

A Convention for Transliteration of Persian into the Latin alphabet
(Writing Persian using the Latin alphabet)
Bijan Bakhshi

Introduction

More and more Iranians are using the Internet these days. There are sophisticated computer software tools around which enable us to write in Persian on most computers and send Persian e-mails. Although it's becoming easier and easier to use these tools, their use amongst Persians (Iranians) is not yet universal.

For the majority it is still easier and more practical to use Latin as the medium to write Persian computer messages. For the purpose of consistency and better communication between fellow Persian speakers, I suggest the following convention for writing Persian using the Latin alphabet.

This document is probably not the first of its kind; it is also not yet comprehensive. It does not attempt to teach the Persian or English grammar and simply refers to some of the rules. The examples given are biased towards the Persian which is spoken in Iran. It is however applicable to Persian spoken in Afghanistan, Tajikistan and elsewhere.

“The aim of any transliteration is to map one script into another script. It should be loss less such that the reader is able to reconstruct the original spelling of unknown transliterated words.” This is the main aim of the convention.

If you are viewing this document in Microsoft® Word, please de-select: **‘Check spelling as you type’** and **‘Check grammar as you type’** from the following option screen: ***Tools/Options/ Spelling & Grammar***. (Or the equivalent)

Please send your suggestions, corrections or criticisms to:
bijan@f2s.com or bijan1@totalise.co.uk

sepásgozár
Bijan Bakhshi

Most Persian letters have direct equivalents in Latin but for the sake of completeness I've listed them all. I have given many examples, which is probably the best way to demonstrate the convention. Some English translations are provided in cases where it helps to clarify the pronunciation.

Before launching into the more technical topics, here are some examples:

chandi az robaa'iaat-e omar-e khayyaam:

Dar kaargah-e kuzehgari raftam doosh
Didam dohezár kuzeh gooyá va khamoosh
Nágáh yeki koozeh barávord khorooash
Koo koozehgar o koozehkhar o kuzehforoosh

Man bi mey-e naab zistan natvaanam
Bi baadeh keshid-e baar-e tan natvaanam
Man bandeh-ye aan damam ke saaqi gooyad
Yek jaam-e degar begir o man natvaanam

Hargez del-e man ze elm mahroom nashod
Kam mánd ze asrár ke ma'loom nashod
Haftádo do sál fekr kardam shab o ruz
Ma'loom shodam ke hich ma'loom nashod

In kuzeh cho man ásheq-e zári bood ast
Dar band-e sar-e zolf-e negári bood ast
In dast ke bar gardan-e oo mibini
Dasti-st ke bar gardan-e yári bood ast

Ey rafteh be chogán-e ghazá hamchon goo
Chap mikhor o rást miro o hich magoo
Kánkas ke to rá afkanad andar tak o poo
Oo dánad o oo dánad o oo dánad o oo

Sarmast be meikháneh gozar kardam doosh
Piri didam mast o sabooé-i bar doosh
Goftam ze khodá sharm nadári ey pir
Goftá karam az khodá-st ro bádeh benoosh!

chand khat az sorood-e 'ey iraan':

Ey iraan ey marz-e por gohar
Ey khaakat sar cheshmeh-ye honar
Door az to andisheh-ye badaan
Paayandeh maani to jaavedaan
Ey doshman ar to sang-e khaare-i man aahanam
Ján-e man fadaay-e khaak-e paak-e mihanam
Mehr-e to chon shod pisheh-am
Door az to nist andisheh-am
Dar raah-e to key arzeshi daarad in jaan-e maa
Paayandeh baad khaak-e iraan-e maa
...

yek aadres:

Tehran khiábán-e Shomáli dast-e chap koocheh-ye sevvom (rooberooy-e farsh
forooshi-ye Javid), pelák-e haft(7) tabaqeh-ye avval áqá va khánoom-e Parsizaban

1 - The alphabet:

