



Urdu is one of many Indic languages written in the Perso-Arabic script. Urdu's ancestors emerged as literary languages during the time of the Delhi Sultanate (1211–1504) and grew to flourish as literature in the Deccan sultanates and at the Mughal court. Like Modern Urdu, these languages combined Indic linguistic structures and literary genres with Perso-Arabic vocabulary (and some grammatical borrowing) and literary genres. The word “Urdu,” which means “camp,” originally referred to the place where the language we now know as Urdu was spoken (the *Urdū-e Mu‘allā* or “high camp” of the Mughal capital in Delhi). The word was first used to name the language around 1775.

English readers unfamiliar with the Urdu writing system may find it intimidating at first. Three of the main challenges that beginners face are the rules of separation or combination, variability in the pronunciation of certain letters, and the ambiguities that result from the lack of short vowel markers. Fortunately, combination rules are easily mastered, and context or prior knowledge can often resolve ambiguities.

With regard to combination, there are two types of letters in the Urdu script. The first is sometimes called “friendly” because letters of this type connect to the letter that follows them. The second is called “unfriendly” because the letters in this group never connect to the letter that follows them.

***Alif Be* – The Urdu Alphabet**

Urdu learners are often surprised to learn that the Perso-Arabic alphabet shares a common ancestor with the Latin alphabet used to write English. This explains, in fact, why the names of the first letters in the Urdu alphabet, *alif* and *be*, so closely resemble the names of the English word “alphabet.” It’s not a coincidence. The English word comes from Greek, and the names of the first letters in the Greek alphabet, *alpha* and *beta*, are distant cousins of the names for the first two letters in the Urdu alphabet: ا (alif) and ب (be).

The Urdu alphabet can be divided into rows comprising what we will call “families” of letters. The letters in each family share a basic shape. Members of the same family are distinguished from each other by the dots, lines, or other markings that appear above or below the shape. The following chart surveys all the letters.



Urdu Script and Pronunciation

by Gregory M. Bruce

ا
ب پ ت ٹ ٹ
ج چ ح خ
د ڈ ذ
ر ژ ز
س ش
ص ض
ط ظ
ع غ
ف ق
ک گ
ل
م
ن ن
و
ہ ہ
ی ی
ء

Pronunciation

The letters in the Urdu alphabet can be divided into two types. The first are those whose pronunciation is invariable. The second is those whose pronunciation is variable. Let's learn the invariable letters first.

Our first family of invariable letters is the *be* series.

ٹ ٹ ت پ ب
s ʈ t p b

Notes:

The letter ب is pronounced just like an English b.

The letter پ [p] is an unaspirated letter, meaning that no air exits the mouth when it is articulated. This is very different from English, in which initial-position *ps* are always aspirated. Try it. Hold your hand in front of your mouth and say "pencil." You will feel a puff of air. Now say "sip." Notice that the *p* at the end of "sip" does not generate a puff of air. That is because it is unaspirated. In Urdu, the letter پ on its

Urdu Script and Pronunciation

by Gregory M. Bruce

own is always unaspirated. In fact, all letters are unaspirated unless immediately followed by the letter **ا**, which we will discuss later.

The letter **ت** [t] is an unaspirated dental consonant. Your tongue should be pressed up against the back of your upper teeth when you pronounce it. No air should come out of your mouth when you pronounce it letter. Check yourself by holding your palm in front of your mouth.

The letter **ٹ** [ʈ] is an unaspirated retroflex or palatal *t*. It is articulated by rounding the tongue backward, placing the tip of the tongue against the roof of the mouth, and pronouncing *t*. This is an unaspirated sound, so make sure that no air escapes your mouth. Practice distinguishing between **ت** and **ٹ** by saying them in combination with some vowels. Retroflex sounds are some of the most beautiful in Urdu. They are also instantly recognizable as markers of Indic-English accent. Keep this in mind as we proceed.

The letter **س** [s] is pronounced just like the *s* in English “some.” It is only found in words that are borrowed directly or indirectly from Arabic.

