒ] are very similar vowels. They are articulated with the same position of tongue. The only difference between these two is: the former is an unrounded-lip vowel, while the latter is a rounded-lip vowel. I know the perception differences between them are

very subtle, but I indeed think that I heard two distinct vowels in the recording. To determine if both vowels exist in the recording, I take seven sample words for each suspected vowel and used the program Praat to analyze the formants of these recorded words.

Below are the analyzed plots from Praat [Boersma 2016]. In Figure 5 and 6, the upper section is a waveform of these samples (seven words for each vowel); the lower section is the spectrogram and formants. Blue lines represent fundamental frequency. Red points represent the formants.
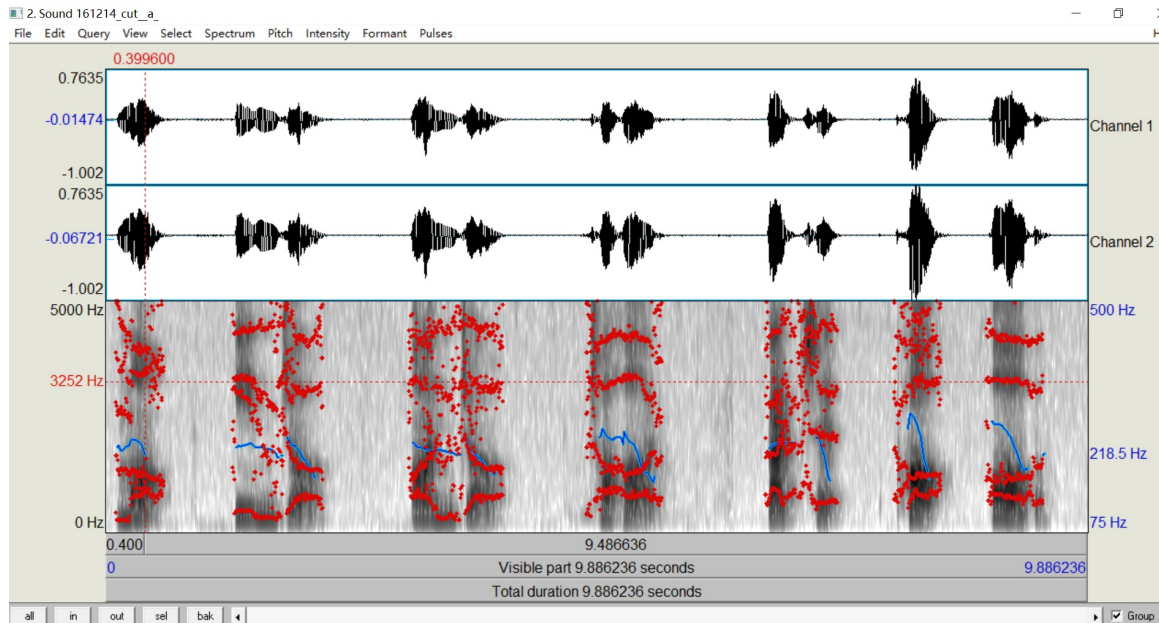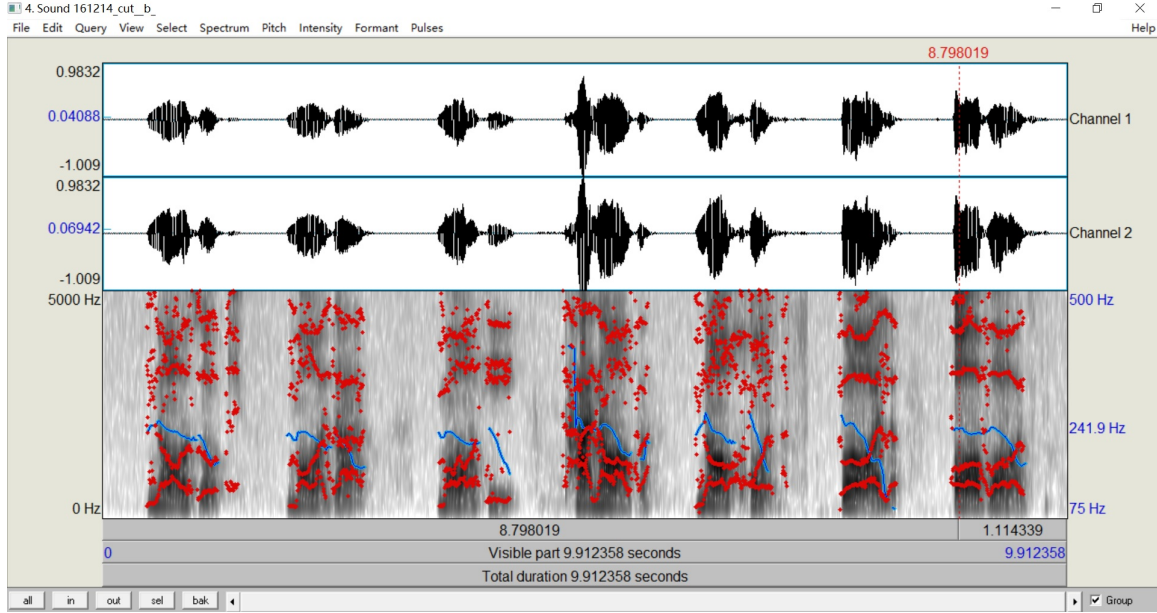

