

Review of the Persian Writing System 2: Tajik(i) Persian

adapted from *Tajiki Textbook and Reader: Second Edition* (2003)

Tājiks began using an expanded adaptation of the Cyrillic script used in Russian in 1940 after an eleven-year experiment with the Roman alphabet. This means that by the end of the twentieth century Tājiks had used a Cyrillic writing system for sixty years in their 1,100-year old tradition of writing and literature. That pre-1929 heritage appeared in an adaptation of Arabic script. Tājiks share their heritage of pre-modern "Persian" writing and literature with Fārsi-speaking Iranians and Dari-speaking Afghans. Tājiks themselves refer to their Perso-Arabic writing tradition as «алифбои ниёгон» [alphabet of the ancestors].

In the 1990s, publication of Tājikī writing in Perso-Arabic dramatically increased, while legislation in Tajikistan called for replacing Cyrillic with Perso-Arabic script. Regardless of the outcome of the call for a return to Perso-Arabic, Tājikī experts will always need to read Tājikī materials in both scripts. This chapter offers a programmed, essentially self-instructional guide to the Perso-Arabic writing system.

* * * * *

§1. Perso-Arabic script consists of thirty-two letters and a handful of other symbols. Four features of this alphabet and writing system are distinctive in comparison with the Cyrillic writing system: (1) it is written from right to left; (2) it is basically cursive, most of the letters joining in typescript to a subsequent letter in the same word; (3) many letters have more than one shape; and (4) dots written above or below letter shapes often distinguish one letter from another.

Here follow Perso-Arabic alphabet letters and other symbols, one-by-one, step-by-step, in keyboarded forms in the context of words and sentences.

§2. The following combination of two letters is pronounced /jâ/ [as in the English word "jaw"] and is read from right to left: ج = ا + ج ←. Most of the letters in the Perso-Arabic alphabet, such as this letter ج called **jîm**, which represents the sound /j/, join in writing to the following letter of the same word.

The addition of another /jâ/ combination to ج produces /jâjâ/, which is written جاجا . Pronounce multisyllabic combinations and noun-adjective words with a stress or accent on the last syllable, unless otherwise indicated. Seven letters in the Perso-Arabic alphabet, among them the letter ا in the foregoing two combinations, do not join in writing to a following letter in the same word. The letter ا is called **alif** and is pronounced /â/ [as in the English word "jaw"] when written in the middle or at the end of a word. The addition of another /j/ sound to جاجا produces /jâjâj/, which is written: جاجاج ←.

Most of the letters in the Perso-Arabic alphabet have more than one letter shape. For example, **jîm**, the letter representing the sound /j/, has two shapes, one for the beginning or the middle of the word (ج) and a second for the end of a word or for the letter written by itself (ج). Some letters, such as the letter **alif**, which appears twice in the combination جاجا and there represents the sound /â/, have only one shape. Here are four combinations presented so far: جاجاج . جاجا . جاج . جا ←

§3. The letter **vâv**, which can be pronounced /û/ [as in the English word "food"] or /ü/ [as in German words spelled with "ü"], now substitutes for the letter **alif** in each of the above combinations, producing /jû/, /jûj/, etc. Except when noted in Cyrillic spelling (i.e., by «ȳ») accompanying Perso-Arabic items, pronounce **vâv** as /û/ in combinations and words:

جو . جوج . جوجو . جوجوج ←
[stream-uȳ]

Like the letter **alif**, **vâv** has only one shape and does not join a following letter in the same word. The seven letters in the Perso-Arabic alphabet which do not join in writing to a following letter in the same word have only one letter shape.

§4. The letter **yâ** is the new letter in the following combinations and there represents the vowel sound /î/ [as in the English word "feet"]. From the right, the first three combinations are /jî/, /jîj/, and /jîjî/:

جى . جيغ . جيغى . جيغيج . جيغو . جيغوج . جيغيج . جاجيج . جياج ←

The letter **yâ** has two shapes: ي for initial-medial position and ى for final-independent position. It also joins in writing to a following letter in the same word, as this combination جيغ shows. The letter **yâ** can also represent the sound /e(y)/ in the middle or at the end of a word. Pronounce **yâ** as /î/ when it functions as a vowel in words in this chapter unless accompanied by Cyrillic spelling featuring the letter «e».

§5. The letter **mîm** represents the sound /m/ and is the new letter in these words:

ما [we/us] . مو [hair-mÿ] . موم [wax] . جام [goblet] . ميم [mîm] ←

Like the letters **jîm** and **yâ**, **mîm** has two letter shapes, م and ى, and joins in writing to the following letter in the same word.

