

# ALEFBĀ-YE 2<sup>OM</sup>

TRANSCRIPTION OF PERSIAN USING THE  
LATIN SCRIPT BASED ON THE TRANSCRIP-  
TION PROCEDURE FOR IRANIAN TOPO-  
NYMIC ITEMS

SPRING 2014

## Contents

Contents.....	2
Foreword .....	2
Sounds and letters .....	2
Alphabet.....	2
Vowels.....	3
Consonants.....	3
Tašdid .....	3
Apostrophe.....	3
Omission, addition and shift of sounds.....	3
Allomorphs .....	3
Free spelling of some proper nouns .....	3
Hyphenation.....	3
Solid, spaced and hyphenated forms .....	4
Spaced form .....	4
Solid form .....	4
Hyphen .....	4
Numbers.....	4
Capitalization .....	5
Foreign words .....	5
Latin script.....	5
Other scripts.....	5
Abbreviations.....	5
General.....	5
Frequently used abbreviations .....	5
Punctuation.....	6
Period (.).....	6
Question mark (?) .....	6
Exclamation mark (!) .....	6
Comma (,).....	6
Semicolon (;) .....	6
Colon (:):.....	6
Quotation marks (""').....	6
Hyphen (-).....	6
Dash (–) .....	7
Ellipsis (...):.....	7
Parentheses ( ) .....	7
Square brackets [ ] .....	7
Slash (/) .....	7
Endnotes .....	8

## Foreword

The Latin script plays in today's world the role of a second script. For the proof of this assertion it is sufficient to take a look at the city and street signs or the Internet addresses in all countries. On the other hand, experience has shown that efforts to teach millions of Iranian young people abroad in reading and writing Persian mostly prove to be unsuccessful, due to the lack of daily contact with the Persian script. It seems that a way out of this dilemma has been found; and that is the use of the Latin script parallel to the Persian script.

Until recently, there was no official standard for the transcription of the Persian using the Latin script. Fortunately, all ministries and government agencies have now been instructed to implement a standard under the name [Transcription procedure for Iranian toponymic items](#)<sup>1</sup>. In this set of rules an attempt has been made to generalize the application of this standard so that one can write any Persian texts on the basis of this standard. Based on this set of rules infrastructure tools for spell checking, transcription of texts, translation of words, and language teaching are under development. Those interested can obtain the current versions of these tools for free at [alefbaye2om.org](http://alefbaye2om.org).

## Sounds and letters

### Alphabet

(1) Alefbā-ye 2<sup>om</sup> has 31 letters, of which 30 have a lower case and a capital form:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
p	q	r	s	t	u	v	w	x	y	z	ā	č <sup>2</sup>	š	ž
P	Q	R	S	T	U	V	W	X	Y	Z	Ā	Č	Š	Ž
,														

(2) Whenever the use of a diacritical mark or apostrophe is not possible due to practical limitations<sup>3</sup>, those signs are omitted:

sedā, seda; hamšahri, hamsahri; možde, mozde; mas'ul; masul

(3) The diacritical characters are together with their base letters in the alphabetical order. The apostrophe does not occur in the alphabetical order:

A (Ā), B, C (Č), D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S (Š), T, U, V, W, X, Y, Z (Ž)

(4) Each letter has a name:

a (ā), be, ce, de, e, ef, ge, he, i, je, ke, el, em, en, o, pe, qe, er, es (še), te, u, ve, dove, xe, ye, ze (že)

(5) Other Latin-based letters can occur for writing foreign words or proper names:

André, Miró

## Vowels

(6) Alefbā-ye 2<sup>om</sup> has 6 vowels a compound vowel:

Sound	Letter	Example
[æ]	a	abr, abzār, sag, dast
[ɒ:]	ā	āb, āftāb, bahār, rahā
[e]	e	emruz, esfand, sedā, sāde
[i:]	i	in, injā, bim, tehrāni
[o]	o	ostād, ordu, boz, do
[ow]	ow <sup>4</sup>	owbāš, jelow, pišrow, mowz
[u:]	u	bu, dust, jāru, xubi

(7) Short vowels are "a", "e" and "o". Long vowels are "ā", "i", "u" and "ow".