Figure 5. Samples of [ɑ]

Figure 6. Samples of [ɒ]

In Table 1 below, I enumerate the seven sample words for each suspected vowel, and used Praat to compute the specific formant values in theses samples. As what I percept, the words with number 4, 9, 10, 25, 39, 80, 150 has a vowel [ɑ], and the words with number 35, 42, 45, 61, 79, 84, 141 has a vowel [ɒ]. They are corresponding to the Figure 5 and 6 respectively.

Table 1. Formant analysis of [ɑ] and [ɒ]

Vowel	Number	Word	IPA	F1	F2	F3
ɑ	4	we	[mɑ]	845	1263	3392
	9	here	[indʒɑ]	723	1447	3022
	10	there	[ændʒɑ]	718	1452	2710
	25	four	[tʃahɑl]	790	1304	3278
	39	child	[bɑtʃʰe]	852	1769	2549
	80	foot	[pʰɑ]	805	1218	3245
	150	water	[ɑb]	791	1197	3263
ɒ	35	thin	[nɒzok]	813	1206	3127

42	mother	[mɒdɑl]	756	970	3297
45	fish	[mɒhi]	773	1183	3304
61	rope	[tʰɑnɒb]	863	1222	3143
79	fingernail	[nɒhõn]	763	1200	3118
84	wing	[bɒl]	721	1219	3083
141	sing	[ɒvɒz]	714	1094	3206
		[ɒvɒz]	721	1154	3157

As we know, F1 is throat cavity length related. F2 is mouth cavity length related. Throat cavity and mouth cavity are the features, and their combination of formants determines the different vowels qualities. From Table 1, we observe that F1 values in the first group (suspected [ɑ] group) range from 718 Hz to 852 Hz, while F1 values in the second group (suspected [ɒ] group) range from 714 Hz to 863 Hz. Apparently, the F1 for the two groups are basically in the same level. This fact indicates that the tongue positions of both groups are very low and in the same zone. However, F2 of the first group ranges from 1197 Hz to 1769 Hz, while F2 of the second group ranges from 970 Hz to 1222 Hz. Although the two ranges overlap slightly with each other, in general the F2 of the second group is still significantly lower than the first group. This observation tells us that the mouth cavity of the second group should be noticeably longer than the first group. From these analyses, we conclude that the two groups should be different vowels. Moreover, we know from the "cheatsheet" that the average F1 for [ɑ] is 850 and F2 is 1500, which falls into the range of the first group above. Therefore, we suspect that the first group has the vowel [ɑ]. We also know that [ɒ] is a rounded-lip vowel, compared to the counterpart unrounded-lip vowel, it has a slightly longer mouth cavity because of the rounded lips. From the observation above that the mouth cavity of the second group is longer, it echoes our assumption that the second group has the vowel [ɒ].

Although the above analysis supports my argument that there exists a [ɑ] vowel in the recordings, it cannot be easily concluded that standard Persian has this vowel. It might as

well be her personal accent or her habits of articulation that leads to this experiment result.

4. Consonants

According to our textbook, the principal airstream processes are:

"Pulmonic egressive: lung air pushed out under the control of the respiratory muscles

Glottalic egressive: Pharynx air compressed by the upward movement of the closed glottis

Glottalic ingressive: downward movement of the vibrating glottis; pulmonic egressive airstream may also be involved

Velaric ingressive: mouth air rarefied by the backward and downward movement of the tongue"

Persian speakers rely on Pulmonic egressive mechanisms to articulate.

In an overview of Persian consonants, it is based on an Arabic script, which is a consonantal system. The classification of Persian consonants is given in Table 3 below.

Table 3. Consonants

	Bilabial	Labiodental	Alveolar	Post Alveolar	Palatal	Velar	Uvular	Glottal
Plosive	p b		t d			k g	ɢ	ʔ
Nasal	m		n					
Trill			r					
Fricative		f v	s z	ʃ ʒ		x ɣ		h
Approximant					j			
Lateral approximant			l					

The classification is based on place of articulation (vertical columns) and manner of articulation (horizontal rows).

4.1. Stops

From my observation, Persian has seven stop consonants. They are:

- [p] 29 (wide, [pʰɑhn]) and [b] 34 (narrow, [bolik]),
- [t] 2 (you (singular), [tʰo]) and [d] 23 (two, [d̥o]),
- [k] 11 (who, [kʰi]) and [g] 73 (ear, [guʃ]),
- [ɢ] 59 (flower, [ɢol])

[p], [t] and [k] are voiceless and always strongly aspirated in all positions. [b] and [g] are voiced and we can hear them slightly palatalized sometimes.