§6. The letter called **tâ** represents the sound /t/ [as in the English word "team"]. **Tâ** has two shapes: in initial or medial position it is written ت ; in final or independent position it is written ت. It also joins in writing to the next letter in the same word, as in the word توت [berry], pronounced /tût/. Both the initial-medial and the final-independent shapes of this letter have two dots over an elbow shape. The final-independent form of **tâ** is about three times as long as the initial-medial form and features the elbow shape at both ends. The letter **tâ** appears in all of the following combinations with the five letters previously learned (**alif**, **jîm**, **mîm**, **vâv**, and **yâ**):

تا [till] . تو [inside-tȳ] . تى . تاج [crown] . توج . تيغ . تام [complete, full] .
توم . تيم [covered bazaar] . تيتا . تيتو . تيتى . تات . توت [mulberry] . تيت ←

A comparison of the initial-medial shape of **tâ** to the initial-medial shape of **yâ** shows that the two letters belong to the same "family" of letters, combining the elbow shape and dots above or below it. Six letters combine this elbow shape with from one to three dots above or below it.

§7. The letter **zâ**, which represents the sound /z/, is similar to the letter **vâv** (و) in that they both have a single curved shape (ر) and do not join to the next letter in a word. The letter **zâ** appears in these combinations:

زات . زوت . زيت . تازاج . توزاج . تيزاج . ←
 توز . موز . جوز . ميز [table] . تاز . جاز . تيز [fast, sharp- te3]

§8. The letter **sîn** represents the sound /s/, has two shapes, س and س, and joins the next letter in the same word. The letter **sîn** appears in these words:

سارا [Sara] . سام [Sam] . ساز [musical instrument] . جاسوس [spy] . سيم [wire, silver]

The final-independent shape of س, when compared to the final-independent shapes of ج , م , ت , and ي , reveals an important fact about letters with multiple shapes: most of them have a larger or longer shape in final-independent position than in initial-medial position. But Perso-Arabic script does not have "capital" letters at the beginning of words.

§9. The letter **lâm** represents the sound /l/ [as in the English word "love"]. **Lâm** has two shapes: ل and ل , and joins with a subsequent letter in the same word, as these words show:

ليمو [lemon] . سال [year] . لال [dumb] . مال [property] ←

The letter **lâm** combines in a distinctive way with the letter **alif**. The word لال [dumb] illustrates the initial **lâm-alif** combination; the combination ميلا , pronounced /mîlâ/, illustrates medial **lâm-alif**.

§10. The letter **tâ-yi 'arabî**, which has only the single shape ط and which joins to a subsequent letter in the same word, represents the same /t/ sound as **tâ** (تت). In a number of cases, more than one letter represents a single sound, although such letters are not usually interchangeable in spelling words. The letter **tâ-yi 'arabî** appears in these words.

طاس [basin, bowl] . طوطى [parrot-تُطَي] . طول [length-تُول] ←

§11. The letter **kâf**, which has two shapes: ك in initial-medial position and ك in final-independent position, represents the sound /k/ [as in the English word "cream"]. These words illustrate the letter **kâm**:

كيسـت [who is ...?] . تاجـيك [Tajik] . تاك [vine] . تاكزار [vineyard] . ←
 كام [palate; desire] . كالا [goods] . كى [who?]

§12. The letter **fâ** has two shapes: ف and ف , joins in writing to the next letter in the same word, and represents the sound /f/. It actually resembles the two shapes of the letter **tâ** (تت). But **fâ** has a loop and a single dot above it, whereas **tâ** combines the basic elbow shape with two dots above. The letter و , which can represent the sounds /û/ and /ü/ when written in the middle or at the end of a word, should not be confused with initial-medial **fâ** (فف), which has a dot above it and joins the next letter in the same word. Sometimes the loops of **mîm**, **vâv**, and **fâ** appear in keyboarded texts as circular dots with the loop inked in. These words illustrate the letter **fâ**:

[film]	فیلم	[foot]	فوت	←
[kâf]	کاف	[elephant]	فیل	[enough] کافی

§13. The three letters **alif**, **vâv**, and **yâ** can represent the vowel sounds â(w)/, /û/-/ü/, and /î/-/e(y)/, respectively, when these letters appear in the middle or at the end of a word, as in these examples.

[enough] /kâfi/	کافی	[crown] /tâj/	تاج	← /â/
[spy] /jâsûs/	جاسوس	[mulberry] /fût/	توت	← /û/
[who?] /ki/	کی	[movie] /film/	فیلم	← /î/

A fourth letter used to represent a vowel sound is **hâ**, written ه, which can represent the vowel sound /a/ [as in the English word "fad"], but only as the last letter of a word. The letter **hâ**, pronounced /â/, appears in these words:

[bowl] /kâsâ/	کاسه	[tulip] /lâlá/	لاله	←
---------------	------	----------------	------	---

The letter **hâ** can also represent the sound /h/ [as in the English word "hot"] at the end of a word, for example, in these words:

[straw] /kâh/	کاه	[mountain-κÿx]	کوه	[black] /siyâh/	سیاه	[moon, month] /mâh/	ماه	←
---------------	-----	----------------	-----	-----------------	------	---------------------	-----	---

The letter **hâ** at the end of a word usually represents the sound /â/ when the letter before it is a consonant and the sound /h/ when the letter before it is a vowel. The letter **hâ** is one of a number of letters which can represent more than one sound.

