This manual describes and illustrates the writing system of modern Persian. It is modeled after Frank A. Rice's "Classical Arabic: The Writing System." Five sections concerning the writing system are provided: (1) printed Persian, (2) typewritten Persian, (3) handwritten Persian, (4) the combination of words, and (5) punctuation. (RL)
THE WRITING SYSTEM
OF
MODERN PERSIAN

Herbert H. Paper
and
Mohammad Ali Jazayery

Spoken Language Services, Inc.
THE WRITING SYSTEM
OF MODERN PERSIAN

Herbert H. Paper
and
Mohammad Ali Jazayery
The research and compilation of which this work is a result were brought to completion under a subvention from the Board on Overseas Training and Research (Ford Foundation)
This manual, originally prepared by Professor Herbert H. Paper, of the University of Michigan, is closely modelled on Frank A. Rice's *Classical Arabic: The Writing System* (Foreign Service Institute, 1952). It has been revised to some extent by the following persons: Charles A. Ferguson, Carleton T. Hodge, Mohammad Mehdandi-Nejad, and Mohammad Ali Jazayery. The notes on typewritten Persian, handwritten Persian, the combination of words, and punctuation are Mr. Jazayery's.
# TABLE OF CONTENTS

## PRINTED PERSIAN

<table>
<thead>
<tr>
<th>Page Range</th>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>General</td>
<td>1</td>
</tr>
<tr>
<td>1-2</td>
<td>The Alphabet</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Vowels</td>
<td>3</td>
</tr>
<tr>
<td>4-16</td>
<td>Non-Connectors</td>
<td>4</td>
</tr>
<tr>
<td>5-16</td>
<td>Connectors</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>Ligatures</td>
<td>16</td>
</tr>
<tr>
<td>18-26</td>
<td>Special Signs</td>
<td>16</td>
</tr>
<tr>
<td>27</td>
<td>Numbers</td>
<td>19</td>
</tr>
</tbody>
</table>

## TYPEWRITTEN PERSIAN

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-21</td>
</tr>
</tbody>
</table>

## HANDWRITTEN PERSIAN

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>22-24</td>
</tr>
</tbody>
</table>

## COMBINATION OF WORDS

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-26</td>
</tr>
</tbody>
</table>

## PUNCTUATION

<table>
<thead>
<tr>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>27-30</td>
</tr>
</tbody>
</table>
PRINTED PERSIAN

GENERAL

0. Persian is customarily written and printed in an adapted form of the Arabic alphabet consisting of 32 letters and a number of signs. It differs from its Arabic original in that it has added four new letters to take care of Persian sounds which do not exist in Arabic. In general, all words of Arabic origin, as well as a number of Arabic phrases common in Persian, appear in their original Arabic orthography.

While the letters of the alphabet are basically the same in form whether written, printed, or typed, there are slight differences. For our purposes, three main types may be distinguished: printed, typed, and handwritten (which is sometimes, though not so frequently as in earlier times, used in lithography or photo-printing). In the present treatment we are mainly concerned with the printed variety of the Persian alphabet. However, a short note on handwritten Persian has been added to serve as an introductory guide to the reading of this style. The typewritten form is also briefly discussed.

1. Persian is written and printed from right to left.
2. All of the consonants, but only three of the six simple vowels, are indicated.
3. The manner of writing and printing is cursive, and most of the letters are connected to preceding and following letters within the same word.
4. There are no capital letters.
5. Most of the letters have four variant shapes, depending on whether or not they are connected to preceding and/or following letters. Later on the four shapes will be called initial, medial, final, and independent.
6. Some letters are identical in basic form and are distinguished from one another by the arrangement of one, two, or three dots placed above, below, or inside the basic form.

THE ALPHABET

1. The following list shows the letters of the Persian alphabet in the shape they have when they stand alone (= independent), that is, not connected to a preceding or following letter. This is the conventional order of the alphabet according to which the dictionaries are arranged.
<table>
<thead>
<tr>
<th>Letter</th>
<th>The Persian Alphabet</th>
<th>Transcribed by</th>
</tr>
</thead>
<tbody>
<tr>
<td>ا</td>
<td>'ælef</td>
<td>'</td>
</tr>
<tr>
<td>ب</td>
<td>be</td>
<td>b</td>
</tr>
<tr>
<td>س</td>
<td>se</td>
<td>s</td>
</tr>
<tr>
<td>د</td>
<td>dal</td>
<td>d</td>
</tr>
<tr>
<td>ز</td>
<td>zal</td>
<td>z</td>
</tr>
<tr>
<td>ر</td>
<td>re</td>
<td>r</td>
</tr>
<tr>
<td>ز</td>
<td>ze</td>
<td>z</td>
</tr>
<tr>
<td>ژ</td>
<td>ţe</td>
<td>ţ</td>
</tr>
<tr>
<td>س</td>
<td>sin</td>
<td>s</td>
</tr>
<tr>
<td>ش</td>
<td>šin</td>
<td>š</td>
</tr>
<tr>
<td>ص</td>
<td>sad</td>
<td>s</td>
</tr>
<tr>
<td>ض</td>
<td>zad</td>
<td>z</td>
</tr>
<tr>
<td>ت</td>
<td>ta</td>
<td>t</td>
</tr>
<tr>
<td>ص</td>
<td>za</td>
<td>z</td>
</tr>
<tr>
<td>ین</td>
<td>'ynn</td>
<td>ţ (or no symbol)</td>
</tr>
<tr>
<td>ق</td>
<td>qeyn</td>
<td>q</td>
</tr>
<tr>
<td>ش</td>
<td>še</td>
<td>ţ</td>
</tr>
<tr>
<td>ق</td>
<td>qaf</td>
<td>q</td>
</tr>
<tr>
<td>خ</td>
<td>kaf</td>
<td>k</td>
</tr>
<tr>
<td>گ</td>
<td>gaf</td>
<td>g</td>
</tr>
<tr>
<td>ل</td>
<td>lam</td>
<td>l</td>
</tr>
<tr>
<td>م</td>
<td>mim</td>
<td>m</td>
</tr>
<tr>
<td>ن</td>
<td>nun</td>
<td>n</td>
</tr>
<tr>
<td>و</td>
<td>vav</td>
<td>v, u, o, ow</td>
</tr>
<tr>
<td>ی</td>
<td>he (hævvaéz)</td>
<td>h, e</td>
</tr>
<tr>
<td>ی</td>
<td>ye</td>
<td>y, i</td>
</tr>
</tbody>
</table>
Notice that in a number of instances several letters represent the same sound in Persian. In these cases, only one of these letters ordinarily occurs in words of Persian origin, while all of them can be found in Arabic loanwords. However, in a few cases, even words which are not Arabic borrowings are spelled with letters otherwise peculiar to Arabic. A number of Persian proper nouns (both place names and personal names) are spelled with typical Arabic letters by some people and with the ordinary letters by others. The present tendency is in favor of the latter type of spelling. The same situation exists in the case of words borrowed from languages other than Arabic (especially Greek and Turkish).

