# A guide for expatriates learning to read Pashto

Adam Baker and Najibullah Sadiqi (نجیب‌الله صدیقي)

January 17, 2016

> This document is a work in progress. It is based on "A guide for expatriates learning to read Dari," which is the corresponding guide for learning to read Dari. Pashto and Dari are quite different, but since both use the Arabic script, there is also overlap in the learning process. We hope that this guide will grow and expand as we receive feedback from people learning to read Pashto.

The purpose of this booklet is to serve as an introduction to the written form of the Pashto language, for people who have already acquired some competence in spoken Pashto. Learning to read Pashto is helpful for a variety of reasons, not least in that it helps one to master new vocabulary that might not be reinforced in daily conversations.

Learning to read Pashto is challenging, because of the lack of resources. The best introductory materials at this point are probably the BBC books (information about these is given at the end of this booklet), or Pashto school books published by the government. We can hope that as literacy becomes more widespread, more effective introductory books will become available. We also hope in the future to offer a more structured approach, along the lines of what is provided by the the LCP publication, *A workbook for reading and writing Dari*. That book introduces the conventions of the alphabet, and has writing and listening exercises. Until those resources are available, this booklet remains the only introductory guide. It provides a big-picture overview of the Pashto alphabet, with explicit instruction on certain important points. In learning to read Pashto, it's possible to become lost in the details, but having an overview of the alphabet and writing conventions can prevent that from happening.

As with learning the spoken language, there is absolutely no substitute for time spent on task: only by reading can one become a good reader. Learning the facts *about* reading Pashto that are presented here is no substitute for actually reading Pashto, any more than reading a book about riding a bicycle is a substitute for actually learning to ride.

## 1   Why should a foreigner learn to read Pashto?

Learning to read Pashto is not a small undertaking, so it is appropriate here to consider a few of the many motivations that one might have for learning to read Pashto.

**Learning to read Pashto introduces more formal speech styles.** Some people think that the Pashto spoken on the street is the "true" form of Pashto, and that the literary language is something else entirely. In fact, things are not so cleanly cut. The written form of Pashto is the model for the spoken speech style. This is the variety of Pashto that is encountered in situations involving any formality at all: formal speeches, the way people talk on the radio or television, and when people are reading aloud.

**Learning to read Pashto helps in learning abstract vocabulary.** Being able to read in Pashto opens up new opportunities for language study. It is difficult to learn abstract terms well, because they're not used very often in daily conversation. Reading printed texts can be helpful in mastering abstract vocabulary, since one can rehearse it as often as one wishes by reading.

**Learning to read Pashto helps in learning relevant technical vocabulary.** A language teacher can't teach vocabulary that s/he isn't familiar with. To learn more technical vocabulary—say, related to professional activities—it can be helpful to read printed texts. As with the previous point, one can rehearse newly learned vocabulary much more easily in print than with a language teacher.

**People who can read can use a dictionary.** There are good bilingual Pashto dictionaries available. All of these, of course, are oriented toward the written form of Pashto. Learning to read makes these resources accessible.

**Learning to read in a new language is somewhat easier.** A maxim of literacy is that a person only learns to read once. That is, once one has learned the idea that a letter can correspond to a sound, one has done much of the hard work of learning to read. Learning to read a second language is easier. Pashto has a reasonably clear writing system. It is slightly more complicated than Spanish, for instance, but certainly much less opaque than English. The experience of the first author was that learning to read a cursive, right-to-left script was not as difficult as he thought it would be.

**Learning to read Pashto makes it easier to learn to read other languages, like Dari.** The greater part of the difficulty in learning to read Pashto is learning to read the cursive Arabic alphabet. Since this same script is used for other languages—and particularly, for all other languages in Afghanistan—learning to read in other languages is easier. There are no new letters to learn for Dari, for instance.

**Pashto has hundreds of years of literary history.** In spite of currently high rates of illiteracy in Afghanistan, Pashto is a language of literature. To be illiterate in Pashto is to miss a significant portion of the language, and therefore to be ignorant of a significant portion of the culture.

With these motivations in mind, the new reader should be ready to plunge enthusiastically into the study of written Pashto. The nature of the task is that most of the difficulty lies in the early lessons. Therefore, perseverance is required at the beginning. After a certain point, however, the only difficulty is the occasional unfamiliar word.

## 2 The Alphabet

Pashto is written with a modified version of the Arabic alphabet, which presents three challenges to new learners.[1] First, the Arabic alphabet is written from right to left, instead of from left to right. Second, the letters are often joined together; it is a matter of practice to be able to pull the letters apart in one's head. Third, some vowels are not written. This turns out not to be as much of a problem as one might think it would be, though of course there is a learning curve.

There are forty-four Pashto letters to learn. They are shown in Table 1. A few observations:

- A Pashto letter is composed of "base shape" plus little marks above or below. Several letters have the same base shape and are differentiated only by their dots and/or a circle (e.g., «ب», «پ», «ت», «ټ», and «ث»). One, two, or three dots are used; these are placed either above or below the base shape. The little marks are part of the letter, so it is important to note their number and location. The little circle (e.g., in «ټ») is added to letters that make retroflex sounds (sounds where the tongue turns backward), and also to turn [k] into [g].

- Each letter has a sound. With only five exceptions, each letter always makes the same sound (setting aside different pronunciations in different dialects).

- Each letter is shown with a 'half' and 'full' form (in Pashto, «نیمه توری» [nima torej] and «پوره توری» [pura torej], respectively). These are the forms that a letter takes when it is written in various parts of the word. The half forms occur at the beginning and middle of a word. The full forms occur elsewhere. To be more complete, the shape of a letter can be said to have an isolated, initial, medial, and final form. This is the nomenclature used below, though it is worth noting that the final form is usually just the full form with a connecting line, and that the medial form is usually just the initial form with a connecting line. The different shapes of a letter are called "contextual variants."

- Generally, the contextual variants resemble one another. There are only a handful of letters for which the variants are particularly different.

- The Pashto alphabet has an order, which is used for alphabetization. Some tips for learning the order, which should be done eventually, are given in Appendix A.

---

[1]People who have previously learned to read Arabic need to make some adjustments. Pashto does not have contrastive vowel length, so there is no correspondence between the length of the vowel and whether it is written or unwritten; instead, some vowels are simply not written. Many consonants that have different pronunciations in Arabic have a single pronunciation in Pashto. There are also some new symbols to learn, since Pashto has certain consonant sounds that Arabic does not.