§14. The symbol ه whether it represents the sound /a/ or the sound /h/ is the final-independent shape of a letter whose initial shape is ه and whose medial shape is ه. In its initial and medial shapes the letter **hâ** represents the sound /h/, as in these combinations:

هیه	-	ماهی	-	میهی	-	ماه	-	لاله	←
[tulip]		[month]		/mihi/		[fish]		/hihâ/	

§15. A letter similar to و and ز, insofar as it has only one letter shape and is not joined in writing to a following letter in the same word, is د, called **dâl**, which represents the sound /d/. The letter **dâl** appears in all of these words:

دود	-	بود	-	دید	-	سود	←
[profit]		[he/she saw]		[he/she/it was]		[smoke]	

§16. A letter shaped exactly like سین (سس) but taking three dots above it is **shîn**, which has the two shapes ش and ش and represents the sound /sh/ [as in the English word "shoe"]. **Shîn** appears in these words:

[urine]	شاش	[mouse]	موش	[happy]	شاد	[shah]	شاه	←
---------	-----	---------	-----	---------	-----	--------	-----	---

§17. A letter shaped exactly like **kâf** (كك) and taking two lines over it is **gâf**, which has the two shapes گ and ک and represents the sound /g/ [as in the English word "go"]. The /g/ sound of **gâf** is pronounced clearly in every position **gâf** may occur in a word (just as the letter **kâf** for /k/); whereas in English, the sound /g/ is often only slightly pronounced in syllable-final and word-final positions. The letter **gâf** appears in these words:

گاہ - گاہ گاہ - گیج - گیلان - گوساله - گاف
 /gâf/ [calf-گوساله] [cherry-گیلاس] [bewildered] [from time to time] [time]

§18. A letter shaped exactly like **jîm** (جج), but taking three dots below it, is **che**, which has the two shapes چ and چ represents the sound /ch/ [as in the English word "church"]. It appears in these words:

چاہ [well] چی [what?] چیز [thing] چیست [what is it?]

§19. A letter shaped exactly like **tâ** (تت) and **yâ** (یی) in their initial-medial shapes is **nûn**, which exhibits one dot above its two shapes ن and ن and which represents the sound /n/. Each of these letters, **tâ**, **yâ**, and **nûn**, has a distinctive final-independent shape: **tâ** (تت), **yâ** (یی), and **nûn** (نن). The letter **nûn** appears in the following words:

نام - نان - نیم، نیمه - ماشین، ماشینه - سینه - شادمان - نادان، نادانی
 [ignorant, ignorance] [happy] [breast, chest] [car] [half] [bread] [name]

§20. All of the letters in the Perso-Arabic alphabet are considered primarily to present consonant sounds. However, the kinds of words illustrated thus far demonstrate that the three letters **alif**, **vâv**, and **yâ**, when not the first letter of a word, may represent, respectively, the vowel sounds /â/, /û/-/ü/, and /î/-/e(y)/. And when the letter **hâ** is the last letter of a word, it often represents the vowel sound /â/.

If a word begins with the vowel sound /â/, /û/-/ü/, or /î/-/e(y)/—except in the case of Arabic loanwords beginning with the letter ‘ayn followed by one of these three sounds (as described below)—, it is written as follows:

/â/ آ = ا + ا آزاد [free] /âzâd/ آسیا [Asia] /âsiyâ/
 /û/- /ü/ او = و + ا او [he/she, him/her-y]
 /e(y)/ - /î/ ای = ی + ا این [this] /în/ ایستگاه [bus stop]

∞ When the letter **alif** begins a word, it is considered a consonantal pause called a glottal stop before the word's first sound. Thus, the word آسیا, pronounced /âsiyâ/, begins with an **alif** written horizontally representing /ʔ/ followed by a second **alif** written vertically, representing the vowel sound /â/.

آن [that] آنها [they, those] آن جا [there] آسمان [sky]

∞ When the letter **vâv** begins a written word, it represents a /v/ sound, as in these words:

وادی [valley] ویژه [visa] ویتامین [vitamin]

⇒ When the letter **yâ** begins a written word, it represents a /y/ sound, as in these words:

← یونانی [Greek] یاد [memory] یا [or - ē]

⇒ The letter **vâv** in the middle or at the end of a word can represent /û/, /ü/, /v/, /u/. For example:

[profit]	суд	سود	[smoke]	дуд	دود	← /û/
[mouse]	муш	موش	[spy]	чосус	جاسوس	
[friend]	дӯст	دوست	[he/she]	ӯ	او	← /ü/
[short]	кӯтоҳ	کوتاه	[ear]	гӯш	گوش	
[cow]	гов	گاو	[valley]	водӣ	وادی	← /v/
[you-1]	ту	تو	[two]	ду	دو	← /u/

⇒ The letter **yâ** in the middle or at the end of a word can represent /i/, /e(y)/, /y/, or /ay/ [between the vowel sound in the words "sigh" and "say"]. For example:

[What is it?]	чист	چیست	[table]	миз	میز	← /i/
[waist middle]	миён	میان	[black]	сиёҳ	سیاه	
[no(t), nothing]	ҳеҷ	هیچ	[fast, sharp]	тез	تیز	← /e(y)/
[he/she/it is not]	нест	نیست	[we build]	месозем	میسازیم	
[Asian] /âsiyâyi/	осиёӣ	آسیایی	[tea]	чай	چای	← /y/
[he/she]	вай	وی	[when?]	кай	کی	← /ay/

§21. Read these sentences.