Our next family is the *jīm* series:

ج	چ	جھ	جج
kh	h	ch	j

Notes.

The letter **ج** [j] is pronounced just like the English *j* in “juice.”

The letter **چ** [ch] is an unaspirated *ch* as in the *ch* in “sandwich.” Be sure that no air escapes your mouth when you pronounce it.

The letter **ح** [h] is pronounced just like a breathy English *h* as in “horse.” It is only found in words borrowed directly or indirectly from Arabic.

The letter **خ** [kh] is a scratchy sound made in the throat. Linguists and philologists often identify it as a Perso-Arabic sound (as opposed to an Indic one). However, there are many Indic words in Urdu that are spelled and pronounced with it. For example, **چٹخارا** [chaṭṭkhārā] (the taste in one’s mouth), **پٹاخا** [paṭākhā] (firecracker), and so on.

The *dāl* series:

Urdu Script and Pronunciation

by Gregory M. Bruce

ذ	ڈ	د
z	ḍ	d

Notes. The letter د is a dental d and the voiced counterpart to ت. It is pronounced by pressing the tip of the tongue against the back of the upper teeth and vibrating the vocal cords.

The letter ڈ is the retroflex d and the voiced counterpart of ٹ. It is pronounced by bending the tongue backward and pressing it against the roof of the mouth, and vibrating the vocal cords.

The letter ذ [z] is pronounced just like the English z in “zap.” Some philologists have identified this letter exclusively with Arabic, but Persian words like پذیر [pazīr] have been spelled with it for centuries.

The *re* series.

ڑ	ز	ژ	ر
zh	z	ṛ	r

The letter ر [r] is a flapped r. The tip of the tongue flaps upward to make quick contact with the ridge behind the upper front teeth.

The letter ڑ [ṛ] is a retroflex flap. Rather than flapping upward, the tongue starts in a retroflex position (similar to the position for ٹ) and flaps forward and downward. Along the way, the underside of the tip of the tongue makes contact with the ridge behind the front upper teeth before the tongue comes to rest in a horizontal, flat position in the mouth.

The letter ز [z] is pronounced just like a z in English “zap.” This is considered a generic z in Urdu, which means that, in addition to being found in the Arabic and Persian words that use it, it is also how z sounds in English borrowings are spelled.

The letter ژ [zh] is a Persian letter, somewhat rare in Urdu, and is pronounced like the s in “television.”

The *sīn* family.

ش	س
sh	s

Urdu Script and Pronunciation

by Gregory M. Bruce

The letter س [s] is pronounced like an English *s* in “some.” This is considered the generic *s* in Urdu. It is therefore used to spell words borrowed from English, Sanskrit, and so on, in addition to the Arabic and Persian words that are spelled with it.

The letter ش [sh] is pronounced just like *sh* in the English “shop.”

The *svād* family.

ض	ص
z	s

The letter ص [s] is pronounced just like the English *s* in “some.” It is only found in words that are borrowed directly or indirectly from Arabic.

The letter ض [z] is pronounced just like the English *z* in “zap.” It is only found in words that are borrowed directly or indirectly from Arabic.

ظ	ط
z	t

The letter ط [t] is an unaspirated dental *t* and is identical in pronunciation to ت. With rare exception, it is only found in words that are borrowed directly or indirectly from Arabic.

The letter ظ [z] is pronounced just like the English *z* in “zap.” It is only found in words that are borrowed directly or indirectly from Arabic.

The letter *ghain*.

غ
gh

The letter غ [gh] is the voiced counterpart to the letter *kh*. Make the scratchy sound with your throat, but vibrate your vocal cords while you do it, and you have *gh*. Linguists and philologists have conventionally held that this is a Perso-Arabic sound not found in Indic words. However, there are many Indic words in Urdu that are spelled with it and pronounced accordingly—for example, غنڈہ [ghundā] (gangster).

Urdu Script and Pronunciation

by Gregory M. Bruce

Fe and qāf.