| IPA Name | Olson | IPA | 'Half' | 'Full' |
|---|---|---|---|---|
| alɛf | A, a | ʌ, a | ا | ا |
| bɛ | b | b | بـ | ب |
| pɛ | p | p | پـ | پ |
| tɛ | t | t | تـ | ت |
| ʈɛ | T | ʈ | ټـ | ټ |
| sɛ | s | s | ثـ | ث |
| ʤim | j | dʒ | جـ | ج |
| tʃɛ | ch | tʃ | چـ | چ |
| dzɛ, ze | z | dz | ځـ | ځ |
| tsɛ | s | ts | څـ | څ |
| hɛ | H | h | حـ | ح |
| xɛ | kh | x | خـ | خ |
| dʌl | d | d | د | د |
| ɖʌl | D | ɖ | ډ | ډ |
| zʌl | z | z | ذ | ذ |
| rɛ | r | r | ر | ر |
| ʈɛ | R | ɽ, ɻ | ړ | ړ |
| zɛ | z | z | ز | ز |
| ʒɛ | zh | ʒ | ژ | ژ |
| ʝɛ/gɛ/ʐɛ | g | ʝ/g/ʐ | ږ | ږ |
| sin | s | s | سـ | س |
| ʃin | sh | ʃ | شـ | ش |
| çin/xin/ʂin | x | ç/x/ʂ | ښـ | ښ |
| sʌd | s | s | صـ | ص |
| zʌd | z | z | ضـ | ض |
| tuj | t | t | طـ | ط |
| zuj | z | z | ظـ | ظ |
| ajn | — | — | عـ | ع |
| ɣajn | gh | ɣ | غـ | غ |
| fɛ | f | f | فـ | ف |
| qʌf | q | q | قـ | ق |
| kʌf | k | k | کـ | ک، ك |
| gʌf | g | g | گـ | ګ |
| lʌm | l | l | لـ | ل |
| mim | m | m | مـ | م |
| nun | n | n | نـ | ن |
| ɳun | N | ɳ | ڼـ | ڼ |
| wʌw | U, O, w | u, o, w | و | و |
| hɛ | H, e, a | h, ɛ, a | هـ | ه |
| narma jɛ | I, y | i, j | یـ | ى |
| tsarganda jɛ | I | i | يـ | ي |
| uʝda jɛ | E | e | ېـ | ې |
| dzwarakəj wʌla jɛ | əy | əj | یـ | ی |
| ʔamrija jɛ | əy | əj | ئـ | ئ |

Table 1: The entire Pashto alphabet, in alphabetical order, with the various forms of the letters, and pronunciations in IPA and Olson script. Start reading on the right. Pronunciations separated by a slash (/) vary between dialects; those separated by a comma are different possible pronunciations of a letter, irrespective of the dialect.

3

## 2.1 Connecting and Non-connecting letters

The Pashto alphabet is divided into connecting letters and non-connecting letters. Connecting letters always join to a following letter in the word. When it connects to a following letter, its shape changes. The connected form of the letter—which is always narrower—is called the [nima torej] 'half' form; the unconnected form is called the [pura torej] 'full' form. On the other hand, non-connecting letters *never* join to a following letter. «ن» [nun] is an example of a connecting letter: it *always* connects to a following letter in the same word. «و» [waw] is an example of a non-connecting letter: it *never* connects to a following letter. The example below shows how this works out for the word «نون» [nun] 'the letter nun', which is spelled ن-و-ن [nun-waw-nun].

(1) ن + و + ن = نون

«ن» [nun] is a connecting letter, and since there is a letter following it («و» [waw]), the «ن» [nun] connects to it, changing its shape in the process: نو. But «و» [waw] is a non-connecting letter, so it does not connect to the second «ن» [nun]. The second «ن» [nun] does not have any letter after it, so it is written in its isolated form. It should be kept in mind that a connecting letter *always* connects to the following letter, even if that letter is itself non-connecting. In the example above, the first «ن» [nun] connects to the «و» [waw], even though «و» [waw] is a non-connecting letter. "Non-connecting" means only that a letter does not connect to the next letter *following* it.

The majority of letters in Pashto are connecting. The non-connecting letters of Pashto are shown below. It might help to notice that these are the letters «ا», «د», «ر», and «و», plus all the letters formed by adding little marks to them.

(2) ا د ذ ډ ر ړ ز ژ ږ و

A connected script such as Pashto's poses challenges for people who are used to reading Latin script. The first challenge is that there are more forms of the letters to recognize. In Latin script, letters have two shapes: the upper and lower cases. There are more potential shapes of Pashto letters.[2] For the most part, however, these letter shapes resemble one another very closely.

A second difficulty is that the *widths* of letters varies in Pashto. A 't' in Latin script is the same size whether it occurs at the beginning, middle, or end of a word (e.g., 'tack', 'atom', and 'cat'). But in Pashto, the widths of the letters varies considerably: the initial «ب» [bɛ] in «بد» [bad] 'bad' is much narrower than the final «ب» [bɛ] in «سبب» [sabab] 'reason.' For this reason, although it is tempting to think of the isolated form of the Pashto word as "basic," it is actually better to think of the initial form of the Pashto letter as the basic one. If one is trained to recognize «س» as the letter [sin], then it is more difficult to recognize it when it occurs word-initially—as in the word «سل» [səl] 'hundred'. That is because one of the most visually prominent features of the isolated form «س» [sin] is the "swash" at the end of the letter. But in the initial and medial positions there are no swashes. If «س» [sin] is learned as the basic form, however, it is still easy to recognize it at the end of a word, for instance in «سل» [las] 'ten'. The initial and medial forms are also much more frequently encountered in texts. It takes some time for the eye to get used to recognizing the letter shapes with their different widths, however.[3]

Finally, it should be emphasized that in Pashto, not every break in the cursive script corresponds to a word break.[4] Single words such as «نون» or «جومات» have "gaps" in them, but are still just one word. Spaces inside of words are narrower than spaces between words. It is natural for a beginner to have difficulty in identifying where words begin and end.[5]

## 2.2 Connecting consonants

The connecting consonants are described below in groups, according to the basic shapes of the letters. The letters «ب» [bɛ], «پ» [pɛ], «ت» [tɛ], «ټ» [ʈɛ], and «ث» [sɛ] are identical except for the position and number of

---

[2]Note that the two languages are not quite parallel. In Latin-script languages, upper case letters mark the beginnings of sentences and certain kinds of words. In Pashto, the initial/medial/final/isolated forms are used irrespective of what kind of word is being written: only the position of the letter in the word is significant.

[3]A tangential note on word length: Latin-based ideas of what a short and long word look like do not carry forward in reading Pashto. This is because of varying widths of letters, and also because some vowels are not written. The words «سبب» and «حرکت» have a similar width on the printed page, but the former has two syllables ([sabab] 'reason') and the latter has three ([harakat] 'movement').

[4]This is unlike cursive scripts in the Latin alphabet: for those, all letters are connecting, so every gap corresponds to a word break.

[5]Unfortunately, in printed text this difficulty is often exacerbated, with spaces inside of words that are as wide as the spaces between words. It is sometimes assumed that this is a problem with the Arabic script, but in fact it is merely poor font design. The problem never arises in a handwritten text.
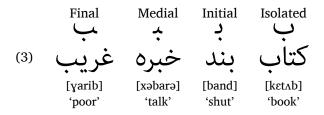
their dots and the presence of absence of the small circle: thus they have the same base shape. The contextual variations in the base shapes of the letters for these four letters are all the same: if one can recognize a «ب» [bɛ] in the middle of a word, one can also recognize a «پ» [pɛ].

This is a good time to start learning to recognize and name the letters. It it important to pay attention to the little marks. One wouldn't want to learn «ب» [be] as "the letter that's shaped like a bowl," because that's an equally good descriptor of «پ» [pɛ], «ت» [tɛ], «ټ» [ʈɛ], and «ث» [sɛ] as well.