<table>
<thead>
<tr>
<th>Persian Sound</th>
<th>Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>z</td>
<td>ze</td>
</tr>
<tr>
<td></td>
<td>zal</td>
</tr>
<tr>
<td></td>
<td>zad</td>
</tr>
<tr>
<td></td>
<td>Chiefly in Arabic loan words</td>
</tr>
<tr>
<td>s</td>
<td>sin</td>
</tr>
<tr>
<td></td>
<td>se</td>
</tr>
<tr>
<td></td>
<td>Chiefly in Arabic loan words</td>
</tr>
<tr>
<td>t</td>
<td>te</td>
</tr>
<tr>
<td></td>
<td>ta</td>
</tr>
<tr>
<td></td>
<td>Chiefly in Arabic loan words</td>
</tr>
<tr>
<td>h</td>
<td>he</td>
</tr>
<tr>
<td></td>
<td>høttë</td>
</tr>
<tr>
<td></td>
<td>Chiefly in Arabic loan words</td>
</tr>
<tr>
<td>q</td>
<td>qeyn</td>
</tr>
<tr>
<td></td>
<td>qaf</td>
</tr>
<tr>
<td></td>
<td>Chiefly in Arabic loan words</td>
</tr>
<tr>
<td></td>
<td>'ælëf</td>
</tr>
<tr>
<td></td>
<td>'eyn</td>
</tr>
<tr>
<td></td>
<td>Chiefly in Arabic loan words</td>
</tr>
</tbody>
</table>

2. The following shows the Persian letters in their alphabetical order and written from right to left. Note the placing of the letters relative to the line of writing.

اببتتختیچخذدرز
سشستظعغفقبکل
منوید

VOWELS

3. Persian has six vowels: a, æ, e, i, o, u. Two of them, i and u, are regularly indicated in the writing by the letters ی (ye) and و (vav) re-
spectively. One, a, is regularly indicated by ֝ (‘a’eléf) in the middle or at the end of a word. At the beginning it is sometimes indicated by a special mark ֝ over ֝ (‘a’eléf): ֝. The other three vowels are not usually represented in the writing. Four special signs, described in Section 18, may be used to indicate ֜, ֝, and ֤, or absence of vowel, but these signs are rarely used in ordinary written or printed Persian. Occasionally ֤ is indicated by י (vav); and at the end of a word ֝ is sometimes indicated by א (he hœARRAY).

There is also a special sign to indicate double consonants (see Section 23). This sign is not ordinarily used.

NON-CONNECTORS

4. The following letters are non-connectors. They do not have a curvise linking to a following letter, though they do connect to a preceding letter if the preceding letter is a connector.

֝ 'a’eléf is characteristically a vertical stroke.
֝ dal
֝ zal is exactly like dal, but with one dot above it.
֝ re
֝ ze is exactly like re, but with one dot above it.
֝ że is exactly like re, but with a dot triangle above it.
֝ vav

At the beginning of a word, the letter ֝ (‘a’eléf) indicates ֝. A vowel naturally follows, as Persian has no words beginning with two consonants. ֝ is represented by ֝, ֝ by ֝. ֝ has a special symbol over the ֝ (‘a’eléf): ֝ (see Section 19.10). ֝, ֝, and ֝ are ordinarily written with ֝ (‘a’eléf) alone.

In the middle of a word, ֝ (‘a’eléf) most often represents the vowel ֝, but it may also represent ֝. At the end of a word, ֝ (‘a’eléf) nearly always indicates the vowel ֝ (see Section 24).

The symbol ֝ (Section 19.11) for ֝ is also used in the middle and at the end of words.

The letter י (vav) at the beginning of a word represents the consonant ֜: elsewhere it may represent ֝, ֝, ֝, or ֝.

罅 ‘æz ‘from’ ֝ do ‘in’ ֝ dad ‘he gave’
א ‘u ‘he’ ֝ do ‘two’ ֝ 드 ‘dard ‘pain’
א ‘a’eléf ‘value’ ֝ dez ‘fortress’ ֝ do ‘a’ ֝ da ‘medicine’
֝ ד ‘dud ‘smoke’
CONNECTORS

5. All other letters connect both to preceding letters (except when they are non-connectors) and to following letters. These may be arranged according to their basic forms.