- | | | |
|------------------|-------------------------|---------------------|
| Вай тоҷик. | [He/she is Tâjik.] | ۱. وی تاجیک. |
| Онҳо тоҷик. | [They are Tâjik.] | ۲. آنها تاجیک. |
| Ӯ шод нест. | [He/she is not happy.] | ۳. او شاد نیست. |
| Мо тоҷик нестем. | [We are not Tâjik.] | ۴. ما تاجیک نیستیم. |
| Мо дӯстем. | [We are friend(s).] | ۵. ما دوستیم. |
| Чой ин чо нест. | [There is no tea here.] | ۶. چای اینجا نیست. |
| Вай кист? | [Who is he/she?] | ۷. وی کیست؟ |
| Он чист? | [What is that?] | ۸. آن چیست؟ |

§22. Use of **alif** in front of **vâv**, **yâ**, or another **alif** is not the only way to represent the sound /û/, /ü/, /î/, /e(y)/, or /â/ at the beginning of a word. The letter '**ayn**', which has four shapes, ع عع , also serves this function. The letter '**ayn**' represents a glottal stop, which can precede cited vowel sounds at the beginning of a word. For example, the following words begin with one of the above vowel sounds:

/âj/ عاج [ivory-оч]	/îd/ عيد [holiday-ид]	/âli/ عالى [excellent, supreme - олий]
---------------------------	-----------------------------	--

Wherever the letter '**ayn**' appears in a word, it signals a pause or break in the flow of sounds. For example, the word « شيعى » [Shi'ite Muslim], is pronounced /shî'î/. Another form of the same word is « شيعه », pronounced /shî'a/.

§23. Remaining letters in the Perso-Arabic alphabet are mainly "family" members of already introduced letter types. Their distinguishing features relate mostly to the number and placement of dots associated with their shapes.

§24. Three other letters share with **tâ** (ت), **nûn** (ن), and **yâ** (ي) a similar initial-medial shape. They are **bâ**, **pâ**, and **sâ**.

shapes		sound	name
final-independent	initial-medial		
ب	ب	/b/	← bâ
پ	پ	/p/	← pâ
ث	ث	/s/	← sâ

Final-independent shapes of these three letters are also similar to that of **tâ** (ت), whereas the other two members of this family of six letters, **nun** (ن) and **yâ** (ي), have distinctive final-independent shapes. The letters **bâ**, **pâ**, and **sâ** (which represents the same sound as **sîn**) appear in these words:

[apple-себ]	سيب	[grandmother]	بيبي	[bank]	بانک
[jeep]	جيپ	[end, down-поён]	پايان	[leg]	پا
[legacy-мерос]	ميراث	[second (of time)-сония]	ثانيه	[(the) second]	ثانى

§25. The **jîm** family of letters has four members. **Jîm** itself (جج) and **che** (چچ) have already been described. The other two members are: **hâ-yi huttî** and **khâ**:

shapes		sound	letter-name
final-independent	initial-medial		
ح	ح	/h/	↓ hâ-yi huttî
خ	خ	/kh/	↓ khâ

Hâ-yi huttî (حح) is one of two letters used to represent the sound /h/. The other is **hâ-yi havvaz** (ههه), which can also represent the sound /â/ at the end of a word. Although both of these /h/ letters are called **hâ**, their full names distinguish them from each other:

hâ-yi hutfî and **hâ-yi havvaz**. The letter **khâ** (خخ) represents a sound between /h/ and /k/ that is foreign to speakers of English. This sound is like "ch" in German or Scottish words such as "nicht" and "loch." Speakers of English need to practice pronouncing **khâ** to avoid mistakenly pronouncing /h/ or /k/ instead. The following words illustrate the letters **hâ-yi hutfî** (حح) and **khâ** (خخ):

حال - حاجی - خاک - خاکی - خوب - ساخت - سیخ - میخ
 [nail-mex] [kabob skewer] [structure] [good, well] [khaki] [dirt, dust] [Haji] [condition]

§26. In some common words, such as the following, which feature the letter **vâv** (و) after the letter **khâ** (خ), the **vâv** is unpronounced, but always written.

خوآب - خواست - خویش - خوانا - خوست
 [sleep, dream] [will, wish] [relative, self] [readable]

§27. The letter **khâ** (خخ) is one of four consonant letters which represent sounds unfamiliar to speakers of English. The other three are: **râ**, **ghayn**, and **qâf**.

independent	shapes			sound	letter name
	final	medial	initial		
ر	ر	ر	ر	/r/	râ
غ	غ	غ	غ	/gh/	ghayn
ق	ق	ق	ق	/q/	qâf

The letter **râ**, which has a single shape, looks like **zâ** (ز) without a dot and does not join to a following letter in the same word, represents the sound /r/, but not the English 'r'. Tâjîkî /r/ is a flap or trill, which means that it is formed by touching the top of the tongue to the front part of the roof of the mouth one or more times in mouthing the letter.