## Consonants

(8) Alefbā-ye 2<sup>om</sup> has 23 consonants:

Sound	Letter	Example
[b]	b	baste, bāz, abr, tab
[tʃ]	c	cerā, cand, āluce, gac
[d]	d	dar, derāz, pedar, sard
[f]	f	fanar, fer, sefid, kaf
[g]	g	gāv, gaz, tagarg, sag
[h]	h	hasti, hamin, rahā, rāh
[dʒ]	j	jib, jāru, bāje, kaj
[k]	k	kam, kenār, bikār, fandak
[l]	l	leng, lagad, mālidan, kacal
[m]	m	mu, mār, hamin, setam
[n]	n	nāz, narm, benām, tan
[p]	p	pās, pas, topol, gap
[ɣ]	q	qam, qalam, raqam, duq
[r]	r	rāst, raftan, sarā, tar
[s]	s	sib, sābun, tasbit, kas
[ʃ]	š	šab, šib, nešastan, farš
[t]	t	tāb, tāze, otu, taxt
[v]	v	vālā, vazeš, navid, gāv
[x]	x	xāb, xam, boxāri, paxme
[j]	y	yek, yār, māye, ney
[z]	z	zard, zohr, gozašt, rezāyat
[ʒ]	ž	žarf, žāle, vāže, kaž

[ʔ] ' ba'd, e'temād, mas'ul, sari'

## Tašdid

(9) In Arabic loanwords sometimes a consonant is emphasized. This phenomenon is referred to as *tašdid*. A consonant with *tašdid* is written twice:

mokarrar, tamannā, moaddel, ezzat

(10) A final consonant usually has no *tašdid*. Displaying *tašdid* follows the pronunciation:

xat, dastxat, xatti, xattāt

## Apostrophe

(11) Apostrophe is not displayed at the beginning of a word:

ābād, andišidan, onoq, azemat

(12) Apostrophe is almost never displayed between two vowels. Here, the respective vowels are pronounced separately. The apostrophe is only present between two vowels, if he has *tašdid*:

Aliābād, nowāmuz, sāat, moallem, miād, fa'āl

(13) Displaying the apostrophe follows the pronunciation otherwise<sup>5</sup>:

Firuzābād, badandiš, por'azemat, bad'onoq, jāme', jāmee, ketāb-e jāme-e ravānšenāsi, badi', badii

## Omission, addition and shift of sounds

(14) The omission, addition and shift of sounds in a word, is directly reflected in the writing<sup>6</sup>:

natawān, natvān; mehrbān, mehrabān, pišrow, pišravi

## Allomorphs

(15) Some Persian words have different allomorphs, i.e. they are pronounced differently. Spelling allomorphs follows the pronunciation:

cešm, cašm; ju, juy; be, bed; jelow, jolow, jelo, jolo

(16) Some allomorphs exist only in the colloquial language. The spelling of these allomorphs also follows the pronunciation:

bārān, bārūn; digar, dige; jāru-yam, jāru-m

## Free spelling of some proper nouns

(17) The spelling of some proper nouns, especially the name of a person, a company, a brand or a product may vary from this set of rules<sup>7</sup>:

Arash, Āraš; Saipa, Sāypā

## Hyphenation

(18) The transfer of a word part onto the next line is possible provided the sounds of a syllable cannot be separated in one word:

xā-ne; le-bāshā; lebās-hā, be dalil-e, be da-lil-e

## Solid, spaced and hyphenated forms<sup>8</sup>

### Spaced form

(19) Simple infinitives and verbs are written in the spaced form in all kinds of compound infinitives and verbs<sup>9</sup>:

yād gereftan, yād nagereftan, yād migiram, yād nemigiram; dar raftan, dar naraftan, dar miravam, dar nemiravam

(20) Proper names emerged from conjunctions with *e* and *o* are written in the spaced form, even if they have only one stress<sup>10</sup>:

Meydān-e Naqš-e Jahān, Xāvar-e Dur, Kašk-o Bādemjān

(21) If there is doubt whether a compound is present and/or the potential compound does not take any affix<sup>11</sup>, then the spaced form is possible. In this case partial compounds are written in the solid form:

piš az, pišaz; banā bar, banābar; banā bar in, banābarin; va gar na, vagarna; na tanhā, natanhā; dar vāqe', darvāqe'; be sor'at, besor'at; conān ke, conānke; hamin ke, haminke

### Solid form

(22) Prefixes and suffixes (whether flecional or derivational) are part of a word and are written in the solid form:

guyand, miguyand, nemiguyand, daftarhā, āqāyān, bozorgtar, bālātarin, bozorgi, dāneš, dānešmand, bimārestān, hamkār, benām, bāadab, bikār