6. b, n, y, t, p, s. These letters are all characterized by an upward stroke with one to three dots above or below. The medial forms are as follows:

   \[
   \begin{array}{ccc}
   b & y & p \\
   n & t & s
   \end{array}
   \]

   In their final and independent forms, there is significant variation only in the basic forms of b, n and y.

6.10 b (be) has a single dot below.

6.11 b- has the initial shape:  

   باز 'open'  
   بيد 'bad'  

6.12 -b- has the medial shape:  

   ببر 'tiger'  
   بتب 'axe'  

6.13 -b has the final shape:  

   زيب 'ornament'  
   تيب 'fever'  

6.14 b has the independent shape:  

   تاب 'glow'  
   ناب 'pure'  

6.20 n (nun) has one dot above.

6.21 n- has the initial shape: ;
6.22 -n- has the medial shape:  
bnayd 'foundation'
panfr 'cheese'

6.23 -n- has the final shape: ن
ن تن 'body'
 تن 'root'

6.24 n has the independent shape: ن
نان 'bread'
نادان 'ignorant'

6.25 Written medial or final nb is pronounced mb:  
زاً 'wasp'

6.30 y (ye) has two dots below (when initial or medial). The letter ی (ye) at the beginning of a word represents the consonant y; elsewhere it may represent i or y. The sequence ای (‘a’leml-ye) at the beginning of a word often represents ‘i, and sometimes it represents ‘ey.

6.31 y- has the initial shape: ی

6.32 -y- has the medial shape: ی
نیز 'also'
پنفر 'cheese'

6.33 Final ye ی is deeper than the final forms of the other letters treated in Section 6 (except n). It is also characterized by the curving connecting line. Notice that the two dots are ordinarily not written below the final and independent forms.

6.34 y has the independent shape: ی
روزی 'a day'
پاری 'play, game'

6.35 At the end of certain Arabic loan words the letter ی represents the sound a. You will have to learn the words that are written with ی for a instead of ی (‘a’leml). (See Section 26).
6.40 t (te) has two dots above.

6.41 t- has the initial shape: 

- tabot 'coffin'
- tebár 'axe'

6.42 -t has the medial shape: 

- pačo 'blanket'
- doktor 'doctor'

6.43 -t has the final shape: 

- tæbýet 'education'
- poast 'skin'

6.44 t has the independent shape: 

- tabot 'coffin'
- ahsabat 'feelings'

6.50 p (pe) has three dots arranged in a triangle below.

6.51 p- has the initial shape: 

- bap 'foot'
- por 'full'

6.52 -p- has the medial shape: 

- tepid 'palpitated'
- bebæer 'fly'

6.53 -p has the final shape: 

- pip 'tobacco pipe'
- tip 'brigade'

6.54 p has the independent shape: 

- bab 'Pope'
- tup 'cannon'

6.60 s (se) has three dots above.

6.61 s- has the initial shape: 

- sabet 'proved; fixed'
- saerbet 'wealth'

6.62 -s- has the medial shape: 

- nesr 'prose'
- mesal 'example'
6.63 -s has the final shape: ت
  'sbees 'futile'
  moshellés 'triangular'

6.64 -s has the independent shape: ت
  varés 'heir'
  vorrás 'heirs'

7. j, c, x, h. These letters, like those of Section 6, have the same basic form but differ only in the arrangement of dots.

7.10 j (jim) is characterized by a dot placed below the basic form.

7.11 j- has the initial shape: ج
  ja 'place'
  joz 'except'

7.12 j- has the medial shape: ج
  bijan 'lifeless'
  raenjidaan 'to be offended'

7.13 j has the final shape: ج
  pøenj 'five'
  berénj 'rice'

7.14 j has the independent shape: ج
  taj 'crown'
  borj 'tower'

7.20 x (xe) is characterized by a dot placed above the basic form.

7.21 x- has the initial shape: خ
  xar 'thorn'
  xøer 'donkey'

7.22 -x- has the medial shape: خ
  bæxt 'luck'
  poxtén 'to cook'

7.23 -x has the final shape: خ
  bix 'root'
  yøex 'ice'

7.24 x has the independent shape: خ
  xøex 'thread'
  zaøæex 'chin' (literary)
The sequence xa is in certain words written خا . The (vav) here is an historical writing and has no sound value.

Note that the letter خ (vav) after a خ (xe) sometimes represents the vowel o.

-c (ce) is characterized by three dots placed below the basic form.

-c has the initial shape: جاب  
cap 'printing'

cap 'left'

-c has the medial shape: بيجيدن  
picidæn 'to twist'

-c has the final shape: بيج  
pic 'twist, screw'

-c has the independent shape: بوج  
puc 'vain, futile'

-h (he hotif) has the same shape in all positions as j, c, and x except that it has no dots of any kind and is simply the basic form of this series.

-h has the initial shape: ح  
hejáz 'Hejaz'

hez 'political party'

heyz 'animal'

h has the independent shape: ح  
nuh 'Noah'

has the initial shape: ح  
heb 'explanation'
8.11 -s has the medial shape: نیست ‘is not’

8.12 -s has the final shape: بست ‘twenty’

8.13 -s has the independent shape: نیست ‘is not’

8.20 -ة has the initial shape: شیش ‘perhaps’

8.21 -ة has the medial shape: بشت ‘back’

8.22 -ة has the final shape: خوش ‘self’

8.23 -ة has the independent shape: خوش ‘happy’

9. ꞌk and ꞌg (kaf and gaf) have the same basic form. Kaf has one oblique stroke in its initial and medial forms.