### 2.2.1 The ب-group

The ب-group is composed of the letters «ب» [bɛ], «پ» [pɛ], «ت» [tɛ], «ټ» [ʈɛ], and «ث» [sɛ]. These correspond respectively to the sounds [p], [b], [t], [ʈ], and [s]. Of these, «ث» [sɛ] is encountered only very infrequently.

The various forms of «ب» [bɛ] (representing of the ب-group) are shown below, written in isolation and also in context. At this point the other letters are unfamiliar, but try to identify the «ب» [bɛ] in each word. The pronunciation is given for each word, but it's not important to pay attention to that; it just bothers some people not to have it.
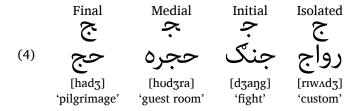
|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ـب | ـبـ | بـ | ب |
| (3) | غریب | خبره | بند | کتاب |
|  | [ɣarib] | [xəbarə] | [band] | [kɛtʌb] |
|  | 'poor' | 'talk' | 'shut' | 'book' |

The isolated form is written when the letter follows a non-connecting letter (in the «نون» [nun] example above, this is the «و» [waw]).[6] The initial form occurs when the letter is at the beginning of a word, or if the letter follows a non-connecting letter. For instance, in the word «کمربند» [kamarband], an initial «ب» [bɛ] is used because the preceding letter («ر») is non-connecting. The medial form occurs when the letter follows a connecting letter, and is followed by a letter. The final form occurs at the end of a word, following a connecting letter.

Notice how «ت» makes the [t] sound, and «ټ» makes the [ʈ] sound. It's as if «ت» is the 'plain' [t] and the little circle makes it retroflexed.[7] You can use this as a mnemonic to help you remember the letters.

This group also illustrates the principle that it is better to think of the initial form as basic, with the isolated (or final) forms having an extraneous swash to complete the letter. Learning «ب» [bɛ] as the basic form is not helpful, since the letter has such a different appearance in the initial and medial forms.

### 2.2.2 The ج-group

The ج-group consists of the letters «ج» [dʒim], «ځ» [dzɛ],[8] «څ» [tsɛ], «چ» [tʃɛ], «ح» [hɛ], and «خ» [xɛ], which correspond to the sounds [dʒ], [dz], [ts], [tʃ], [h], and [x] (though see the note below on dialect variation with «ځ» [dzɛ]). The positions of letters in the ج-group are shown below.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ـج | ـجـ | جـ | ج |
| (4) | حج | حجره | جنګ | رواج |
|  | [hadʒ] | [hʊdʒra] | [dʒaŋg] | [rɪwʌdʒ] |
|  | 'pilgrimage' | 'guest room' | 'fight' | 'custom' |

Here again, it is advantageous to view the initial/medial forms as basic, and the "tail" as something required just for the isolated and final forms. Try not to focus on the tail of this letter, since it doesn't occur in all positions.

The [h]-sound is commonly dropped in spoken Pashto, so the presence of «ح» [hɛ] in the written form can be surprising. For instance, in the word «امتحان» [ɛmtɛhʌn] 'test', the spoken realization of the «ح» [hɛ] is more like a [j]: [ɛmtɛjʌn].
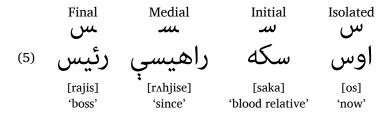
---

[6]The isolated form also appears when the letter is being referred to, such as "the letter «ب»."

[7]All of the letters with the circle follow this pattern, except that «ګ» is just [g], not a retroflex.

[8]In some dialects of Pashto, and in some words, «ځ» [dzɛ] can make the [z] sound.

### 2.2.3 The س-group

The three letters of the س-group are «س» [sin] and «ش» [ʃin], and «ښین» [çin] which represent the sounds [s], [ʃ], and [ç] (though see below for dialect variation with «ښین» [çin]). Previously it was seen that «ث» [sɛ] also represents the sound [s]; «س» [sin] is much more commonly used, however.
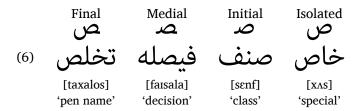
| Final | Medial | Initial | Isolated |
|:-----:|:------:|:-------:|:--------:|
| س | سـ | ســ | س |
| رئیس رايس (5) | راهيسي | سكه | اوس |
| [rajis] | [rʌhjise] | [saka] | [os] |
| 'boss' | 'since' | 'blood relative' | 'now' |

The challenge in identifying these letters is that the vertical lines in the letters also resemble the medial forms of the ب-group letters. Identifying individual letters in words such as «سبب» [sabab] 'reason' (ب + س + ب) is challenging at first.

The letter «ښین» [çin] makes different sounds in different dialects. [ç] is the Ghazniwala pronunciation. [ʂ] is the Kandahari pronunciation. Eastern dialects have [x]. In this guide we have used [ç] consistently, since that at least gives «ش» a consistent pronunciation letter, which you can adjust for your own dialect. Note that if your dialect has [x]—i.e., if «ښه» 'good' has the same sound at the beginning as «خلک» 'people'—then you will need to memorize the spellings of words with the [x] sound.
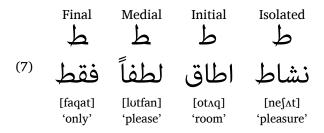
### 2.2.4 The ص-group

The letters «ص» [sʌd] and «ض» [zʌd] are fairly rarely encountered. «ص» [sʌd] is an uncommon representation of [s]—for which «س» [sin] is more commonly seen. «ض» [zʌd] represents [z], for which «ز» [zɛ] is more commonly used. These letters are used in certain words of Arabic origin.

| Final | Medial | Initial | Isolated |
|:-----:|:------:|:-------:|:--------:|
| ص | صـ | صـ | ص |
| تخلص (6) | فيصله | صنف | خاص |
| [taxalos] | [faɪsala] | [sɛnf] | [xʌs] |
| 'pen name' | 'decision' | 'class' | 'special' |

### 2.2.5 The ط-group

The letters «ط» [tuj] and «ظ» [zuj] represent [t] and [z], respectively. As with the previous group, these are the less common representations of these sounds, occurring mainly in words of Arabic origin. The letters are virtually identical in all positions.

| Final | Medial | Initial | Isolated |
|:-----:|:------:|:-------:|:--------:|
| ط | ط | ط | ط |
| فقط (7) | لطفاً | اطاق | نشاط |
| [faqat] | [lʊtfan] | [otʌq] | [neʃʌt] |
| 'only' | 'please' | 'room' | 'pleasure' |

### 2.2.6 The ع-group

The letters «ع» [ajn] and «غ» [ɣajn] form the ع-group. «ع» [ajn] is commonly encountered, but it is often silent. In careful speech, it may be pronounced as a glottal stop ([ʔ]; the middle sound in *uh-oh*).[9] The letter «غ» [ɣajn] is the [ɣ] sound. As can be seen below, the contextual variants of this letter are fairly distinct.