Persian Alphabet
Names

Examples of use with English translation in brackets

<i>alef</i>	<u>see Persian Vowels and Equivalents</u>
<i>be</i>	<i>boro</i> (go), <i>bad</i> (bad)
<i>pe</i>	<i>pedar</i> (father), <i>pesar</i> (son, boy),
<i>te</i>	<i>tamaam</i> (all, finished), <i>raftan</i> (to go)
<i>se</i>	<i>mosbat</i> (positive)
<i>jim</i>	<i>jaam</i> (cup), <i>jooraab</i> (socks)
<i>che</i>	<i>chonkeh</i> (because), <i>koochak</i> (small), <i>nooch</i> (sticky)
<i>he</i>	<i>hammaam</i> (bath), <i>rahmat</i> (mercy), <i>rooh</i> (spirit, ghost)
<i>khe</i>	<i>khiaabaan</i> (street), <i>makhloot</i> (mixed), <i>yakh</i> (ice)
<i>daal</i>	<i>dar</i> (door), <i>kodaam</i> (which), <i>bad</i> (bad)
<i>zaal</i>	<i>zehn</i> (mind, opinion), <i>mazkoor</i> (mentioned)
<i>re</i>	<i>rooz</i> (day), <i>aaraam</i> (tranquil), <i>maar</i> (snake)
<i>ze</i>	<i>zard</i> (yellow), <i>mazeh</i> (taste), <i>marz</i> (border)
<i>zhe</i>	<i>zharf</i> (deep), <i>mozheh</i> (eyelash), <i>ezhdehaa</i> (dragon), <i>dezh</i> (castle, camp)
<i>sin</i>	<i>sib</i> (apple), <i>mes</i> (copper), <i>mesvaak</i> (toothbrush)
<i>shin</i>	<i>shir</i> (milk, tap, lion), <i>ashk</i> (tear), <i>aash</i> (Persian soup)
<i>saad</i>	<i>sadr</i> (top), <i>sarih</i> (clear), <i>saf</i> (queue, line)
<i>zaad</i>	<i>zedd</i> (against), <i>zakhim</i> (thick)
<i>tein</i>	<i>teb</i> (medicine), <i>matloob</i> (pleasant)
<i>zein</i>	<i>zarf</i> (dish), <i>zolm</i> (cruelty)
<i>ein</i>	<u>see Persian Vowels and Equivalents</u>
<i>ghein</i>	<i>ghar</i> (cave), <i>ghaz</i> (goose), <i>ghayeb</i> (absent)
<i>fe</i>	<i>fil</i> (elephant), <i>aafat</i> (parasite), <i>ma'roof</i> (famous)
<i>qaaf</i>	<i>qashoq</i> (spoon), <i>maqsad</i> (destination)
<i>kaaf</i>	<i>koja</i> (where), <i>maskan</i> (place), <i>ashk</i> (tear)
<i>gaaf</i>	<i>garm</i> (warm), <i>agar</i> (if), <i>barg</i> (leaf)
<i>laam</i>	<i>lab</i> (lip), <i>kesalat</i> (being unwell), <i>bolbol</i> (nightingale)
<i>mim</i>	<i>moosh</i> (mouse), <i>amr</i> (order), <i>aaraam</i> (tranquil)
<i>noon</i>	<i>namak</i> (salt), <i>bini</i> (nose), <i>dahan</i> (mouth)
<i>ve</i> (consonant)	<i>vatan</i> (home country), <i>navakhtan</i> (to play music)
<i>ve</i> (vowel)	<u>see Persian Vowels and Equivalents</u>
<i>he</i>	<i>hamegi</i> (altogether), <i>hameh</i> baa <i>ham</i> (altogether)
<i>ye</i>	<u>see Persian Vowels and Equivalents</u>

Note: There is no need to use the following letters to represent any of the Persian letters: C, U, X. 'W' however can be used for Arabic words which need to be pronounced correctly, or by Kurdish speakers.

2 Persian Vowels and Equivalents:

The Arabic script, which is used in Persian writing, is Consonantal Alphabetic or Abjad. In these scripts vowels are not written and consequently two identical words can have two different meanings & pronunciations then depend on the context within a sentence.

Because of this, when using the Latin alphabet to represent Persian script, i.e. transliterating Persian script into Latin script, it is very important to explicitly represent vowels to stop any confusion.

2.1 Pure vowels

There are three pure vowels in Persian, which are represented by:

- **The ‘a’ sound**

The same sound as **a** in the English words ‘*add*’ or ‘*bad*’. Persian Examples

abr (cloud)
bad (bad)
maskan (place)
sabz (green)

- **The ‘e’ sound**

The same sound as **e** in the English words ‘*medical*’ or ‘*bed*’. Persian Examples

resideh (ripe)
del (heart)
sefr (zero)
pesar-e man (my son)
mey (wine),
meidaan (square),
peidaa (seen, found),
ney (straw, cane)
key (when)

Don’t spell as ‘*may*’

Don’t spell as ‘*midaan*’

Don’t spell as ‘*nay*’

Don’t spell as ‘*kay*’