9.10 ꞌk has the initial shape: كاب ‘book’

9.11 ꞌk has the medial shape: شکار ‘hunting’

9.12 ꞌk has the final shape: یک ‘one’
9.13 \( k \) has the independent shape: ك

pak 'pure'

ğak 'dust'

9.20 gaf differs from kaf only in having two parallel oblique strokes in all its forms.

9.21 g- has the initial shape: 

kerek 'wolf'

gozesten 'to pass'

9.22 -g- has the medial shape: 

jegar 'liver'

tægærg 'hail'

9.23 -g has the final shape: 

rig 'sand'

jæng 'war'

9.24 g has the independent shape: 

kerek 'wolf'

tægærg 'hail'

10. f and q (fe and qaf) differ only in the number of dots and in their final and independent forms, where qaf has a deeper loop than fe; the final form of fe is hollow, while that of qaf is not.

10.10 f- has the initial shape: 

fezand 'child'

ostasen 'to fall'

10.11 -f- has the medial shape: 

seffd 'white'

jost 'pair'

10.12 -f has the final shape: 

sherif 'noble'

keef 'discovery'

10.13 f has the independent shape: 

bærf 'snow'

bærf 'letter of the alphabet'

10.20 qaf differs from fe in having two dots placed above instead of one.

10.21 q- has the initial shape: 

tork 'Turk'

teryak 'opium'

tork 'Turk'

tægærg 'hail'

jæng 'bell'

bozorg 'big'

baerg 'leaf'

fekr 'thought'

'afser 'officer'

næft 'oil'

tofseeng 'rifle'

kæsf 'dirty'

kif 'purse, bag'

naf 'navel'

SAE
10.22 -q- has the medial shape: ٣

نار qaebr 'grave'

قزر qern 'century'

10.23 -q- has the final shape: ق

تاق saeqqa 'water-carrier'

منقار menqar 'beak'

10.24 q has the independent shape: ق

بته saeqq 'east'

اتناث 'etefaq 'incident''

II. 1 (lam) is characteristically a vertical stroke. It differs from ِ (‘eelaf) chiefly by connecting to a following letter.

11.10 1- has the initial shape: ل

لین ländsen 'London'

لخت loxt 'naked'

لاباس lebás 'clothes'

11.11 -ل- has the medial shape: ل

جلد jeld 'skin, hide'

جلو jelów 'front'

تلمع talex 'bitter'

كلر gaelú 'throat'

11.12 -ل- has the final shape: ل

كل gol 'rose'

بل pol 'bridge'

نيل nil 'Nile'

بل ball 'nightingale'

11.13 ل has the independent shape: ل

دل del 'heart'

بول pul 'money'

دببال dombál 'rear'

12. m (mim) is characteristically a small circle, often merely a dot on a downward break in the line.

12.10 m- has the initial shape: م

ما ma 'me'

مادر madar 'mother'

مئز miz 'table'

12.11 -م- has the medial shape: م

محمد mohammad 'Mohammed'

مبار bimár 'sick'

ست saemt 'direction'

شمي šimi 'chemistry'
12.12 -m has the final shape: م
nim ‘half’
ketābem ‘my book’
SIM ‘wire’
raštām ‘I went’
12.13 -m has the independent shape: م
teemām ‘whole’
salām ‘hello’
sevvom ‘third’
merdōm ‘people’
13. h (he ḥayyēz) does not have any one characteristic form.
13.10 h- has the initial shape: ه
hēr ‘each, every’
hēt ‘seven’
8 ‘eight’
deḥān ‘mouth’
13.11 h- has the medial shape: ه
behtār ‘better’
baḥār ‘spring’
beḥēt ‘paradise’
cehēl ‘forty’
13.12 h has the final shape: ه. The final h (ه) often indicates the vowel e.
se ‘three’
rafeh ‘gone’
7 rhobe ‘radish’
texē ‘board, plank’
13.13 h has the independent shape: ه. The independent h (ه), too, often indicates the vowel e.
rah ‘path, road’
ejazē ‘permission’
piadē ‘on foot’
deh ‘village’
14. s and z (sad and zad) have the same basic form. The sad is dot-less; the zad has one dot above.
14.10 s- has the initial shape: س
sabūn ‘soap’
sēhīn ‘correct’
sobh ‘morning’
sed ‘hundred’
14.11 -s- has the medial shape: س
teṣsim ‘decision’
bāsrē ‘Basra’
sāmēr ‘Egypt’
sīsād ‘three hundred’
14.12 -s has the final shape: س
neqs ‘defect’
naqēs ‘defective’
nāfs ‘person’
morēxxēs ‘permitted to leave’
14.13 s has the independent shape: س
14.20 -z- has the initial shape: صضر امة 'loss'

14.21 -z- has the medial shape: a هضم qezavet 'judgment'

14.22 -z- has the final shape: ضنيف نعئز 'violation'

14.23 -z- has the independent shape: ضطاش رياض 'Riyadh'

15. t and z (ta and za) have the same basic form. The ta is dotless; the za has one dot above.