(23) The solid form is the default form for writing compounds. Compounds not treated specially in this set of rules are written in the solid form<sup>12</sup>:

ketābxāne, toxmemorq, goftogu, pākkon, barfpākkon, yāddāšt, qulpeykar, qadboland, sarxorde, azxodgozašte, conin, conān, yekdigar, xištan, ānce, injā

### Hyphen<sup>13</sup>

(24) Whenever a word has no unbound use due to a sound change or the use of an allomorph, it is written in the hyphenated form together with that word, with which it is read directly together:

ke az, k-az; va in, v-in; be u, bed-u; man rā, ma-rā; ketāb rā, ketāb ro, ketāb-o

(25) Phonetically bound words having no stress (enclitics) are written in the hyphenated form together with their preceding words<sup>14</sup>. It does not matter whether

they build a compound or not. The position of the inserted sound (mediator sound) is after the hyphen:

doxtar-emān, man-o to, pesar-i, daftar-at; bidār-im, šāh-ā, ci-st, pedar-ešān-and, ketāb-e xub-i-st, xāne-ye man, amme-vo xāle, sedā-yi, jāru-yat, xeyl-i, xeyl-i ādamhā, xeyl-ihā

(26) In unfamiliar, occasion-related or longer compounds, one or more hyphens are set. The hyphens have to be set so that the partial compounds are common and the meaning of the word is not impaired:

raveš-e sang-dar-miyān, kam-dardesarsāz, tāze-be-dowrān-reside

(27) Words of an expression are hyphenated, if they are used as a word due to a need:

hamin man-bemiram-nāz-nakonhā

(28) A letter or an abbreviation is followed by a hyphen, if it is inflected, derived or compounded as a word due to a need:

n-om, g-hā, bz.-i, p-dār

(29) In compounds having two equal rank elements or compounds with repetition of words or semi repetition of words, a hyphen can be set:

irāni-ālmāni, rafte-rafte, jurāb-murāb

### Numbers<sup>15</sup>

(30) The components of a number less than one thousand are only separated by the conjunction *o*:

haft (7), davāzdah (12), panjāh-o hašt (58), nohsad-o haftād-o panj (975)

(31) Thousand, million, billion, ... are written separately:

Se milyun-o haštsad-o panjāh hezār-o pānsad-o cehel (3,850,540)

(32) For ordinals only the last part takes a suffix:

haftom (7<sup>om</sup>), davāzdahom (12<sup>om</sup>), panjāh-o haštom (58<sup>om</sup>), nohsad-o haftād-o panjom (975<sup>om</sup>), do hezārom (2,000<sup>om</sup>), bist-o panj hezārom (25,000<sup>om</sup>)

(33) Fractions and decimal numbers are written in the solid form:

sepanjom (3/5), bistosepanjom (23/5), bist-o sepanjom (20 3/5), dohezārom (2/1,000), bistopanjhezārom (0.025), bist-o dosadom (20.02), bistodosadom (0.22)

(34) Compounds consisting of numbers and other words are written in the solid form:

panjruze, bistopanjsāle, haftādopanjtāyi, sadhezārnafare, yāzdahdarsadi

(35) Estimating numbers are written in the hyphenated form:

se-cāhār nafar, do-se dāne, haftād-haštād tā, cehel-panjāh darsad

(36) Numbers in digits are written in the solid form together with suffixes but in the hyphenated form together with other words:

27om, 27i, 5-ruze, 25-sāle, 75-tāyi, 100,000-nafare, 11-darsadi

(37) The omission of digits in a number is directly reflected:

sāl-e 1393, sāl-e 93

## Capitalization

(38) The beginning of a sentence is capitalized:

Man raftam. U goft: "Man raftam." Xarid-e xub-i bud: Yek pirāhan-o yek šalvār xaridam.

(39) The beginning of a title or headline is capitalized. The capitalization of the other words of the title or the headline is possible:

Havādes-e emruz, Havādes-e Emruz

(40) Forms of address are capitalized before and after proper nouns<sup>16</sup>:

Ali Āqā, Āqā-ye Bahrāmi, Maryam Xānom, Xānom-e Panāhi, Šāh Abbās, Rezā Šāh, Karim Xān, Xāje Nasir

(41) Each main word from a group representing a proper noun is capitalized:

Sārā Panāhi, Tehrān, Sepidrud, Šāhnāme, Sāzmān-e Melal-e Mottahed, Jang-e Jahāni-ye Dovvom, Zereškolow bā Morq