- **The soft ‘o’ sound**

As in the soft ‘o’ in the Persian word *por* (full). In English there are many combinations of *o*, *u*, *w* and *a* as in: *boat*, *out*, *autumn*, *poor*, *road*, *blow* etc., but none of them accurately represent the ‘o’ part of this word. Here use a single ‘**o**’ to achieve the desired effect. More examples:

dom (tail)
mosbat (positive)
yek do seh ... (one two three...) Don’t spell as ‘*dau*’ or ‘*daw*’
man o to (me and you) Don’t spell as ‘*tau*’ or ‘*taw*’
polo (Persian style steamed rice) Don’t spell as ‘*polow*’

2.2 The strong ‘O’ sound or the soft form of ‘ve’

This form of ‘o’ is explicitly written and strictly speaking is not a vowel, as in *bood* (was), or *khoob* (good). Use ‘**oo**’ or ‘**u**’ and NOT other forms like ‘ou’ or ‘ow’. The choice should be made based on suitability, Examples:

<i>rooz</i> , or <i>ruz</i> (day)	<u>Don’t spell as ‘rouz’</u>
<i>door</i> or <i>dur</i> (far)	<u>Don’t spell as ‘dour’</u>
<i>goor</i> or <i>gur</i> (grave)	<u>Don’t spell as ‘gour’</u>
<i>rood</i> (river)	<u>Don’t spell as ‘roud’</u>
<i>doost</i> (friend)	<u>Don’t spell as ‘doust’</u>

2.3 The ‘aa’ sounds

There is no exact equivalent in Latin, to replace the letter *a* in the Persian word *nam* which means ‘name’ in English. *nam* with this spelling can also be read to mean ‘damp’. You can argue that the reader can guess which word is meant from the context of the sentence. It is however much clearer to have a distinct representation for each vowel, for this (**aa**) or (**á**) should be used. This way ‘*naam*’ (name) is differentiated from ‘*nam*’ (damp). If you cannot reproduce **á** (Ctrl+Alt+a) on your keyboard use **aa**. More examples:

<i>dam</i> (close, near)	
<i>daam</i> OR <i>dám</i> (domesticated animals)	
<i>kam</i> (little)	
<i>kaam</i> OR <i>kám</i> (palate, aim)	
<i>bad</i> (bad)	
<i>baad</i> OR <i>bád</i> (wind)	<u>Don’t spell as ‘bod’</u>

2.4 The yeh (‘i’) sound

There are three categories:

- Use of simple ‘**i**’. Examples:
 - miyeh* (fruit)
 - maahi* (fish)
- Where a word ends in ‘i’ and it is preceded by a vowel or ‘h’, in these cases use ‘**i**’. Examples:

<i>dampaa-i</i> (slippers)	<u>Don’t spell as ‘dampaa-ee’</u>
<i>gedá-i</i> (action of begging)	<u>Don’t spell as ‘gedá-ee’</u>
<i>bivafá-i</i> (unfaithfulness)	
<i>májerájoo-i</i> (adventure seeking)	
<i>paareh-i</i> (a part of, a piece of)	
<i>pahloo-i</i> (the one next to, next one)	
<i>goftéh-i</i> (a saying OR you’ve said)	
<i>tanhá-i</i> (loneliness)	
- Where a word starts or ends with ‘i’ sound or, has a strong ‘i’ in the middle of it, in these cases use ‘**y**’. Examples:
 - saayeh* (shade)
 - áyandeh* (future)
 - royaa* (dream)
 - mey* (wine)

2.5 The ‘*ein*’ sound (the 21st letter of Persian alphabet) & ‘*hamzeh*’

‘*ein*’ or ‘*ain*’ and the recently abandoned ‘*hamzeh*’ have no equivalents in Latin. There are three categories of words, which use these letters:

- Words starting with ‘*ein*’. For these, letters *aa*, *a*, *e* and *o* should be used. Examples:
aasheq (in love, lover)
atr (perfume)
edaalat (justice),
omr (life)
omq (depth)
eynak (spectacles, eyeglasses)
- Words with ‘*ein*’ in the middle or end. For these use *é* (Ctrl+Alt+e) or the apostrophe (‘ ’), if you cannot reproduce *é* on your keyboard. Examples:
baéd OR *ba’d* (next)
eétebár OR *e’tebaar* (credit)
motmaéen OR *motma’en* (confident, sure)
daévat OR *da’vat* (invitation)
sáéat OR *saa’at* (time, watch)
mosáéed OR *mosaa’ed* (favourable, agreeable)
moéattar OR *mo’attar* (fragrant, perfumed)
moéádel OR *mo’aadel* (equivalent)
moéaddel OR *mo’addel* (average, mean)
ejtemáé OR *ejtemaa’* (society, gathering)
ejtemáé-i OR *ejtemaa’-i* (sociable)
ertefaaé OR *ertefaa’* (height)
páéin OR *pá’in* (low, bottom)
aléán OR *al’án* (now, this minute)

3 Stressed letters (tashdid)

Any letter in a Persian word, which has ***tashdid*** on it, is simply repeated. Examples:

avval (first)
ammaa (but)
moddat (duration)
sonnat (tradition)

4 Prepositions, Possessive and Determiners.

Liaison - Where a word, ends in a vowel (or 'eh') and is followed by another sub-word, which starts with a vowel, the vowel (or 'eh') of the first word is dropped,

Use '-i', '-e' or '-ye' if a word ends in a vowel. The position of '-' is very important as it, links the preposition, determiner or the possessive to the intended word.