The letter **ghayn** (غغغ), which looks like 'ayn (ععع) with a dot above it, represents the sound /gh/, basically the same as the French 'r'.

The letter **qâf** (قق), which looks like **fâ** (فف) with two dots in its initial-medial shape and like **nûn** (نن) in its final-independent shape, represents the sound /q/, a sound between /k/ and /g/ foreign to speakers of English. The sound which **qâf** represents is best learned through imitation of native speakers. The following words illustrate the letters **râ** (ر), **ghayn** (غغغ), and **qâf** (قق):

توریست - فارسی - گیتاره
 [tourist] [Persian language of Iran] [guitar]
 غاز - تیغ - باغ
 [goose] [sword, dagger, blade-ref] [garden]
 قالی - قالین - موسیقی - طاق
 [oriental carpet] [music]

§28. The following nine letters in the Perso-Arabic alphabet have only one shape:

vâ و - zâ ظ - tâ ط - zhâ ژ - zâ ز - râ ر - zâl ذ - dâl د - alif ا

Of these letters, only **tâ-yi 'arabî** (ط) and **zâ-yi izghî** (ظ) join to the next letter in a word. The letter **zâl** (ذ) looks like **dâl** (د) with a dot added and represents the sound /z/ just like **zâ**. **Zhâ** (ژ) looks like **râ** (ر) and **zâ** (ز) with three dots and represents the sound /zh/. The letter **zâ** (ظ), which looks like **tâ** (ط) with a dot added, also represents the sound /z/ just like **zâl** and **zâ**. These items illustrate the letters **zhâ** (ژ), **zâl** (ذ), and **zâ** (ظ).

ژولیده - میژه - سوژه - ژاله - زال - ظا - ذات
 [dishevelled] [eyelash] [plot of story] [hail] /zâl/ /zâ/ [essence, nature]

§29. The letters **sâd** (صص) and **zâd** (ضض) are a similarly shaped pair of letters, the dot over **zâd** being the distinguishing feature, which represent the sounds /s/ and /z/ respectively. **Sâd** is one of three letters which represent /s/. The other two are **se** (ثث) and **sîn** (سس). **Zâd** is one of four letters which represent /z/. The other three are **zâl** (ذ), **zâ** (ز), and **zâ-yi izghî** (ظ). The following words illustrate the letters **sâd** (صص) and **zâd** (ضض):

صاف - صوفی - صیغه - خاص - راضی
 [clear, pure] [Sufi] [Shi'ite marriage contract, verb mood] [special] [satisfied]

§8.30. Here follows the whole Perso-Arabic alphabet in alphabetical order. Its memorization facilitates the use of dictionaries.

shape(s)		name	letters		shape(s)	name
<i>naskh</i>	typescript		<i>naskh</i>	typescript		
ص		sâd	ا		alif	
ض		zâd	ب		bâ	
ط		tâ, te-yi arabî	پ		pâ	
ظ		zâ-yi izghî	ت		tâ	
ع		ayn	ث		sâ	
غ		ghayn	ج		jîm	
ف		fâ	چ		che	
ق		qâf	ح		hâ-yi huttî	
ک		kâf	خ		khâ	
گ		gâf	د		dâl	
ل		lâm	ذ		zâl	
م		mîm	ر		râ	
ن		nûn	ز		zâ	
و		vâv	ژ		zhâ	
ه		hâ-yi havvaz	س		sîn	
ی		yâ	ش		shîn	

§31. Read the following list of words, checking Cyrillic script versions when in doubt. Review foregoing descriptions of Perso-Arabic letters in cases of pronunciation questions. Words appear in Perso-Arabic alphabetical order.

орзу	[wish, desire]	آرزو	←	об	[water]	آب
Осиё	[Asia]	آسيا		осмон	[sky]	آسمان
офтоб	[sun(shine)]	آفتاب		ошхона	[kitchen, dining area]	آشخانه
оху	[deer]	Аҳо		он, онҳо	[that, those]	آن, آنها
ин чо	[here]	این جا		ин	[this]	این
боб	[chapter]	باب		бо	[with]	با
борон	[rain]	باران		бобо	[grandfather]	بابا