(42) Proper nouns are not capitalized in derivations and compounds:

Xodā, Xodāyān, bāxodā, xodāyi, xodāšenās

## Foreign words

### Latin script

(43) Loanwords are written according to their Persian pronunciation:

pitzā, tāksi, rādiyo, oktobr, sigār

(44) The name of a person, a company, a brand or a product is written 1-1:

Wiliam Jones, BMW, New York Times, iPod, Windows

(45) If other foreign words are not written according to the Persian pronunciation for any reason, they are written in italics:

Niyoyork, *New York*; Landan, *London*

### Other scripts

(46) All words are written according to their Persian pronunciation:

moallem, cāy, Žāpon

(47) If the foreign word is not a loanword or the name of a person, a company, a brand or a product, then it will be written in italics:

*ader ka'san va nāvelhā, yo'refa mo'meno bessimā*

(48) The Arabic elements *abu* and *al* are always written in the solid form together with the following word. The element *al* is also written in the solid form together with the previous word, if there is a syllable merging between the two words. The element *ebn* is written in the solid form together with the previous word, if there is a syllable merging between the two words. It is written in the spaced form with the following word in conjunctions with *e*:

Abolqāsem, Abuali Sinā, Al'erāqi, Nasireddin, Ziyāolhaq, sariosseyr, Hoseynebn-e Ali, Isabn-e Ja'far, Ebn-e Batute

## Abbreviations

### General

(49) If a word is abbreviated with capital letters, then no point is set. Otherwise, an abbreviation ends with a dot:

Tehrān; TEH, Teh.

(50) All main elements should play a role in the abbreviation of a group or a compound. Compounds written in the solid form are considered as one word. Otherwise, every word is abbreviated, but the abbreviation is written in the solid form:

cāhārrāh, cr.; hejri-ye šamsi, h.š.

(51) With regard to the case-sensitiveness, the abbreviations having a dot follow their written out forms:

Dužite Panāhi, Du. Panāhi; Tehrān, Teh.; cāhārrāh, cr.

### Frequently used abbreviations

(52) For the calendar, the following abbreviations are available:

x. (xoršidi); h.x. (hejri-ye xoršidi); h.š. (hejri-ye šamsi); h.q. (hejri-ye qamari); m. (milādi, pas az milād); p.m. (piš az milād)

(53) For the Iranian names of the months, the following abbreviations are available:

FAR, far.; ORD, ord.; XOR, xor.; TIR, tir; MOR, mor.; ŠAH, šah.; MEH, meh.; ĀBĀ, ābā.; ĀZA, āza.; DEY, dey; BAH, bah.; ESF, esf.

(54) For the Christian names of the months, the following abbreviations are used:

ŽĀN, žān.; FEV, fev.; MĀR, mār.; ĀVR, āvr.; ME, me; ŽUA, žua.; ŽUI, žui.; UT, ut; SEP, sep.; OKT, okt.; NOV, nov.; DES, des.

(55) For the names of the weekdays, the following abbreviations are used:

ŠA, ša.; YŠ, yš.; DŠ, dš.; SŠ, sš.; CŠ, cš.; PŠ, pš.; JO, jo. yā ĀD, ād.

(56) For the time, the following abbreviations are used:

bd. (bāmdād), 09:15 bd.; bz. (ba'dazzohr), 02:00 bz.

## Punctuation

### Period (.)

(57) A narrative sentence or an indirect question ends with a period:

Man raftam. Az man porsid, cerā xošhāl-am.

(58) The decimal part of a number is separated by a period:

25.05, 2.735

### Question mark (?)

(59) A question ends with a question mark:

Kojā raft? Cegune? Na?

### Exclamation mark (!)

(60) An exclamation mark can be placed after an exclamation, a command, a wish, or a request:

Ce qašāng! Komak! Āfarin! Dar rā beband!

### Comma (,)

(61) A comma indicates mainly a short break in a sentence. This break can especially

- occur between a main and a subordinate clause:

Vaqt-i resid, Mahin hanuz ānjā bud.

- replace conjunctions *va* or *yā*:

Sobhhā boland mišavi, sobhāne-at rā mixori va be dānešgāh miravi. Rāmin, Narges yā Šivā ham mitavānad komak-at konad.

- include a brief explanation:

Rāmin, dust-am, xeyli lāqar-ast.

- occur after an exclaimed noun:

Ey ostād-e bozorgvār, az šomā mamnun-am.