- **'of' Preposition.** Examples:
vasat-e khiábán (middle of the road)
- **Other Prepositions.** Examples:
az ánjá tá injá (from there to here)
az key? (From when)
az ki? (From who/whom)
az-am, az man (from me)
az-at, az to (from you)
az-ash, az-án, az-oo (from him/her/it)
az-amoon, az má (from us)
az-atoon, az shomá (from you)
az-ashoon, az ánhá, az-ishán (from them)
zir-esh, bálá-sh, pá'in-esh, roo-sh, nazdik-esh, pahl-oosh
- **Possesives.** Examples:
ketáb-e man (my book)
tabaqeh-ye maa (our floor)
hamsáye-ye shomá (your neighbour)
pedar-am OR pedar-e man (my father)
pesar-at OR pesar-e to (your son)
dast-esh OR dast-e oo (his/her hand)
ketáb-emoon OR ketáb-e má (our book)
lebaas-atoon OR lebás-e shomá (your clothes)
máshin-eshoon OR máshin-e aanhaa (their car)

chesm-há-ye man OR chesm-háam
chesm-há-ye to OR chesm-hát
chesm-há-ye oo OR chesm-hásh

chesm-há-ye má OR chesm-há-moon
chesm-há-ye shomá OR chesm-há-toon
chesm-há-ye áhá OR chesm-há-shoon
bará-ye má (for us)
- **Determiners.** Examples:
mard-e khoob (the good man)
mard-i khoob (a good man)
derakht-e boland (the tall tree)
derakht-i boland (a tall tree)
ketáb-e khoob (the good book)
ketáb-i khoob (a good book)

khiábán-e avval (the first street)
tabaqeh-ye dahom (the tenth floor)
too-ye yakhchál (in the fridge)
roo-ye zamin (on the ground)

5 Comparative and Superlative Adjectives

Examples:

bozorg, bozorg-tar, bozorg-tarin
koochak, koochak-tar, koochak-tarin
beh, beh-tar, beh-tarin

6 The Conjugate AND

The equivalent of 'and' should be denoted by 'o' or by 'va'. Don't use '-since 'and' separates words. Examples:

man o Bahman (me and Bahman)
in o oon (this and that, this person and that person)
ra'is va sekreter-ash (the boss and his secretary)

7 Plural forms

There are two ways to do this:

- Use '-haa' or '-há'. Examples:
bacheh-haa OR bacheh-há
khiaabaan-haa OR khiábán-há
miz-haa OR miz-há
- Use 'aan' or 'án' without '-'. Examples
doostaan OR doostán
mardomaan OR mardomán

8 Long words

Long words can optionally be broken by a '-' so that they can be read easier. The position of '-' should however be carefully selected. Examples:

aláqeh-mandán instead of *aláqehmandán*
dánesh-ámoozan instead of *dáneshámoozan*
dánesh-mandán instead of *dáneshmandán*
bar-resi (examination, inspection)
jost-o-joo (search)
goft-o-goo (debate)
raft-o-ámad (comings and goings)
zad-o-khord (conflict, flight)

9 Verb forms

Here is a list all possible (I think) tenses of 'raftan' (to go) for the first person:

rafteh-boodam, raftam, rafteh-am, miraftam, dáshtam-miraftam
miravam, daram-miravam, kháham-raft, beravam

10 Colloquial, slang or every day conversation (*zabán-e ámiáneh*)

The reader should not take this topic as a standard for speaking. This is covered for the sake of completeness.

In colloquial Persian, the third person of the verb '*boodan*' (To Be i.e. 'is') is '*ast*' or '*hast*', and is almost always dropped. There are two cases:

- **Simple Present Tense (*zamán-e hál*):** This is always shortened to '*e*', except when there is a liaison.

Examples without liaison:

'*in sib bad ast*' (this apple is bad) becomes '*in sib bad-e*'.

'*in ketáb-e man ast*' (this is my book) becomes '*in ketáb-e man-e*'

The same convention is used in some part of topic 4 above. Don't use '*in sib badeh*', or '*in ketáb-e maneh*'. 'h' should not be here.