There is glottal stop in Persian too, but we cannot find it in our word list. I found from other literatures two complementary examples:

- Ali (male name): [ʔɑli]
- impression: [tæʔsil]

4.2. Nasal

There are only two nasal consonants in Persian:

- [m] 4 (we, [mɑ]) and [n] 207 (name, [nɑm])

4.3. Trill

According to papers, Persian has an alveolar trill [r]. But I cannot find it in our examples. Furthermore, I ask my consultant and she thinks that Persian does not have a trill. This may be another problem of her personal accent.

4.4. Fricative

Persian has 9 fricative consonants. They are:

- [f] 98 (blow, [fut]) and [v] 204 (and, [vɑ])
- [s] 24 (three, [se]) and [z] 36 (woman, [zɑ̃n])
- [ʃ] 41 (husband, [ʃohɑl]) and [ʒ] 57 (root, [ʒiʃe])
- [x] 51 (tree, [delɑxt]) and [ɣ] 68 (horn, [ʃɑɣ])
- [h] 17 (all, [hɑme])

4.5. Approximant

Persian has only one approximant:

- [j] 22 (one, [jek])

4.6. Lateral Approximant

Persian has only one lateral approximant:

- [l] 34 (narrow, [bolik])

References

[Windfuhr 1979] Windfuhr, G. L. (1979). Persian. In B. Comrie (Ed.). (1987). The World's Major Languages (pp. 523-547): Oxford UP.

[Hall 2007] Phonological characteristics of Farsi speakers of English and l1 Australian English speakers' perceptions of proficiency. *Unpublished MA thesis. University of Curtin, Croatia.*

[Praat 2016] Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.23, retrieved 12 December 2016 from http://www.praat.org/

Attachments

Table 2. Word List

Number	Word	IPA
1	I	[mæ̃n]
2	you (singular)	[tʰo]
3	he	[u]
4	we	[mɑ]
5	you (plural)	[ʃomɑ̃]
6	they	[ænhɑ]
7	this	[ĩn]
8	that	[ɑ̃n]
9	here	[indʒɑ]
10	there	[ændʒɑ]
11	who	[kʰi]
12	what	[tʃʰi]
14	when	[kʰei]
16	not	[nɑ]
17	all	[hɑme]
18	many	[heili]
19	some	[bɑlhi]
20	few	[kʰɑmi]
21	other	[digɑl]
22	one	[jek]
23	two	[d̥o]
24	three	[se]
25	four	[tʃɑhɑl]
26	five	[pʰɑ̃ndʒ]

27	big	[bozolg]
28	long	[bolãnd]
29	wide	[pʰɑhn]
30	thick	[koloft]
31	heavy	[sãngĩn]
32	small	[kʰutfʰak]
33	short	[kutɑh]
34	narrow	[bolik]
35	thin	[nɒzok]
36	woman	[zãn]
37	man (adult male)	[mɑld]
38	Man (human being)	[bɑʃal]
39	child	[batʃʰe]
40	wife	[zãn]
41	husband	[ʃohɑl]
42	mother	[mɒdɑl]
43	father	[pædɑl]
44	animal	[heivãn]
45	fish	[mɒhi]
47	dog	[sɑg]
48	louse	[ʃepʰeʃ]
49	snake	[mɑl]
51	tree	[delɑxt]
52	forest	[dʒãngɑl]
53	stick	[tʃʰub]
54	fruit	[mive]
55	seed	[dɑne]
56	leaf	[bɑlg]
57	root	[ʒiʃe]
59	flower	[ɢol]
60	grass	[ɑlɑf]

61	rope	[tʰɑnɒb]
62	skin	[pʰust]
63	meat	[guʃt]
64	blood	[hun]
65	bone	[ostʰohãn]
68	horn	[ʃɑɣ]
69	tail	[dõm]
70	feather	[phãl]
71	hair	[mu]
72	head	[sɑl]
73	ear	[guʃ]
74	eye	[tʃʰeʃm]
75	nose	[bini]
76	mouth	[dɑhãn]
77	tooth	[dãndãn]
78	tongue	[zabãn]
79	fingernail	[nɒhõn]
80	foot	[pʰɑ]
81	leg	[pʰɑ]
82	knee	[zɑnu]
83	hand	[dɑst]
84	wing	[bɒl]
85	belly	[ʃekʰãm]
86	guts	[ʒude]
87	neck	[gɑldãn]
88	back	[kʰɑmɑl]
89	breast	[sine]
90	heart	[ɢɑlb]
91	liver	[dʒegɑl]
92	drink	[nuʃidãn]
93	eat	[holdãn]

95	suck	[mɑkʰidã̃n]
96	spit	[tʰof]
98	blow	[fut]
99	breathe	[nɑfɑs]
141	sing	[ɒvɒz]
148	moon	[mɑh]
150	water	[ɑb]
152	river	[ʒud]
161	fog	[mæh]
163	wind	[bɑd]
177	night	[ʃɑb]
180	warm	[ɢɑlm]
181	cold	[sɑld]
186	bad	[bɑd]
204	and	[vɑ]
207	name	[nɑm]