15.10 t- has the initial shape: طلاطب tseb 'physician'

15.11 -t- has the medial shape: طسل سمل sell 'pail'

15.12 -t- has the final shape: طربط رابط xett 'line'

15.13 -t- has the independent shape: طامح tseebiyat 'caution'

15.20 -z- has the initial shape: طظر zahr 'vase'

15.21 -z- has the medial shape: طناظار نظر نظر nezarat 'supervision'

15.10 t- has the initial shape: ططل ططط telē 'gold'

15.11 -t- has the medial shape: طتعل نمط met 'meter'

15.12 -t- has the final shape: طربط رابط xett 'line'

15.13 -t- has the independent shape: طامح tseebiyat 'caution'

15.20 -z- has the initial shape: طظر zahr 'vase'

15.21 -z- has the medial shape: طناظار نظر نظر nezarat 'supervision'

xolús 'sincerity'

xass 'particular'

ziyafet 'banquet'

vazéh 'clear, plain'

enqeza 'expiration'

mozér 'harmful'

mehze 'only for'

qezr 'debt'

maeqrus 'indebted'
15.22 -z has the final shape: 
\[ \text{خذأ حافظ} \]
\[ \text{xodahařez 'Good-bye!} \]
\[ \text{حرف 'protection'} \]
\[ \text{hefż 'protection'} \]

15.23 z has the independent shape: 
\[ \text{حظي} \]
\[ \text{tælæfiz 'pronunciation'} \]
\[ \text{hafız 'Hafizz'} \]

The following basic shape ('eyn) represents (that is, either a glottal stop or stricture, or a lengthening of the preceding vowel). Initially it functions much like 'eléf, but occurs only in words of Arabic origin. 'eyn and qeyn are alike in their basic form, but qeyn has one dot above.

16.10 - has the initial shape: 
\[ \text{غاءت} \]
\[ \text{'aqel 'wise'} \]
\[ \text{'ebärät 'phrase'} \]

16.11 - has the medial shape: 
\[ \text{سعت} \]
\[ \text{'aejfb 'strange'} \]
\[ \text{'ajmat 'social'} \]

16.12 - has the final shape: 
\[ \text{'ابع} \]
\[ \text{'ayy 'attempt'} \]
\[ \text{'efel 'verb'} \]

16.13 has the independent shape: 
\[ \text{مصارع} \]
\[ \text{mozäré 'present (tense)' \}
\[ \text{دايغ} \]
\[ \text{defi 'defense'} \]
\[ \text{اجتمااع 'ejtemaţ 'gathering; society'} \]

16.20 q- has the initial shape: 
\[ \text{غاز} \]
\[ \text{qaz 'goose'} \]
\[ \text{غذا} \]
\[ \text{qaşı 'food'} \]

16.21 -q- has the medial shape: 
\[ \text{منتر} \]
\[ \text{maq 'brain, marrow'} \]
\[ \text{بئل} \]
\[ \text{bæq'el 'armpit'} \]

16.22 -q has the final shape: 
\[ \text{بغي} \]
\[ \text{miq 'cloud, mist'} \]
\[ \text{مشغول} \]
\[ \text{maeq'ul 'busy'} \]

16.23 q has the independent shape: 
\[ \text{داوغ} \]
\[ \text{dorůq 'lie, false'} \]
\[ \text{داوغ} \]
\[ \text{dæmq 'nose'} \]
LIGATURES

17. Ligatures are a combination of two letters into a specialized form which is not a part of the regular Persian alphabet. They are limited to the following combinations: lam-‘æl̄ef, kaf-‘æl̄ef, gaf-‘æl̄ef, kaf-lám, and gaf-lám. When the sequences kaf-lam-‘æl̄ef and gaf-lam-‘æl̄ef appear, only two of the letters may be combined into a ligature. The remaining letter is printed in its usual form.

17.10 The combination lam-‘æl̄ef, when in isolated position, is printed:

لا
lastik ‘rubber’
laqṣān ‘vulture’
lak ‘sealing wax’
لأ ‘owlad ‘child’

17.11 When in final position, the combination lam-‘æl̄ef has this shape:

ال
خلال ‘interval’
( literary)
خلال ‘agriculture’
فلحه ‘difference’
طلا ‘gold’

17.20 The combinations kaf-‘æl̄ef and gaf-‘æl̄ef, when in isolated position, are printed ك and ك. When in final position, they are printed and ك respectively.

كما
kaqæz ‘paper’
gam ‘step’
‘enkar ‘denial’
‘nak ‘glance’

كما
kæl̄em ‘cabbage’
‘ælkol ‘alcoholic’
‘ængol ‘parasite’
‘ænngol ‘English’

كما
šekl ‘shape’
šekl ‘shape’
šekl ‘shape’
šekl ‘shape’
tawwakkol ‘trust (e.g., in God)’

gol ‘flower’

17.30 The combinations kaf-lám and gaf-lám are written ك and ك initially, ك and ك medially, ك and ك finally, and ك and ك when independent.

كما
kæl̄em ‘cabbage’
šekl ‘shape’
šekl ‘shape’
tawwakkol ‘trust (e.g., in God)’

gol ‘flower’

17.40 The combinations kaf-lam-‘æl̄ef and gaf-lam-‘æl̄ef are more commonly printed ك and ك. Kama, kama, kama, kama, kama, kama, kama, kama, kama, kama.

كما
kæl̄eq ‘crow’
šekl ‘he’
šekl ‘he’
šekl ‘he’
tawwakkol ‘trust (e.g., in God)’

bekolabł ‘to the pear’

SPECIAL SIGNS

In Sections 3 and 4 some of the means for indicating vowels and diphthongs were mentioned. In the following nine sections, examples and combinations will be discussed.
18. The vowels ə, e, and o may be represented by the following signs:

ae (placed above the preceding consonant):

ئر 'soak'

e (placed under the preceding consonant):

کرم 'worm'

o (placed above the preceding consonant):

پر 'full'

Absence of vowel may be indicated by the sign ' placed above the consonant to which it belongs:

پدر 'father'

These four signs are, however, hardly ever used in printed (or written) Persian. These signs are sometimes called 'vowel points'.