бог	[garden]	باغ	←	бозор	[bazaar, market]	بازار
бист	[twenty]	بیست		буй	[smell, odor]	بوی
пул	[money]	پول		по	[foot, leg]	پا
пир	[old (adj), elder (n)]	پیر		печ	[twist, turn]	پیچ
точик	[Tâjik (person)]	تاجик		точ	[crown]	تاج
сония	[second (of time)]	ثانية		тут	[mulberry]	توت
чой	[place]	جای		чон	[life, soul]	جان
чой	[tea]	چای		чоп	[type, print(ing), edition]	چاپ
чиз	[thing]	چиз		чи	[what]	چی
холо	[now]	حالا		хол	[condition]	حال
хоб	[sleep] /khâb/	خواب		хок	[dirt]	خاک
хуб	[good, well]	خوب		хост	[wish, will] /khâst/	خواست
лору	[medicine]	دارو		хӯрок	[food]	خوراک
рост	[straight, right]	راست		дӯст	[friend]	دوست
роҳ	[way, road]	راه		розй	[satisfied, in agreement]	راضی
рӯз	[day]	روز		рӯ	[face]	رو
русй	[Russian (e.g., lang.)]	روسی		рӯзнома	[newspaper]	روزنامه
зуд	[soon, fast]	زود		риш	[beard]	ریش
жола	[hail]	ژاله		зиёд	[a lot, much]	زیاد
соя	[shade]	سایه		сол	[year]	سال
шом	[evening]	شام		сиёх	[black]	سیاه
соф	[clear, smooth]	صاف		шоҳ	[king, shah]	شاه
форсй	[Fârsi (language)]	فارسی		тӯтй	[parrot]	طوطی
қофия	[rhyme]	قافیه		филм	[film, movie]	فیلم
кӯҳ	[mountain]	کوه		кор	[work]	کار
кй	[who?]	کی		кай	[when?]	کی
гӯш	[ear]	گوش		гов	[cow]	گاو
лиму	[lemon]	ليمو		лола	[tulip]	لاله
мошина	[car]	ماشینه		мошин	[car]	ماشین
муш	[mouse]	موش		моҳ	[month, moon]	ماه
митинг	[meeting]	میتینگ		миён	[middle, waist, center]	میان
нон	[bread]	نان		ном	[name]	نام
ёд	[memory]	یاد		ё	[or]	یا

§32. The Perso-Arabic alphabet has shortcomings as a representation for the contemporary Tājīkī language: (1) it uses the letter « و » to represent «в», «у», and «ӯ»; (2) it uses single letter « ى » to represent «й», «и», «...й», «э», «ай», and «...е»; (3) it does not represent short vowel sounds in such words as « من » [ман], « شما » [шумо], and « فكر » [фикр]; and (4) it does not represent the sound «и» in «изофат» constructions.

§32.1. Listen to a modelling of these words featuring « و » in Perso-Arabic transcription vis-à-vis «у» or «ӯ» in the Cyrillic writing system.

гӯш	[ear]	گوش	дур	[far, distant]	دور
шӯхӣ	[joke, prank]	شوخی	нур	[light]	نور
кӯча	[street]	کوچه	русто	[village]	روستا
рӯз	[day]	روز	хуб	[good, well]	خوب
дӯст	[friend]	دوست	пул	[money]	پول
кӯх	[mountain]	کوه	суд	[profit, use]	سود
ӯ	[he/she]	او	рус	[Russian (person)]	روس

§32.2. Listen to a modelling of these words featuring « ى » in Perso-Arabic vis-à-vis «ай», «и», «...й», «э», or «...е» in Cyrillic.

пеш	[before]	پیش	кй	[who(m)?]	کی
зебо	[beautiful]	زیبا	чй	[what?]	چی
тез	[quick(ly, speedy)]	تیز	миз	[table]	میز
чег	[scream, call]	جیغ	бист	[twenty]	بیست
шер	[lion]	شیر	шир	[milk]	شیر
ҳеч	[nothing, any(thing)]	هیچ	риш	[beard]	ریش
дер	[late]	دیر	сим	[wire, silver]	سیم
шева	[dialect]	شیوه			
реша	[root]	ریشه	дайр	[monastery]	دیر
Эрон	[Iran]	ایران	вай	[he/him, she/her]	وی
эчодкор	[creator]	ایجادکار	кай	[when?]	کی
эзад	[God]	ایزد	кайхо	[long ago]	کیها
эрод	[comment]	ایراد			

§32.3. Although **hâ-yi havvaz** often represents /â/ when written at the end of a word and the letter **vâv** sometimes represents /u/ when written in the middle or at the end of a word, words that contain one of the three sounds /a/, /u/, and /i/ appear in Perso-Arabic script as if the sounds did not exist in the word. For example, the word for "hello," «салом», is written سلام (or s+l+â+m); the word for "book," «китоб», is written کتاب (or k+t+â+b); and the word for "god," «худо», is written خدا (or kh+d+â). Pronunciation of such words require memorization in order to read Perso-Arabic texts.

Three symbols exist to indicate that an /a/, /u/ or /i/ sound follows a particular written consonant. These symbols are:

for /a/, a mark called **fathá** or **zabár**, written $\overset{_}{\text{ـ}}$
 for /u/, a mark called **zammá** or **pîsh**, written $\overset{\u}{\text{ـ}}$
 for /i/, a mark called **kasrá** or **zîr**, written $\overset{_}{\text{ـ}}$

A fourth symbol, called **sukûn** and consisting of a small circle written above a consonant letter, indicates that the consonant so marked is not followed by a vowel sound. Here follow examples of these symbols in a group of Arabic loanwords sharing the same trilateral consonant root system (س + ل + م).