---

[9]The "proper" Arabic pronunciation is a voiced pharyngeal fricative [ʕ], or as a pharyngealized glottal stop [ʔˤ]. This distinctive sound might be produced by educated speakers, or those who have had training in Arabic.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ع | ڪ | ڬ | ع |
| (8) | منع | تعلیم | عید | موضوع |
|  | [mana] | [talim] | [id] | [mawzo] |
|  | 'prohibition' | 'education' | 'Eid' | 'subject' |

### 2.2.7 The ف-group

«ف» [fɛ] and «ق» [qʌf] correspond to the sounds [f] and [q].[10] While these letters are grouped here, it is notable that the isolated and final forms of the letters differ, in that «ق» [qʌf] has a rounder and shorter "tail," while that of «ف» [fɛ] is longer and less curved.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ف | ڧ | ڧ | ف |
| (9) | مکلف | منفي | فرق | مصروف |
|  | [mokalaf] | [manfi] | [farq] | [masruf] |
|  | 'bound, charged (masc.)' | 'negative' | 'difference' | 'busy' |

### 2.2.8 The ک-group

«ک» [kʌf] and «ګ» [gʌf] represent the sounds [k] and [g]. The only difference between these two letters is the little circle.[11]

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ک | ک | ک | ک |
| (10) | توشک | مشکل | کشر | ناک |
|  | [toʃak] | [muʃkɛl] | [kəʃər] | [nʌk] |
|  | 'mattress' | 'difficult' | 'younger' | 'pear' |

### 2.2.9 Other connecting consonants

The remaining connecting consonants all have unique base shapes. They are «ل» [lʌm], «م» [mim], «ن» [nun], and «ه» [he].

«ل» [lʌm] makes the [l] sound. In reading this letter, it is important not to confuse it with «ا» [alɛf], which is similar in that it has a single tall vertical stroke. The key to differentiating the two letters is that «ل» [lʌm] is a connecting letter, so in the initial and medial positions it *always* connects to a following letter, whereas «ا» [alɛf] *never* does. In the final and isolated forms, «ل» [lʌm] has a tail, whereas «ا» [alɛf] does not.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ل | لـ | لـ | ل |
| (11) | بل | قلم | لند | بندول |
|  | [bəl] | [qalam] | [land] | [bandawəl] |
|  | 'other' | 'pen' | 'short' | 'to shut' |

One final complication for «ل» [lʌm] and «ا» [alɛf] is that when the two letters connect, they are written with a special ligature, a single letter that represents both the original ones. The lam-alef sequence is written «لا» in isolation, or «لا» when it is joined. For instance, the name of the letter «ل» [lʌm] is written «لام». The word [halʌl] is written as «حلال».

«م» [mim] makes the [m] sound; its forms are shown below.

---

[10]A silly mnemonic to remember which is which: [f] is made with the lips; [q] is made in the back of the throat. There is only one pair of lips (one dot in «ف» [fɛ]), but two tonsils in the back of the throat (two dots in «ق» [qʌf]).

[11]Though «ک» is the usual letter for the [g] sound, some writers recommend using «ګ» instead, which is the letter that is used to write Dari.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
| (12) | مٖ | مٖ | مٖ | مٖ |
|  | سٖم | ایمان | مالٖته | لٖرٖم |
|  | [sam] | [imʌn] | [mʌlʈa] | [laɽəm] |
|  | 'correct' | 'faith' | 'orange' | 'scorpion' |

«ن» [nun] makes the [n] sound, and—following the usual pattern of the little circle—«ن» makes the [ŋ] sound. The forms of «ن» [nun] are shown below. As seen below, in the initial and medial positions this letter acts like it belongs in the ب-group, but the medial and final forms are quite different.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
| (13) | ن | ‎ـنـ | ن | ن |
|  | نن | منل | نرس | طوفان |
|  | [nən] | [manəl] | [nars] | [tofʌn] |
|  | 'today' | 'to accept' | 'nurse' | 'storm' |

The letter «ه» [hɛ] alternately represents a vowel or a consonant. In word-initial or word-medial position, it represents the sound [h]. At the end of the word, it can either be [a], or [ə]. The former is illustrated by «مینه» [mina] 'love', the latter by «زه» [zə] 'I'.

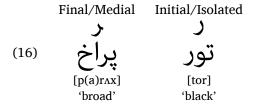|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
| (14) | ـہ | ھ | ھ | ه |
|  | ونه | بهرنی | هګی | زره |
|  | [wəna] | [bahranej] | [hagəj] | [zɽə] |
|  | 'tree' | 'foreigner' | 'egg' | 'heart' |

There is an alternate medial form of «ه» [hɛ] that is often used in handwriting, and occasionally seen in print: instead of «بهرنی» [bahranej], for instance, «ه» [hɛ] is written as «بهرنی».

## 2.3 Non-connecting consonants

### 2.3.1 The د-group

The sounds «د» [dʌl], «ډ» [dʌl] and «ذ» [zʌl] make the [d], [ɖ] and [z] sounds. «د» [dʌl] is the only way to write [d] and «ډ» [ɖʌl] is the only way to write [ɖ], but «ذ» [zʌl] is a less common way to write [z]. The medial/final form has a somewhat different shape than the initial/isolated form, as shown below.

|  | Final/Medial | Initial/Isolated |
|---|---|---|
| (15) | ـد | د |
|  | دروند | دغه |
|  | [drund] | [daɣa] |
|  | 'heavy' | 'this' |

### 2.3.2 The ر-group

The ر-group is comprised of «ر» [rɛ], «ړ» [ɽɛ], «ز» [zɛ] and «ژ» [ʒɛ], and «ږ» [jɛ]. «ر» [rɛ] makes the [r] sound. «ړ» [ɽɛ] makes the [ɽ] sound.[12] «ز» [zɛ] makes the [z] sound; it is the most common letter used to represent this sound. «ژ» [ʒɛ] makes the [ʒ] sound. «ږ» [jɛ] makes different sounds, depending on the dialect (see below). For these letters the medial/final and initial/isolated forms are nearly identical.

---

[12]At the end of a word, this letter makes a drawn-out r-sound, which can be written with the IPA letter [ɻ].

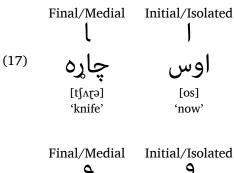|  | Final/Medial | Initial/Isolated |
|---|---|---|
|  | رـ | ر |
| (16) | پراخ | تور |
|  | [p(a)rʌx] | [tor] |
|  | 'broad' | 'black' |

The letter «ر» [jɛ] makes the [j] sound in Ghazniwala Pashto, the [ʐ] sound in Kandahari Pashto, and the [g] sound in Eastern Pashto. The example words in this book assume the [j] pronunciation. As with other sounds that vary between dialects, it may be necessary for you to memorize spellings. If «ژومنخ» 'comb' starts with the same sound as «ګډ» 'mixed' in your dialect, you will have to memorize the spellings.

This concludes the discussion of the consonant letters. Since the vowels are a little more complicated, it would be worth taking a break at this point to review the consonants: their shapes, their names, and their sounds. It is a good idea have a working knowledge of the consonants before moving on to the next section.