- occur between a weekday and a date:

šanbe, 31<sup>om</sup>-e farvardin-e 1393; ša., 31-e far.-e 93

(62) A comma is placed after each three digits to the left for numbers with at least four digits:

2,000; 2,025.05, 2,000,000

### Semicolon (;)

(63) A semicolon separates sentences and phrases stronger than a comma but not as much as a period:

Mardom kār rā dust dārand; be dun-e ān fekr mikonand, ke zende nistand.

(64) The semicolon is used to separate the elements of a series, if they include different pairs/groups:

Dar in dānešgāh mitavān reštehā-ye pezeški-yo dandānpezeški; hoquq, eqtesād va jāmeešenāsi xānd.

### Colon (:)

(65) A colon occurs before a direct speech:

Āmuzgār porsid: “Ke pāsox rā midānad?”

(66) A colon occurs before a complementary series of elements separated by comma:

Mā be cand ciz niyāz dārim: sālon, musiqi, qazā va nušidani.

(67) A colon occurs before the translation of a word:

Tašxis: bāzšenāsi.

(68) A colon occurs before a sentence of a dialogue:

Navid: Cerā narafti? Narges: Vaqt nadāštam.

(69) Announcements and reasons follow a colon:

Šab rā ānjā gozarāndam: Māšin-am xarāb šode bud va ta'mirgāh baste bud.

(70) The hours and minutes are separated by a colon in time expressions:

07:45 bd.; 18:30

### Quotation marks (“”)

(71) Quotations and direct speech are enclosed by quotation marks:

“Doruq bozorgtarin gonāh be šomār miraft.” Āmuzgār porsid: “Ke pāsox rā midānad?”

(72) Text parts are highlighted by quotation marks in order to comment them on or to indicate distance to them.

First time used new terms are also highlighted by quotation marks:

bā “va” yā “yā”; Mahnāz fekr mikonad, ke to “afsorde” šodi. Morād-e mā yek onsor-e tarkibi-st, ke be ān “suratsāz” miguyand. Suratsāzhā bar do gune-and ...

### Hyphen (-)

(73) A hyphen replaces *tā* or *az ... tā* when specifying a range of values:

Negāh konid be s. 21-38. Dehxodā (1257-1334); Sāāt-e kār: ša.-cš., 09:00-18:00; Qatār-e Tehrān-Tabriz

(74) A hyphen replaces *be* or *bar* between two numbers:

Esm-e šaxs 1-1 nevešte mišavad. Irān 2-1 Koveyt rā šekast dād.

(75) Repetitions are avoided by a hyphen:

jodā- va sarhamnevisi

(76) A hyphen separates the month from the day and the year of a date:

31-01-1993, 31-01-93

### Dash (–)

---

(77) The dash indicates a pause in speech:

Hame tamām-e ruz dar entezār budand – yekdaf'e āmad.

(78) An longer explanation of the foregoing is enclosed by dashes:

Emruz sobh – hanuz dāštam bā mādar-am sobhāne mixordam – be man telefon kard.

(79) A speaker change may be indicated by a dash:

– Be pedar gofti? – Bale.

(80) A dash occurs before the name of the author or the source of a quotation:

“Doruq bozorgtarin gonāh be šomār miraft.” – Herodot

### Ellipsis (...)

---

(81) Parts of the text that are not important or can be guessed by the reader could be left by an ellipsis. A period is not needed for ellipsis at the end of the sentence:

Pišnahād-e Nasrin-o ... pazirofte šod. Bāzi-ye xub-i bud ... Fardā cekār mikoni?

### Parentheses ()

---

(82) The parentheses are used for explanatory additions or additional options. If a complete sentence is enclosed in parentheses, then the sentence period is placed before the second parenthesis:

ru(y) ≈ ru yā ruy; In ketāb rā (moteassefāne) hanuz naxānde bud. Diruz be bāzār raftim. (Parvin ham āmade bud.)

### Square brackets []

---

(83) Square brackets are used to indicate a replacement option:

ce[a]šm ≈ cešm yā cašm

### Slash (/)

---

(84) Alternatives can be expressed by a slash. If one of the alternatives include a space, then the slash is enclosed by spaces:

Vorudi-ye estaxr barā-ye kudakān/bozorgsālān 8,000/12,000 Tumān-ast. Dar bāre-ye safar be Āfriqā-ye Jonubi / Keniyā hanuz tasmim nagereftim.