سَلَام . اِسْلَام . مُسْلِم . تَسْلِيم . سَلَامَت . سَالِم .
 [hello] [Islam] [Muslim] [surrender] [health] [healthy]

In most cases, as long as the signs for /a/, /i/, and /u/ appear consistently, the **sukûn** sign is unnecessary. Of course, none of these signs ordinarily appears in Perso-Arabic writing; their use restricted to dictionaries and to situations where an uncommon word appears or some confusion about pronunciation of a word is anticipated. For students of Tâjîkî, they are useful in taking notes and recording pronunciation of what is heard and read.

If a word begins with one of the three sounds /a/, /u/, or /i/, it begins in writing with either an **'ayn** or an **alif**, both of which represent a glottal stop. In this situation, the appropriate short vowel sign can appear to indicate the vowel sound associated with the **'ayn** or **alif**. Examples involving an initial **alif** or **'ayn** (used only in those Arabic loanwords with their original spelling intact) include:

أَلْكُل - أَفْسَر - أَتْم - أَمَلِت - أُپْرَا - أُرِكْسْتِر - اِمْرُوَز - اِمپِرَاتُور - اِنِرژِي
 [alcohol] [officer] [atom] [omelette] [opera] [orchestra] [today-имрӯз] [emperor] [energy]
 عَرَب - عَزِيْز - عَذْر - عُمْر - عِبَارَه - عِبَارَت
 [Arab] [dear] [excuse, apology] [life(time)] [phrase] [consisting of]

The sign **fathá/zabár** can be used with **sukûn** to distinguish the sound /i/ as written (i.e., the letter **yâ** with no sign) from /ay/, also written with **ye**. For example:

كى [who?] . كِي [when?] . كِيهَا [for a long time] . كى

The sign **fathá/zabár** can be used with **sukûn** to distinguish the sound /û/ or /ü/ as written (i.e., the letter **vâv** with no sign) from the sound /av/. For example:

[mountain] кӯх	کوه	[far, distant] дур	دور
[how] чӣ тавр	چӣ طَوْر	[period of time] давра	دَوْرَه

§33. For practice reading Perso-Arabic transcriptions of words with short vowel sounds, turn to Section §4.11 (above), which presents Tâjîkî words and phrases in Cyrillic and fully vowel-pointed Perso-Arabic scripts along with English equivalents.

§34. Besides thirty-two alphabet letters, three other important written symbols appear in Perso-Arabic writing, called **tashdîd**, **tanvîn**, and **hamzá**.

§34.1. The symbol **tashdîd** (تَشْدِيد) or **shadda** is a small "w"-shaped character placed over a consonant letter signifying the doubling of that consonant. Doubling a consonant means doubling the length of time in pronouncing a consonant, as would occur if the English word "bookcase" were pronounced with both "k" and "c" clearly enunciated. Although **tashdîd** is often omitted in writing, words which have it are still pronounced with the doubled consonant. Otherwise, what is said may be hard to understand or may be a different word. The vast majority of instances of **tashdîd** occur in words of Arabic origin. Here are some common examples:

муҳим(м)	[important] /muhîmm/	مُهْمٌ	аммо	[but] /ámmâ/	أَمَّا
тараққиёт	[progress] /taraqîyât/	تَرْقِيَّاتٌ	албатта	[certainly] /albátta/	أَلْبَتَّةَ
шахсият	[personality] /shakhsîyyát/	شَخْصِيَّاتٌ	ҳаҷ(ҷ)	[hajj] /hajj/	حَجٌّ
муддат	[period of time] /muddát/	مُدَّتٌ	қад(д)	[height] /qadd/	قَدٌّ
ҳад(д)	[limit] /hadd/	حَدٌّ	маҳалла	[neighborhood] /mahallá/	مَحَلَّةَ
муаллим	[teacher] /mu'allîm/	مُعَلِّمٌ	ҳуҷҷат	[document, reason] /hujjat/	حُجَّتٌ

Appearing mostly in words of Arabic origin, **tashdid** often signals specific Arabic loanword patterns, as in these examples of common patterns.

муҷаррад	[unmarried]	مُجَرَّدٌ	муаллим	[teacher]	مُعَلِّمٌ
мураббаъ	[square(d)]	مُرَبَّعٌ	муҳаррир	[editor]	مُحَرَّرٌ
мусаллаҳ	[armed]	مُسَلَّحٌ	муаллиф	[writer]	مُؤَلِّفٌ
муваққат	[temporary]	مُؤَقَّتٌ	мусаввир	[illustrator]	مُصَوِّرٌ
мутахассис	[expert]	مُتَخَمِّصٌ	таваҷҷӯх	[attention]	تَوَجُّهُ
мутафаккир	[thinker]	مُتَفَكِّرٌ	таваллуд	[birth]	تَوَلَّدَ
муташаккил	[organized]	مُتَشَكِّلٌ	танаффус	[break from work]	تَنَفُّسٌ
мутакаббир	[proud]	مُتَكَبِّرٌ	тавассут	[intermediation]	تَوَسُّطٌ