# 3  Vowels

While the correspondences between the consonant letters and their sounds is straightforward, vowels are somewhat more complicated. The Arabic script provides three vowel symbols: «ا» [alɛf], «و» [wʌw], and «ی» [narma jɛ]. In Pashto, there are nine simple (monophthongal) vowels and four diphthongs. As a result, many vowels are not written. Approximately: vowels other than [ʌ], [i], and [u] are not written. The reader must infer when a vowel is present. When an experienced reader sees the word «بند», which has no vowel letters, s/he knows intuitively that the word is [band] 'shut'. For a beginner, this may seem like a recipe for chaos, since «بند» might also be pronounced as [bɛnd], [bɛnɛd], [bʊnd], [bʊnad], [bənəd], [bənd], etc.; in practice, it is not that bad, because there's usually only one actual word that the letters can spell.

First, we review the positional forms of the three vowel letters. Of the three, «ا» [alɛf] and «و» [waw] are non-connecting, and vowels of the ی-group are connecting.

|  | Final/Medial | Initial/Isolated |
|---|---|---|
|  | ـا | ا |
| (17) | چاره | اوس |
|  | [tʃʌɽə] | [os] |
|  | 'knife' | 'now' |

|  | Final/Medial | Initial/Isolated |
|---|---|---|
|  | و | و |
| (18) | غوا | ولې |
|  | [ɣwʌ] | [wale] |
|  | 'cow' | 'why' |

The shapes of «ا» [alɛf] and «و» [waw] are nearly identical in either context. Recall from the discussion of «ل» [lʌm] that «ل» [lʌm] and «ا» [alɛf] look similar, but can be differentiated in two ways: in the initial/medial forms, «ل» [lʌm] connects to a following letter but «ا» [alɛf] does not; in final/isolated forms, «ل» [lʌm] has a tail, but «ا» [alɛf] does not.

|  | Final | Medial | Initial | Isolated |
|---|---|---|---|---|
|  | ـي | ـيـ | يـ | ی |
| (19) | شخصي | مينه | يو | سړي |
|  | [ʃaxsi] | [mina] | [jaw] | [saɽi] |
|  | 'personal' | 'love' | 'one' | 'men' |

The forms of the ی-group are illustrated with «ي» [tsarganda jɛ], since that letter occurs in all positions in a word. Letters in this group look like ب-group letters in the initial and medial forms. They have more exciting forms for the final and isolated contexts.

Turning now to the pronunciation of these letters, it is important to note that all three vowels letters have multiple possible sounds. This is different to the consonants, which are nearly all consistent. Interpreting the vowel letters is one of the more difficult parts of learning to read Pashto.

## 3.1   The ی-group

There are five vowels that have the base shape «ی»: «ی», «ي», «ېٔ», «ۍ»,[13] and «ئ». These letters make different sounds, but they are also very important in understanding the endings of nouns and verbs. Although the sounds of the letters are listed here, it might be easiest to make the sense of the different sounds if you look over the paradigms in sections 4 and 5.

The letter «ي» [tsarganda jɛ] can be the consonant [j], as in «يو» [jaw] 'one,' the vowel [i], as in «تکلیف» [taklif] 'difficulty.' How can these distinct pronunciations be distinguished? In the *middle* of the word, there is no way to tell, unless you know the word. «تکلیف» could be [taklajaf], but it happens that it's [taklif]. Things are different at the beginning of a word, however. When the sound [i] occurs at the beginning of a word, it is represented by «يا». The «ا» [alɛf] shows that the word begins with a vowel sound. Thus [imʌn] 'faith' is written «ایمان». If there is no «ا» [alɛf] at the beginning, the letter is making the consonant sound [j]. As noted above, [jaw] 'one' is written «يو».

The letter «ی» [narma jɛ] (without dots) occurs only at the end of certain masculine singular nouns and adjectives, and makes the [ej] sound,[14] as in «سړی» [saɽej] 'man (direct singular)'.[15]

The letter «ي» [uʝda jɛ] (with two vertical dots) makes the [e] sound, as in «کیږو» [keju] 'we become'.[16] When this letter occurs at the end of a noun, the noun will always be feminine (cf. Tables 5 and 6).

The letter «ۍ» [dzwarakəj wʌla jɛ] makes the [əj] sound, and occurs only at the ends of feminine nouns, as in «کړکۍ» [kaɽkəj] 'window'.

The letter «ئ» [ʔamrija jɛ] also makes the [əj] sound. It only occurs at the end of verbs as part of the 'you (pl.)' personal ending, e.g., «کیږئ» [kejəj] 'you (pl.) become'.

## 3.2   The vowel «و»

The situation with «و» [waw] is closely parallel to that of «ي» [tsarganda jɛ], except that there is only one of them! It can be the consonant [w], or one of the three vowels [u], [o], or [aw].

At the beginning of a word, «و» [waw] is always [w], as in «واحد» [wʌhɛd] 'Wahed.' If a word begins with one of the vowels, it is written «او», as in «اوبه» [ubə] 'water.'

«و» [waw] can also be used to write [o], as in «تول» [tol] 'all.' As you might expect, when the sound occurs at the beginning of the word it is written «او», as in «اور» [or] 'fire.'

«و» [waw] can also be used to write [aw], as in «موضوع» [mawzo] 'subject.' It's as though there is a silent [a] sound, followed by a «و» [waw] that makes the [w] sound.

In parallel with «ي» [tsarganda jɛ], then, «و» [waw] can be used to write either a consonant or a vowel. If «و» [waw] occurs at the beginning of a word, it is a consonant; at the beginning of a word the vowel is written «او». If it occurs elsewhere, it is a matter of recognizing the word. As a vowel, «و» [waw] can be either [u], [o], or [aw].

## 3.3   The vowel «ا»

The vowel «ا» [alɛf] in the middle or end of the word makes the [ʌ] sound, as in «ما» [mʌ] 'I (oblique)'. At the beginning of the word, things are quite different. As discussed previously, «ا» [alɛf] can be used to write [i], [u], [o], and [aw] at the beginning of a word. When «ا» [alɛf] is written by itself at the beginning of a word, it can be either [a], [ɛ], [o], or [ʊ]. For instance [ɛmtɛhʌn] 'test' is written «امتحان», [aspa] 'horse' is written «اسپه», and [ʊzbak] 'Uzbek' is written «ازبک». One can tell that these are representing one of these three sounds—though not *which* one—because the following letter is not «ی» [jʌ] or «و» [waw].

To illustrate these principles, consider the word «اذان». One can't be certain whether the first «ا» [alɛf] is [a], [ɛ], [o], or [ʊ], but the second «ا» [alɛf] *must* represent [ʌ]. In this case, the word is [azʌn] 'call to prayer'.

---

[13]«ۍ» could technically be considered to have a different base shape, since the little tail is added to the end of the swash.

[14]It makes the [ej] sound in Eastern Pashto, but in other dialects it makes the [aj] sound.

[15]A special note is in order for «ي» and «ی», which are differentiated by the two (horizontal) dots beneath. These are distinct letters in Pashto, but not in Dari and Arabic. (In Arabic, «ي» is used. In Dari, «ی» should be used, but «ي» also creeps in as a typing mistake.) In the middle of a word, they are both written with two dots underneath, so it is really ambiguous whether the letter is «ی» or «ي». If this is confusing then you don't need to think any more about it.