Examples with Liaison:

'*sar-ash peidá ast*' becomes '*sar-esh peidá-s*'

'*in kháneh mál-e oon ágá-ast*' becomes '*in khooneh mál-e oon ágá-s*'

'*Mohammad bálá ast*' becomes '*Mohammad bálá-s*'

- **Present Perfect Tense (*zamán-e mázi-e naqli*):** 'ast' here is used as an auxiliary verb, and since it is always preceded with the past participle of another verb, it simply gets dropped, examples:

'*Ahmad rafteh-ast biroon*' becomes '*Ahmad rafteh biroon*'

'*máshin-e man kharáb shodeh ast*' becomes '*máshin-am kharáb shodeh*'

'*shám-ash rá khordeh ast*' becomes '*shámesh-o khordeh*'

Other examples with liaison (see 4 above),

'*panjerh-ye to*' becomes '*panjer-at*'

'*shisheh-ye oo*' becomes '*shish-ash*'

'*dar jostojoo-ye oo*' becomes '*dar jostojoo-sh*'

'*rooberoo-ye to*' becomes '*rooberoo-t*'

Other examples:

'*in dorost nemishavad*' becomes '*in dorost nemish-e*'

'*Mohammad fardá in rá mibarad*' becomes '*Mohammad fardá in-o mibar-e*'

N.B. There is a case where most Persian speakers sound as if they are using the Simple Past Tense (*mázi*) but they are actually using the Present Perfect Tense (*mázi-e naqli*), Example:

Correct: '*man dah daqiqeh-ye pish sham **khordam**, pas man shám **khordeh-am***'

Colloquial: '*man dah daqiqeh pish sham **khordam**(1), pas man shám **khordam**(2)*'

The difference between 1 and 2 is very subtle and it is the liaison, which makes the two tenses sound almost the same. These cases will be treated as if the speaker is using Simple Past Tense (*mázi*). More examples:

Correct: '*man dah daqiqeh-ye pish sham **khordam**, pas man shám **khordeh-am***'

Colloquial: '*man dah daqiqeh pish sham **khordam**, pas man shám **khordam***'

Correct: '*to dirooz in áqá rá **didi** pas to in áqá rá **dideh-i***'

Colloquial: '*to dirooz in áqá ro **didi** pas to in áqá ro **didi***'

11 Persian Names & Surname

In Persian a name followed by a surname is pronounced the same way as the ‘of’ or the possessive prepositions, you can optionally use the same convention examples:

Abolqásem-e Ferdosi
Ahmad-e Shámloo
Koorosh-e Hakhámaneshi

Or

Abolqásem Ferdosi
Ahmad Shámloo
Koorosh Hakhámaneshi

12 Capitalization

To cover the capitalization rules of English (or other European languages) is beyond the scope of this document, so we will list a subset of these for now: Capitalize the first letter of a word when the word is:

At the start a sentence, or after a full stop

After a question mark

An initial, name, surname and other titles of an individual

Name of a:

Place (Shiraz, Daneshgáh-e Tehran, Bandar-e Abbas, Takht-e Jamshid)

Month (Farvardin, Ordibehesht, Mehr)

Organization (Anjoman-e Honarmandán)

Holiday (Eid-e Norooz, Sizdah-bedar)

Religion (Eslám, Din-e Zartosht, Masihiyat)

Country (Espánia, Armanestán, Hend, Ámriká)

Language (Lori, Torki, Gilaki, Álmáni, Ingilsi)

Work of literature (Sháhnámeh-ye Ferdosi, Boostan-e Sa’di)

Work of art (Labkhand-e Zhokond)

...

...

13 Exception

To make life easier, place and personal names are exempt from this convention. This is because people are already used to spelling these in a certain way. This can be reviewed when the convention is more widely used. Example:

Bahram NOT Bahrám

Abbas NOT Abbás

Iran NOT Irán

Tehran NOT Tehrán

14 Punctuation

Persian punctuation rules should be observed, but use the Latin comma (,) and question mark (?).

15 Mixed language formats

All foreign words and sentences should be in double quotes. Examples:

Ádres-e doost-e man dar Ámrika hast “#23, 500th Street, New Moscow, 12345” va shomáreh-ye telefon-esh ham hast 00111 3567 3456789

When a Persian word is used in a foreign text, use any convention, which makes the word sound correct in that language. For example:

When I was in Tehran I had ‘chelow kabob’

‘*chelow kabob*’ breaks the conventions but should be allowed since for the English reader the correct form i.e. ‘chelo kabaab’ might not be as easy to pronounce.