19. 'ælef mæddi. The formـ, called 'ælef mæddi' (essentially an 'ælef drawn horizontally above a conventional 'ælef) expresses the sound a when in independent or final position. In alphabetical listing, 'ælef mæddi' counts as 'ælef plus 'ælef.'

19.10 'ælef mæddi in independent position:

آب 'water'

آرد 'flour'

آزاد 'free'

آسيا 'Asia'

19.11 'ælef mæddi in final position:

آثر 'memorable deeds' (literary)

آتار 'theater'

20. Initial 'ælef plus vav expresses 'u, 'ow, or 'v: أورنک

owræng 'throne' (literary)

اولاد 'children'

اول 'he, she'

21. Initial 'ælef plus ye expresses 'i, 'y, or (with a vowel, of course) 'ey or 'ey: 

آیران 'Iran'

ایران 'Iran'

آیوان 'palace'

ایشان 'they'

22. hæmzé. The symbol ' , called hæmzé, is sometimes used over ی, ی, and ə, as described in the following sections:

22.10 A glottal stop in the middle of a word is sometimes indicated by the small sign ' (hæmzé) above ی, ی, ə (without dots), and ə.

 Developers

 sa'el 'beggar'

روپسأ 'chiefs'

mæstúl 'responsible'

امیرکاک 'American'
There is a tendency to omit the ْهَمْزَة and use a ی (ye) instead everywhere except in words of Arabic origin.

22.20 ْسِلْف ْهَمْزَة. There are also cases where, in the middle of a word, ْهَمْزَة is written above an ْسِلْف (though its omission in print is frequent) to express the sequence ُّئ. Note that ْسِلْف here does not express the sound ُّأ.

---

خانِی بزرگ ْبَنَی 'the big house'
دَسْتِه کاکُذ ْبَنَی 'the bunch of paper'

Final connected ی may have ْهَمْزَة to represent indefinite -ی: ی 'ی. This is increasingly being written ْی 'ی.

خانِی ْبَنِی 'a house'
دَسْتِه ْبَنِی 'a bunch'

23. ْتَسْدِید. The doubling of consonants, in the middle or at the end of a word, may be indicated by the small sign ّ (ْتَسْدِید) written above the letter to be doubled; but it, too, is often omitted.

ْمَغْلِب ْبَنِی 'first'
ْمَخْفَف مَخْفَفی 'contraction'
ْمَلْم مَلْمی 'teacher'
ْلَفْطْ لَفْطی 'pronunciation'

24. ْتَنْفِن. An ending -ن (of Arabic origin) of certain adverbs is indicated by the following sign ّ (ْتَنْفِن), written usually above a final ْسِلْف. Like ْهَمْزَة and ْتَسْدِید, it, too, is often omitted. Note that in this case, final ْسِلْف does not have the value ُّأ.

ْمَعْلَم 'for example'
ْنَول 'firstly'
ْتَقْرِیب 'almost, approximately'
ْبَل 'beforehand, earlier, before'

25. In older printed material, the final and independent ی, ی and ی, appear, at the end of some words, with two dots above them. These words are invariably of Arabic origin. In modern writings these forms are very rarely found. In words where the dotted ی is pronounced ی, the symbol ی is used.

ْسِلاک 'genuineness'
ْتُاب 'pilgrimage'

In words where an originally dotted ی is retained without dots, as ی, it represents the vowel ی:

ْتُاب 'pilgrimage'

26. Mention was made in Section 6.35 of the fact that final ی is sometimes pronounced as if it were final ْسِلْف, and is pronounced ُّأ.
However, when these words are followed by the syllable -ye (another form of -e, the 'ezafé sign), the letter ye is replaced by 'elélèf. This 'elélèf represents the sound a (which is the same sound as was indicated by the ye), and the letter ye indicates (as it does in other instances) the y of the 'ezafé ending -ye.

NUMBERS

27. In Persian the numerals are written and printed from left to right. They have the following shapes:

- Yek 'one' یک
- Do 'two' دو
- Se 'three' سه
- Kohar 'four' کهار
- Paenj 'five' پنج
- Seš 'six' شش
- Hæft 'seven' هفت
- Hæst 'eight' هشت
- Noh 'nine' نه
- Sefr 'zero' صفر

What has been described in the preceding pages is the Persian writing system as it appears in newspapers, periodicals, books, and other material printed on the present-day presses of Iran. In order to acquaint the reader with typewritten Persian, and to give him a general idea about handwritten Persian, a short note on each follows.

The statements made in the notes on combination of words and on punctuation apply to all three forms of the Persian writing system described in the present paper.
The main difference between printed Persian and typewritten Persian consists in the number of the forms of each letter.

1. All the non-connectors, that is, 'ælēf, dal, zal, re, ze, že, and vav, each of which has two forms in printed Persian (Section 4), and two of the connectors, ta and za, each of which has four printed forms (Section 15), have only one form each on the typewriter keyboard:

2. The letters be, pe, te, ze, nun (Section 6); Jim, ce, he (hoti), xe (Section 7); sin, žin (Section 8); sad, zd (Section 14); fe, qaf (Section 10); kaf, qaf (Section 9); lam (Section 11); and mīm (Section 12), which have four forms each in printing, have only two forms each on the typewriter, an initial-medial form and a final-independent form:

3. The letter he (hævvæz) (Section 13), which is a connector, has three forms in typewritten Persian: initial, $; medial, $; and independent-final, $.

4. The letter ye (Section 6.30) has three forms: initial-medial, $; final, $; and independent, $.

5. Two letters, 'eyn and qeyn (Section 16) have four forms each (just as they do in printed Persian): initial, $; medial, $; final, $; and independent, $.