[16]This sound does not occur at the beginning of a word, so there is no spelling like «يا» at the beginning of a word.

| Meaning | Direct | | Oblique | |
|---|---|---|---|---|
| | Written form | Pronunciation | Written form | Pronunciation |
| I | زه | zə | ما | mʌ |
| you (sg.) | ته | tə | تا | tʌ |
| he | دى | dej | دۀ | də |
| she | دا | dʌ | دي | de |
| we | موږ ، مونږ | munɟ | موږ ، مونږ | munɟ |
| you (pl.) | تاسو | tʌso | تاسو | tʌso |
| they (masc. or mixed) | دوى | duj | دوى | duj |
| they (fem.) | دوى | duj | دوى | duj |

Table 2: Personal pronouns in the direct and oblique case. Note that using the 'you (pl.)' is a way to show respect when individuals are addressed.

| Meaning | Pronunciation | Written form |
|---|---|---|
| tired doctor | stəɽej ɖʌkʈər | ستړى ډاکتر |
| tired doctors | stəɽi ɖʌkʈərʌn | ستړي ډاکتران |
| tired teacher | stəɽej çowunkej | ستړى ښوونکى |
| tired teachers | stəɽi çowunki | ستړي ښوونکي |
| tired Pashtun | stəɽej pəçtun | ستړى پښتون |
| tired Pashtuns | stəɽi pəçtʌnə | ستړي پښتانه |
| big man | ɣaʈ saɽej | غټ سړى |
| big men | ɣaʈ saɽi | غټ سړي |

Table 3: Masculine nouns in the direct case

How is a word written if it begins with the [ʌ] sound? In that case the sound is written with a different letter, «آ». The word [ʌlu] 'potato' is written «آلو». «آ» is called [alɛf mad].

To fill in the picture of vowels at the beginnings of words, recall from the section on consonants that some consonant letters are not pronounced. A word that in spoken Pashto begins with a vowel sound might actually begin with a consonant letter, such as «ه» [hɛ], «ح» [hɛ], or «ع» [ajn]. Examples of this are «عزیز» [aziz] 'Aziz', «هوښیار» [oçjʌr] 'intelligent, savvy', and «حساب» [ɛsʌb] 'arithmetic'.

# 4 Nouns

Pashto nouns can be either masculine or feminine. They are inflected either in direct or oblique case.[17] This section will introduce a few paradigms to introduce the written forms of the pronouns.

Table 2 shows the personal pronouns (I, you, he, she, it, ...) in the direct and oblique cases. The spellings are predictable. 'We' can be written with or without the «ن» («مونږ» or «موږ»). Also, it is worth noting that oblique form of 'he'—«دۀ»—is shown here with a small diacritic called [hamza]. This diacritic is often not written, but when it is written, it helps to identify a form as masculine.

Examples of masculine nouns in the direct case are shown in Table 3. The spellings are just what one would expect, given the pronunciations. This is a helpful paradigm for rehearsing the difference between «ي» and «ى». Similarly, the masculine nouns in the oblique case, given in Table 4, are what one would expect. (This table shows dialectal variation in the form 'big men', which is reflected in the written forms as well.)

Table 5 shows feminine nouns in the direct case. The spellings are predictable, and the use of «ۍ» to write the [əj] sound is illustrated, e.g., in 'big dog'. The oblique forms are shown in Table 6, and are fully predictable. One noteworthy thing from this paradigm is that when the sound [əj] occurs in the middle of a word, it is written with «ي» rather than «ۍ»: as in «سپيو» [spəjo] 'dogs (obl.)'. The [ə] sound is unwritten as usual, and the «ي» makes the [j] sound.[18]

---
[17]If these terms are unfamiliar, "direct" refers to the case used for subjects in present tense, and "oblique" is the other case. For the word 'I' or 'me,' [zə] is in the direct case, and [mʌ] is in the oblique case.

[18]A different way to think of it would be that in the middle of a word «ۍ» looks like «ى» or «ي» in the middle of the word.

| Meaning | Pronunciation | Written form |
|---|---|---|
| tired doctor | stəɽi ɖʌkʈər | ستړي ډاكټر |
| tired doctors | stəɽo ɖʌkʈərʌno | ستړو ډاكټرانو |
| tired teacher | stəɽi çowunki | ستړي ښوونكي |
| tired teachers | stəɽo çowunko | ستړو ښوونكو |
| tired Pashtun | stəɽi pəxtun | ستړي پښتون |
| tired Pashtuns | stəɽo pəxtano | ستړو پښتنو |
| big man | ɣaʈ saɽi | غټ سړي |
| big men | ɣaʈ saɽo | غټ سړو |
| | ɣaʈ saɽjo | غټ سړيو |

Table 4: Masculine nouns in the oblique case

| Meaning | Pronunciation | Written form |
|---|---|---|
| tired doctor | stəɽe ɖʌkʈəra | ستړې ډاكټره |
| tired doctors | stəɽe dʌkʈərʌne | ستړې ډاكټراني |
| big tree | ɣaʈa wəna | غټه ونه |
| big trees | ɣaʈe wəne | غټې ونې |
| big dog | ɣaʈa spəj | غټه سپۍ |
| big dogs | ɣaʈe spəj | غټې سپۍ |
| tired teacher | stəɽe çowunke | ستړې ښوونكې |
| tired teachers | stəɽe çowunke | ستړې ښوونكې |

Table 5: Feminine nouns in the direct case

| Meaning | Pronunciation | Written form |
|---|---|---|
| tired doctor | stəɽe ɖʌkʈəre | ستړې ډاكټرې |
| tired doctors | stəɽo dʌkʈərʌno | ستړو ډاكټرانو |
| | stəɽjo dʌkʈərʌno | ستړيو ډاكټرانو |
| big tree | ɣaʈe wəne | غټې ونې |
| big trees | ɣaʈo wəno | غټو ونو |
| big dog | ɣaʈe spəj | غټې سپۍ |
| big dogs | ɣaʈo spəjo | غټو سپيو |
| tired teacher | stəɽe çowunke | ستړې ښوونكې |
| tired teachers | stəɽo çowunkjo | ستړو ښوونكيو |

Table 6: Feminine nouns in the oblique case

| Meaning | Pronunciation | Written form |
|---------|---------------|--------------|
| keʒ-əm | I become | کېږم |
| keʒ-e | you (sg.) become | کېږې |
| kiʒ-i | he/she/it becomes | کېږي |
| keʒ-u | we become | کېږو |
| keʒ-əj | you (pl.) become | کېږئ |
| kiʒ-i | they become | کېږي |

Table 7: Subject agreement in the present tense

| Meaning | Pronunciation | Written form |
|---------|---------------|--------------|
| I arrived | wə-rased-əm | ورسېدم |
| you (sg.) arrived | wə-rased-e | ورسېدي |
| he arrived | wə-rased-ə | ورسېدهٔ |
|  | wə-rased | ورسېد |
| she arrived | wə-rased-a | ورسېده |
| we arrived | wə-rased-u | ورسېدو |
| you (pl.) arrived | wə-rased-əj | ورسېدئ |
| they (masc.) arrived | wə-rased-əl | ورسېدل |
| they (fem.) arrived | wə-rased-əle | ورسېدلې |

Table 8: Subject agreement in the past tense (intransitive)

# 5 Verbs

This section illustrates the written forms of verbs. Pashto grammar requires us to consider the present tense separately from the past tense, and the past tense of transitive verbs separately from the past tense of intransitive verbs. Table 7 illustrates subject agreement in the present tense for the verb 'to become'. The spellings reflect the pronunciations. Note that, at the end of a verb, the [əj] sound is written with «ئ».