## Endnotes

---

<sup>1</sup> [News agency ISNA, Eblāq-e Šivenāme-ye āvānegāri-ye kolli-ye nāmhā-ye joqrāfiyāyi-ye Irān, 21<sup>om</sup>-e esfand-e 1391.](#)

<sup>2</sup> This letter has only been taken for being conform to the Transcription procedure for Iranian toponymic items. It represents the same sound as "c" and occurs only in geographical names.

<sup>3</sup> For example, in the name of a homepage or in an email address.

<sup>4</sup> The letter "w" occurs only after "o". The compound vowel "ow" is always in one syllable.

<sup>5</sup> In words of Persian origin, the apostrophe is never displayed. The final consonant of the preceding syllable changes to the initial consonant of the next syllable (e.g. *narm* + '*af.zār*' > *nar.maf.zār*). But in Arabic loanwords, this change usually does not take place (e.g. *bad* + '*onoq*' > *bad'onoq*). On the other hand, the apostrophe at the end of a syllable in Arabic loanwords is eliminated, when the next syllable starts with an apostrophe (e.g. *Sam* + '*i*' > *sam. 'i*, *ba.di*' + '*i*' > *ba.di. i*).

<sup>6</sup> In some other languages, the sound omission or contraction is displayed using the apostrophe, which is a letter here. An extra sign for the sound omission does not improve the correct reading. In Persian, the sound omission almost always takes place in some specific and limited cases, allowing the reader to expect them. In the colloquial language, the sound omission takes place so repeatedly in these cases, that an extra sign for that would both challenge the writer as well as affect negatively the readability. (The Spanish, who have a phonetic script, do not use any sign for sound omission or contraction.)

<sup>7</sup> In the Transcription procedure for Iranian toponymic items, the apostrophe occurs between two vowels depending on the corresponding letter in the Persian script (Nāru'i vs. Aliābād).

<sup>8</sup> Basically, the purpose of a space is to mark the boundary between the words. Successive words sometimes form a compound word and therefore must be written together because they each represent a word. However, the determination of a compound word is not always easy.

<sup>9</sup> The non-verbal part is considered independently. Experts do not agree on many cases in which a compound could be present. This rule relieves the writer of the determination of compound verbs.

<sup>10</sup> The capitalization of the elements of these proper nouns prevents ambiguity. If one thinks of, for example, *kašk va bādemjān*, then one writes *kašk-o bādemjān*.

<sup>11</sup> The determination of a compound is usually easy and intuitive: A) *U kam howsele be xarj midehad*. B) *U yek ādam-e kamhowsele-ast*. Compound words have one stress: *kam+howsele* > *kamhowsele*. Compound words have the grammatical characteristics of a simple word. In A) *howsele* is a noun and has no comparative forms. But in B) *kamhowsele* is an adjective: *kamhowsele + tar* > *kamhowseletar*. However, there are few cases in which the determination of a compound is difficult. One of the reasons for this is that monosyllabic prepositions and conjunctions have no stress. If these words together with other words form a compound, the stress cannot be used as a determining characteristic. Another reason is that sometimes it is not absolutely clear whether there exists one or more than one stress. If the potential compound does not take any affix (flecnal or derivational) too, then the determination is even more difficult. This phenomenon usually occurs when a group of words appear in the role of a preposition, conjunction or a true adverb (an adverb, which is neither a noun nor an adjective nor a pronoun). Fortunately, the number of ambiguous cases is very limited.

<sup>12</sup> According to Prefixes and suffixes (whether flecnal or derivational) are part of a word and are written in the solid form:, derived words in compounds are written in the solid form: *rang+in* > *rangin+kamān* > *ranginkamān*. Similarly, derivations of a compound are written in the solid form: *asb+savār* > *asbsavār+i* > *asbsavāri*.

<sup>13</sup> As the reader is not the listener, one should assist him in determining the position of pause and stress. Likewise, longer compounds can represent a challenge to the reader. In addition, there are compounds that are formed due to a temporary need, so that the reader does not expect them therefore.

<sup>14</sup> Consequently, the possessive/objective pronouns, the short forms of *budan* (to be), the conjunctions *e* and *o*, the indefinite article *i* and the exclamation word *ā* are hyphenated. This rule helps in determining the position of the stress: *zamān-e* vs. *zamāne* or *ketāb-i* vs. *ketābi*.

<sup>15</sup> Due to specific characteristics of numbers they are treated entirely in this section.

<sup>16</sup> Proper nouns refer to names of persons, animals, things (products, works and objects), companies, organizations, geographic locations and unique events.