6. Of the ligatures (Section 17), the typewriter has only lam-‘ælēf: $.

7. There is a hæmzé (Section 22) which can be placed above ‘ælēf, vav, or ye. There is also an initial-medial $.
8. The _mæddi_ sign (Section 19), which in printed Persian is part of the type ı, is a separate (dead) key on the typewriter.

9. There is a dotted he hævæz on the typewriter.
HANDWRITTEN PERSIAN

0. As is natural, Persian letters, which have basically the same form in printing, typing, and handwriting, are somewhat different and less easily recognizable in handwriting. This is partly due to the differences among individual writers, but there are also generally accepted written forms of the letters and combinations of letters which are quite different from the printed shapes.

What follows is a description of some of the major differences between a variety of handwritten Persian, which is more or less typical, on the one hand, and printed (and typewritten) Persian, on the other. No attempt will be made to discuss calligraphy or any of a number of other types of handwritten Persian.

1. The three pointed strokes of sin and šin can, in all their forms, be replaced by a slightly curved straight line:

สฉส "head"

However, when there are two successive sins, or šins, or one of each, it is preferred that only one of them be written in the above manner.

2. The final curved stroke of the final and independent sin, šin, sad, sad, and ye may be replaced by a stroke slanting downward to the left and ending in an upward turn to the left. The entire stroke may be angular or slightly curved:

¥¢ "jaime"  "religion"

¥¢ "particular"  "merfs 'ill"

¥¢ "on"  "bi 'without"

3. The curved stroke of the final or independent nun may be replaced by a stroke slanting downward to the left, as in 2 above, but ending in a turn to the right:

¥¢ "an 'that"

¥¢ "man 'I"

4. A preceding connector joins jim, ce, he (hottif), and xe from the top, not from the side, as in typewritten and most printed forms:

¥¢ "baecé 'child"

¥¢ "nex 'thread"
5. A medial be, pe, te, se, nun, and ye before final nun, and sometimes before re, ze, and že, has the shape -I, coming up from the preceding letter and joining the next by a short straight line:

\[ \text{beyn 'between', pir 'old'} \]

6. When followed by final re, ze, že, or one of the final forms described in 1, 2, and 3 above, the pointed stroke of a medial be, pe, te, se, nun, and ye is replaced by a curve:

\[ \text{pir 'old', neebi 'prophet', sir 'milk', cin 'China'} \]

7. Care must be taken not to form a curve such as described in 6 above when joining an initial to a final letter, since the curve may be taken for a letter itself. Compare, for example,

\[ \text{pir 'old', por 'full'} \]

In joining an initial letter to one of the final forms described in 1, 2, and 3 above, the angle between the letters should be rather sharp. Otherwise, any resulting curve may be taken for a letter. Compare, for example,

\[ \text{koš 'Kill!', kešti 'ship'} \]

8. Initial he (hævvez) may be written like a mim with a comma-like addition below it:

\[ \text{honér 'art'} \]

This 'comma' stroke is written from the bottom up:

9. Final he (hævvez) is a simple curve downward, the central part of which is only slightly (if at all) raised:

\[ \text{če 'What?'} \]

10. If necessary, any medial letter and, less frequently, an initial letter, may be stretched, usually to fill in the line:

\[ \text{ketáb 'book', zahér 'apparent'} \]

It is not very common to stretch two successive letters in the same word; nor is it considered permissible to stretch a letter when it is likely to cause ambiguity, as in the case of por 'full', where, if the pe is stretched, the stretching line may be taken for a medial sin (see 1 above), thus resulting in the word pesér 'boy'.

11. All forms of kaf have one oblique bar. This bar is added after the word is written, like the crossing of a t in English. (See also 14 below.) The ' which appears inside the final and independent forms of printed kaf and gaf does not appear in handwriting.

12. The ligature lam-selēf is written ' finally and  independently:

\[ \text{telá 'gold', balá 'above'} \]

Sometimes lam-selēf is written \( \text{or } \):
13. The following observations may be made concerning dots:
(a) Two dots belonging to the same letter can be replaced by a straight line:

\[ \text{'ya or'} \]

(b) Three dots belonging to the same letter can be replaced by a line with a dot above or below it (depending on whether the dots are placed above or below the letter):

\[ \text{'pedar 'father'} \]

\[ \text{'abot 'fixed'} \]

Three dots may also be replaced by a small circle, thus: \[ \text{'ofte} \]

(c) The dots are frequently not placed exactly above or below the letter to which they belong:

\[ \text{'eftezar 'pride, honor'} \]
\[ \text{'boxarf 'stove, heater'} \]

In cases where there is ambiguity, the context usually helps. Compare

\[ \text{'boxarf 'stove'} \]
\[ \text{'nejjarf 'carpentry'} \]

both of which can be written \[ \text{'} \].

(d) Not infrequently, dots are left out altogether. In such cases, the context usually helps in the reading.

14. The dots, haemzés, and all the special signs are added after the whole word, or any part of a word ending in a non-connector, is written.

15. The numerals ١ and ٢ are written ١ and ٢.
COMBINATION OF WORDS

0. Compound words are ordinarily, but not necessarily, written as one word (unless the first word ends in a non-connector). Certain other words are treated as compounds in writing, that is, they are joined together. A few of the commoner cases are listed below:

1. 'this' is sometimes joined to a following short word:
   
   اين
   
   This does not happen when the following word begins with a letter of the group ن س ح.
   