Verb agreement in the past tense intransitive—illustrated in Table 8 for the verb 'to arrive'—is much the same. The difference is that in the past tense, there is a difference in the third person for the masculine and feminine gender. The form for 'he arrived', if written with the «ه», can be written with the hamza marker («هٔ») to emphasize that it is masculine. The hamza is not always written, so you should not rely on it.[19] For the past tense transitive forms, shown in Table 9 for the verb 'to hit', the writing is no more complicated. The spellings are parallel to the past tense intransitive, in spite of the fact that the verb now agrees with the object rather than the subject.

# 6 Remaining complications

- In Pashto the [an] suffix changes a word into an adverb (like the English *-ly* suffix). For instance, [mutʌbɛq] 'corresponding' becomes [mutʌbɛqan] 'correspondingly'. There is a special spelling for this in Pashto, using an «ا»-and-diacritic combination: «اً». The previously mentioned words are spelled «مطابق» 'corresponding' and «مطابقاً» 'correspondingly'.

- (Rare) At the end of words a «ی» [narma jɛ] sometimes makes the [ʌ] sound. An example is «معنی» [mʌnʌ] 'meaning'.

- (Rare) Occasionally a diacritic (called [taʃdid]) is written over a consonant to indicate a lengthened pronunciation «نجّار» [nadʒʌr] 'carpenter'. This is not encountered consistently in print, and it does not

---

[19]The pronunciation of the 3rd person masculine singular ('he arrived') in particular varies quite a bit with dialect, so you may need to make adjustments here for the dialect that you know.

| Meaning | Pronunciation | Written form |
|---|---|---|
| hit me | wə-wahəl-əm | ووهلم |
| hit you (sg.) | wə-wahəl-e | ووهلې |
| hit him/it | wə-wʌh-ə | وواههٔ |
| hit her/it | wə-wahəl-a | ووهله |
| hit us | wə-wahəl-u | ووهلو |
| hit you (pl.) | wə-wahəl-əj | ووهلئ |
| hit them (masc.) | wə-wahəl | ووهل |
| hit them (fem.) | wə-wahəl-e | ووهلي |

Table 9: Object agreement in the past tense (transitive)

change the pronunciation. It follows, however, that if the same letter is written twice in a row, then there must be an intervening vowel, as in «ممکن» [mʊmkɛn] 'possible'.[20]

- (Rare) In certain Arabic words, the letter «أ» makes the [a] sound, as in «تأکید» [takid] 'emphasis.' These words are often written with a plain «ا» [alɛf] instead, however, giving rise to words with an «ا» [alɛf] in the middle of the word that makes the [a] sound: «تاکید».

# 7 Tips for getting started

Probably the most daunting thing about learning to read Pashto is coping with the lack of vowel symbols. Earlier it was noted that the word «بند» could be read with a variety of combinations of vowels, other than the correct pronunciation, [band]. In practice, this ends up not being much of a difficulty. The reason is that most words are clearly differentiated by their consonants; vowels are less important. Indeed, even speakers of English commonly eliminate vowels from SMS messages, and the result is only marginally less intelligible. There are very, very few instances of Pashto words that can be confused because of the absence of written vowels.

It is important to keep in mind that a mature reader does not read letter-by-letter, but word-by-word.[21] That is, the shapes of whole words become familiar as one reads, so that the brain is decoding the whole word, rather than taking apart each letter. This happens naturally with practice, but a new reader lacks that practice. The initial effort must be made to laboriously pick apart the letters of each word, and then to find the meaning. It's undoubtedly tedious, but do keep in mind that it becomes automatic with practice.

With that said, new words are often encountered, and a reader needs skills to figure out what an unfamiliar word is. There is a heuristic rule that may be useful: guess [a]. That is, when ones encounter a sequence of letters must have a vowel, try first to insert an [a].[22] For some words, like «حرکت», guessing [a] produces the right pronunciation, [harakat]. For the word «شرکت», guessing [a] produces the incorrect [ʃarakat], but that might be close enough for you to recognize the word, [ʃɛrkat].

A related problem is knowing where to put the vowels. Even non-native speakers of Pashto have some intuition about what a possible Pashto word is. It is still occasionally necessary to guess, however. For the word «حرکت», for instance, [hɛrkat] is a possible Pashto word—for instance, «شرکت» is pronounced [ʃɛrkat]— but the correct pronunciation of «حرکت» turns out to be [harakat].

The easiest words to recognize are the words known already in the spoken language. With surprisingly little practice, these words are recognized automatically. The more difficult words are the ones learnt for the first time on the printed page. Unless one reads texts from entirely within one's sphere of previous experience, learning new words is part of learning to read. Obviously, broadening one's vocabulary is one of the prime motivators to learn to read: only, be aware that this is also the more laborious and time-consuming part of the reading process.

It's impossible to overstate the important of repetition and practice. A technique that the first author used in learning to read in this script was to read the same passage over and over again. While one might think that reading the same thing repeatedly wouldn't force one's brain to learn the symbols, in fact, the rote procedure of moving one's eyes over the words with understanding is just what is needed to firm up reading skills.

---

[20] An uncommon but systematic exception to this generalization is in compound words, where a double-letter can occur even if there is no intervening vowel. Example: «یادداښت» [jʌddʌçt] 'memo' from the words meaning 'memory' and 'to have'. (Mace 2003: 196)

[21] And even phrase-by-phrase!

[22] [a] is more than twice as common as [ɛ] and more than ten times as common as [o], so it makes more sense to guess [a] than anything else.

At the outset, it helps to adopt some practices that may seem childish: following the words with one's finger and reading aloud. These are helpful practices, though, which is why children use them. If possible, try to find printed text that uses a larger font size than would be used for Latin script. For whatever psychological reason, an unfamiliar script looks tiny even if it is printed in a reasonable font size.

A dictionary is necessary for looking up unfamiliar words, and for checking the spellings of words one knows from the spoken language. Several Pashto-English dictionaries are available.

The most useful for daily use is the *Pashto – English Dictionary* by Dr. Abdul Hamid Bahij (Danish Publishing Association, Peshawar, 2008). For looking up Pashto words starting from English, there is the *English – Pashto Dictionary* (Danish Publishing Association, Peshawar, 2007), and also *Bahij's English – Pashto Dictionary* by Dr. Abdul Hamid Bahij (Danish Publishing Association, Peshawar, 2009). More advanced students can use a Pashto-Pashto dictionary, such as the *Aryan Pashto – Pashto Dictionary* by Kamal Sind (Academy of Sciences of Afghanistan; Kabul), and the *Pashto – Pashto Descriptive Dictionary* (4 volumes; Academy of Sciences of Afghanistan, 2005). Robert Sampson's *A Dictionary of Spoken Pashto* (Peshawar: Interlit Foundation) is an English-to-Pashto dictionary; the words are given in both in Pashtu script and a phonetic alphabet.