ИТТИФОҚ	[union, alliance]	اتِّفَاقٌ	ташаккур	[thanks, thanking]	تَشَكُّرٌ
ИТТИХОД	[union]	اتِّحَادٌ	тасаввур	[imagination/ing]	تَصَوُّرٌ
ИДДАО	[claim, pretension]	ادِّعَا	таҷассум	[embodiment]	تَجَسُّمٌ
ИТТИЛООТ	[information (pl)]	اطِّلَاعَاتٌ	тааччуб	[surprise]	تَعْجَبٌ
			тааллуқ	[connection, relation]	تَعَلُّقٌ

§34.2. The symbol **tanvîn** (تَنْوِين) consists of two **zabâr** marks written in conjunction with an **alif**. **Tanvîn** (أ) is pronounced /ân/, and serves as a suffix indicating that a word is an adverb. For example:

баъдан	[subsequently, afterwards]	بَعْدًا	масалан	[for example]	مَثَلًا
тақрибан	[approximately]	تَقْرِيْبًا	қаблан	[previously]	قَبْلًا
ҳатман	[certainly]	حَتْمًا	ақсаран	[for the most part]	أَكْثَرًا
нисбатан	[relatively]	نِسْبَتًا	одатан	[customarily]	عَادَتًا
амалан	[practically]	عَمَلًا	мазмунан	[in content]	مَضْمُونًا
аслан	[at all, originally]	أَصْلًا	расман	[formally]	رَسْمًا
тамоман	[wholly]	تَمَامًا			

§34.3. **Hamzá** (هَمْزَة) is a written symbol used over medial or final **alif** (أ), **vâv** (و), or the seat of **yâ** (ي) without that letter's two dots. **Hamzá** represents a glottal stop or pause between the sounds which precede and follow it, just as the letter '**ayn** (ع) represents it in such words as شَيْعَة /shî'â/ [Shi'ite Muslim]. In word-initial position, the letter **alif** fulfills the function of **hamzá** as a glottal stop in words whose first sound is a vowel, such as ايران, pronounced /'e(y)rân/, or أْتُمْ, pronounced /'atúm/. Here follow common examples of **hamzá**:

масъала	[problem, issue] /mas'ala/	مَسْأَلَةٌ	масъул	[responsible] /mas'ûl/	مَسْئُولٌ
раис	[president] /ra'is/	رَئِيسٌ	таъкид	[emphasis] /ta'kid/	تَأْكِيدٌ
муассиса	[institution] /mu'assisa/	مَوْسِسَةٌ	оид ба	[concerning] /â'id ba/	عَائِدٌ بِهِ
маъмурият	/ma'mûriyyat/	مَأْمُورِيَّةٌ	мутаассифона	[unfortunately] /muta'assifâná/	مُتَأَسِّفَانَةٌ
	[administration]				

§34.3.1. **Hamza** and '**ayn** letters disappear in the Cyrillic transcription of many words. This means that, as with the lack of indication in Cyrillic transcription of different multiple letters in Perso-Arabic for «с», «х», «т», and «з», Tâjiks unfamiliar with Perso-Arabic script cannot immediately appreciate connections among originally related Arabic loanwords or distinguish between those words and originally Persian words. Here follow examples of common words with **hamza** and '**ayn** in Perso-Arabic not reflected in Cyrillic.

кучой	[from where]	كُجَائِي	раис	[president]	رئيس
амрикой	[American]	أَمْرِيكَائِي	ибтидой	[elementary]	إِبْتِدَائِي
шиносой	[acquaintance]	شِنَاسَائِي	муассиса	[institution] /mu'assisa/	مُؤَسَّسَة
соат	[hour, clock/watch]	سَاعَة	оқибат	[end, result, in the end]	عَاقِبَة
саноат	[industry]	صِنَاعَة	одат	[habit, custom]	عَادَة
зироат	[agriculture]	زِرَاعَة	оила	[family]	عَائِلَة
табиат	[nature]	طَبِيعَة	оид ба	[concerning]	عَائِدَ بِهِ
муаллим	[teacher]	مُعَلِّم	воқеа	[event, accident]	وَاقِعَة
муошират	[social intercourse]	مُعَاشِرَة	ҷомеа	[society]	جَامِعَة
мурочиат	[address]	مُرَاجَعَة	маҷмӯа	[collection, miscellany]	مَجْمُوعَة
мутолиа	[study/ing]	مُطَالَعَة	монеа	[barrier, obstacle]	مَانِعَة
маориф	[education]	مَعَارِف			

иттилоъ - иттилоот	[information (s/pl)]	اطَّلَاع - اَطَّلَاعَات
мавзӯъ - мавзӯи муҳити зист	[subject - environment]	مَوْضُوع - مَوْضُوعٌ مُّحِيطٌ زَيْسْت
авзоъ - авзои кор	[mood, situation - state of affairs]	أَوْضَاع - أَوْضَاعٌ كَار
- авзоаш бечо	[He/she is in a bad mood/low spirits.]	- أَوْضَاعُش بِيَجَا .
мавқеъ - мавқеи муҳим	[situation, moment - important moment]	مَوْقِع - مَوْقِعٌ مُّهِم
нафъ - ба нафи давлат	[in the interests of the state]	نَفْع - بِهِ نَفْعٌ دَوْلَت