   The above statement is true, though much less frequently, of that:
   
   آن

2. What was said above of اين is true of يك 'one', too.

3. The combinations اين، اين، and اين may join a following short word:
   
   اين، اين، and اين

4. If اين follows a preposition ending in a consonant, it may lose its 'eliff, in which case the two words will be combined:
   
   از اين (for بر اين) برین
   
   This is true of an, too; only here it is the meedek (which, the reader will remember, was an 'eliff originally) which is omitted:
   
   از آن (for بر آن) بران

5. Some writers drop one of the two h's (he haevvés) when they add the plural suffix to a word ending in an h (he haevvés) that represents the vowel e (see Section 13.12):
   
   خانهها (for خانه)

   However, even these writers preserve both the h's if dropping one of them may cause confusion, as in the case of
6. When *mi* joins a verb form beginning with an *l* ("ælél") or a *y* (*ye*), the *ælél* or *ye* of the verb root may be dropped:

\[
\text{Mi aaf} + \text{Mi yábi3} \quad \text{(for Mi aff} + \text{Mi yábi3)}
\]

The dropping of *ælél* may be considered as indicating the fact that the glottal stop represented by *l* ("ælél") is omitted, though even when the *ælél* is preserved, it may quite often be ignored in pronunciation.

7. Some of the cardinal numerals one to ten may join certain words following them (mostly the words سال 'year' and ماه 'month'):

\[
\begin{align*}
\text{Mi sál} & \text{ 'one year'} \\
\text{Mi nég-sál} & \text{ 'five-year-old'} \\
\text{Mi nég-sál} & \text{ 'six-month-old'}
\end{align*}
\]

These numerals may or may not join to 100, هزار, 1,000, 'million', and, occasionally, 1,000,000.

8. The word *ast* has two written forms: *ast* and *ast* (or *ast*). Used after a consonant, the latter represents *ast* combined with the preceding word:

\[
\text{xúbaest 'It is good.'}
\]

After a vowel, *ast* represents *st*:

\[
\text{Ketábi3 'It is a book.'}
\]

However, even if the *l* ("ælél") is retained in such cases, it usually has no sound value; thus, *ketábi3* is pronounced exactly like *ketábi3*.
PUNCTUATION

The following punctuation marks are used in Persian:

- period
- question mark
- exclamation point
- colon
- semi-colon
- comma
- dash
- quotation marks
- parentheses
- brackets
- hyphen

These marks are not consistently used in Persian. There is variation among different writers, and inconsistency may occur in the work of a single writer. When they are used, they generally follow Western, often French, usage. The following notes are meant as a guide, not a description of all possible uses.

1. The PERIOD (.) is used at the end of a declarative or imperative sentence, and sometimes after an interrogative one:

   بردرم را دیدم.  "I saw my brother."
   بنشین.  "Sit down."
   کجا میروید.  "Where are you going?"

   Several periods indicate an omission:
   آیا ... انجا نیست.  "Mr. ... is not here."

2. The QUESTION MARK (?) is used, as in English, after questions:

   اسمیان چهست؟  "What is your name?"

   It may also be used after an indirect quotation which is a question itself or of which the original was a question:

   "آور پریسم مرا بیشاند باید؟"  "I asked him whether he knew me or not."

3. The EXCLAMATION POINT (!) is used, as in English, after sentences expressing surprise, astonishment, emphasis, etc.:
This Isphahani is not an Isphahani!

Be sure not to forget!

It is also used after imperatives:

Stand up!

The COLON (:) is used
(a) before a list, when this list is at the end of the sentence:

I bought these books:...

(b) before a quotation, direct or indirect:

My friend said: "Where are you going?"

My friend asked me where I was going.

(Note that quotation marks are optional.)

(c) before an explanatory note referring to the preceding sentence:

I looked everywhere but I didn't find him: he had left the city.

The SEMI-COLON (;) may be used instead of the colon before an
explanatory note (see 4c above):

I looked everywhere but I didn't find him; he had left here.

The COMMA (,) is used to separate phrases, clauses, or words in
a list (usually equivalent to single bar juncture, |):

the books, the pencils, the pens

We went home, had supper, and went to bed.

It may indicate a parenthetical phrase or clause:

Ahmad, my brother, is here.

Sometimes this indication of the phrasing (juncture) is necessary
to avoid ambiguity, as:

This board is black.

This is the blackboard.

In cases like the following, the comma is used to avoid ambiguity
caused by two successive identical words:

This Isphahani is not an Isphahani!

The DASH (—) is used like the comma (see 6 above). It is also used
in dialogue to mark the beginning of a speaker's remarks:
A dash may also be used to indicate an omission, instead of several periods (see 1 above):

'...is not here.'

8. QUOTATION MARKS (« ») may be used
(a) to mark a direct quotation:

'My father said, ‘Give me the book, please.’

(b) to indicate a foreign word, or for emphasis (as in English):

‘Do you know what ‘chair’ means?’

(c) to set off a parenthetical phrase or sentence:

Ahmad «brother» is having his supper.

9. PARENTHESES (()) may be used instead of quotation marks:

'Ahmad, my brother, is having his supper.'

10. BRACKETS ([ ]) are generally restricted to scholarly works, where they are used as in English:

'Shakespeare was one of the great writers of the 14th [sic] century.'

11. The HYPHEN (—) is not commonly used, but occurs occasionally to indicate the division of a word at the end of a line:

'If you like novels, you can take these books home with you.'

However, words are almost never divided at the end of a line in handwriting and rarely in print. When such a division does occur, it is usually done between a root and an affix, or between affixes, providing the break occurs after a non-connector (see Section 4) and the affix to begin the next line starts with a consonant:
I have read many of his poems in recent years.

You may not, for example, divide السالم، 1967. poem, as the latter begins with a vowel.

Compound words, the members of which occur also as separate words, may be divided between the two parts of the compound: 

 innefamفانه = سهبانخانه 'hotel'