## Acknowledgments & Bibliography

Ager, Simon. "Pashto." *Omniglot - writing systems and languages of the world.* Accessed 7th December 2015.
http://www.omniglot.com/writing/pashto.htm

Awde, Nicholas and Putros Samano. 2000. The Arabic Alphabet: How to Read & Write It. Lyle Stuart.

Olson, Randall. 2007. Speaking Afghan Pashto: The Eastern Afghan Dialect. Peshawar: Interlit Foundation.

## A    Learning the order of the Pashto alphabet

It isn't necessary to learn the order of the Pashto alphabet immediately. It is useful, however, for using dictionaries or other resources whose entries are organized alphabetically. Refer to Table 1 for order of the alphabet. This appendix provides a few mnemonic devices that can be used to learn the order more simply. One may take or leave any of these items; they are merely observations that the author has found helpful.

- The letters are grouped by their base shape. First comes «ا» [alɛf], then the letters shaped like «ب» [bɛ], then those shaped like «ج» [dʒim], then «د» [dʌl], then «ر» [rɛ], then «س» [sin], then «ص» [sʌd], then «ط» [tuj], then «ع» [ajn], then «ف» [fɛ], then «ک» [kʌf].

- Within these subgroups, the first letter of the group is often the unmodified version, with the following letters being the "decorated" versions: «س» [sin] precedes «ش» [ʃin]. (This doesn't hold for the ب-group or the ج-group, however.)

- The letters with a circle always come right after the equivalent letter without the circle.

- The «ا» [alɛf] – «ب» [bɛ] order is like the Latin a-b order (and like the Greek alpha-beta order, which itself gives us the word "alphabet"). If you happen to know the term *abjad*, this can be a mnemonic for the first four letter groups: «ا» [alɛf] – «ب» [bɛ] – «ج» [dʒim] – «د» [dʌl].

- The «ک» [kʌf] – «ګ» [gʌf] – «ل» [lʌm] – «م» [mim] –«ن» [nun] – «و» [waw] sequence is like the Latin k-l-m-n-o sequence. («ګ» [gʌf] intervenes, but it must since it is the more decorated form of «ک» [kʌf].)

## B    Learning the numbers

In addition to the letters of Pashto, the numbers need to be learned as well. These are shown in Table 10, along with mnemonic devices for recognizing the shapes. Somewhat surprisingly, numbers in Pashto are written just as they are in Latin script: from left-to-right. Thus, the number «۱۵» is 15, not 51. Note, however, that the overall order remains right-to-left. If someone were referring to a range of 15–20, it would be written «۲۰-۱۵».

The only other complication is that for the numbers 4, 5, and 6, («۴», «۵», «۶») there are alternate forms: respectively, «٤», «٥», «٦». These alternate forms are actually the Arabic varieties of these numbers. They are seen fairly commonly in Afghanistan, however.

| Latin-based | Pashto | Arabic | Mnemonic |
|---|---|---|---|
| 0 | ۰ | | |
| 1 | ۱ | | looks like a 1 |
| 2 | ۲ | | has two strokes |
| 3 | ۳ | | has three strokes |
| 4 | ۴ | ٤ | |
| 5 | ۵ | ٥ | |
| 6 | ۶ | ٦ | |
| 7 | ۷ | | looks like the 'v' in 'seven' |
| 8 | ۸ | | |
| 9 | ۹ | | looks like a 9 |

Table 10: Persian & Arabic numerals.

## C   The vowel accent markers

Pashto can be written with accents to indicate the presence of a vowel. A slanted line under a consonant (called [zɛr]) indicates the presence of the vowel [ɛ] following that consonant. For the word [kɛtʌb] 'book':

(20)   كِتاب

A slanted line *over* a consonant (called [zwar]) indicates the presence of the vowel [a], as in «بند» [band] 'shut' below:

(21)   بَند

The presence of a small diacritic called [peç], which has the shape of a small «و» [waw], represents the vowel [ʊ], as in [tʃʊp] 'shut':

(22)   چُپ

When the diacritics are followed by vowels, it is a diphthong, as in [mawzo] 'subject'.

(23)   مَوضوع

These diacritics are rarely used, and only in very specific situations: in very early elementary school materials, or written text when it is necessary to indicate the pronunciation of a word (e.g., if a word might be confused with another one).[23] The fact that these diacritics were not introduced earlier indicates the attitude of the author toward their use. Although they seem helpful, a new reader shouldn't depend on them. An important part of learning to read Pashto is to learn to identify a word based on the sounds that are written: this can't be learned if one is constantly relying on diacritics.

## D   Ideas for early reading

After the first primers have been completed, what should a new reader begin to read? Based on the principle that in learning one moves from what is known to what is unknown, the following progression may be suggested for a new reader.

**Known subject, known text.** At the very beginning, it is most helpful to read translated texts, since these provide maximum context, thus allowing a focus on the process of decoding rather than of discovering the meaning. These may be stories that are already familiar, or that are also available in English translation. The IAM Language & Culture Programme has a number of short (1–4 page) fairy and folk tales,

---

[23]In fact, the vowel diacritics are most commonly used in the Qur'an to indicate proper pronunciation of the Arabic words. Arabic is a separate language of course, and in fact the diacritics have different interpretations in Arabic.

which are available in English and written Pashto.[24] The organization Operation Mercy has published abridged translations of several Western classics (*Les Miserables, Tom Sawyer,* etc.), which are available for purchase from the Kabul office of Rahmat Publications, and in bookstores around the country. The BBC has published the well-known "BBC books," the plots of which can largely be inferred from the pictures; these can be purchased from the BBC's Kabul office, or from second-hand bookshops.[25]

**Known subject, unknown text.** The next step is to read books about familiar subjects, but in a unfamiliar text. This is a stage, for instance, in which one might wish to read books about one's professional interests. It is not difficult to visit a bookshop, mention a topic to the proprietor, and buy a large number of reasonably priced books pertinent to one's professional interests. Wikipedia articles written in Pashto are also a good resource.[26] One difficulty at this stage is that, depending on the book, the language might be very elevated, and the use of a dictionary might therefore be a constant necessity.

**Unknown subject, known text.** This category covers reading material that is available in Pashto and English, but which addresses subjects that the reader usually does not discuss in Pashto, or at all. This could be any material which can be found in English and Pashto translation. The military propaganda newspaper that usually comes with naan is printed in Pashto, Dari, and English, for instance.

**Unknown subject, unknown text.** This final category, of course, covers the rest of the printed material in the world. The most straightforward sources for new texts are newspapers, bookshops, and Afghan or Pakistani media available on the internet.

---

[24]Find contact information at: `http://iam-afghanistan.org/lcp`

[25]BBC Afghan Education Project, House 271, Street 1, Qalai Najarha, Khair Khana. A caution: the BBC books suffer from very poor typography—particularly in the spacing between non-connecting letters and words—which can make them very hard to read. Some books are also printed with a calligraphic style that is difficult for readers just learning the shapes of the letters. These are not ideal books for beginning readers, but they may be the best option for Pashto.

[26]`http://ps.wikipedia.org/`