ق ف
q f

The letter ف [f] is pronounced just like the English *f* in “finance.”

The letter ق [q] in Modern Standard Urdu is pronounced similarly to the *k* in the English “park,” but rather than sliding the tongue upward into the back of the throat, one slides it straight back and slightly downward to block the air at the base of the throat instead. The result might be described as a fuller, thicker sound resembling a pop. Note, however, that many Urdu speakers in the northern part of South Asia pronounce this letter as an unaspirated *k* and thus do not distinguish it in speech from ک [k] (see below). In the southern part of India, in and around Hyderabad, however, the letter is often pronounced identically with خ [kħ] (see above). Hence, قیّمہ [qīmā] in Hyderabad is often called *khīmā*.

The *kāf* family.

گ ک
g k

The letter ک [k] is an unaspirated consonant pronounced just like the English *k* in “park.” Be sure not to allow any air to escape when you pronounce it.

The letter گ [g] is the voiced counterpart of ک [k] and is pronounced just like the English *g* in “go.”

ل

The Letter ل [l]. This is pronounced just like the *l* in “letter.” Note that, in American and British English, the pronunciation of *l* changes depending on its placement in a word. At the end of words—for example, in “pull”—the so-called “dark *l*” is pronounced using the back of the tongue. In Urdu, the pronunciation of *l* does not change. Whether at the beginning or end of a word, the Urdu ل [l] is always pronounced as is the *l* in “letter.”

The letter م [m].

م

Urdu Script and Pronunciation

by Gregory M. Bruce

m

The letter م [m] is pronounced just like the English *m*, as in “me.”

ن

n

The letter ن [n] is pronounced just like the English *n* in “no.”

Short Vowels and Diacritics

Urdu represents short vowel sounds using three diacritical markings. A short superscript diagonal marker called *zabar* (above) represents the short vowel *a* (pronounced like the *u* in the English “but”). It appears above ت in the following:

(then) [tab] تَب

The short vowel *i* (just like the *i* in the English “bit”) appears as a short subscript diagonal mark called *zer* (beneath). You can find it beneath ب in the following word:

(proven, established) [sābit] ثَابِت

The short vowel *u* (as in “put”) appears as a small superscript loop called *pesh* (front). You can find it above ب in the following word:

(idol; beloved) [but] بُت

Urdu represents the lack of a vowel with either a small superscript circle (called *sukūn*) or half circle (called *jazm*). You can find it in its half-circle form (*jazm*) above ح in the following word.

(discussion, argument) [bahs] بَحْث

One of the most challenging aspects of the Urdu script is that these short vowels are almost never written. Hence in the word مطلب [matlab] (aim; meaning), the two short *as* are nearly always omitted in writing.

Urdu Script and Pronunciation

by Gregory M. Bruce

Prior knowledge and some knowledge of word-building patterns in Urdu are the easiest ways to resolve ambiguity.

The pronunciation of the short vowels is generally regular. However, their pronunciation does vary according to regular rules when they precede and succeed certain letters. When a short vowel precedes a breathy *و* or *ح* [h] that is not succeeded by a vowel, the pronunciation of the short vowel changes according to a regular pattern. In these cases, *ا* [a] and *ی* [i] are both pronounced similar to the *ay* in “say” (and without the bending at the end of the English vowel), which we will represent as *e*. The short vowel *و* [u] is pronounced like the *o* in “open.” Here follow examples:

(gathering) [mehfil]	مخفل
(to say) [kehrnā]	کہنا
(honor) [ehtirām]	احترام
(present) [tohfā]	تُحفہ

The pronunciation of the short vowels similarly changes when in proximity to *ع*. For details, see the discussion of *ع* below. Further, in certain cases, the short vowel markers are not pronounced as short vowels at all, but instead determine the quality of certain mutable long vowels. For details, see the discussions of *و* and *ی* below.

The diacritical marker *ء* (*hamzah*) is used to mark the separation of two vowels of any length. When it precedes the vowel *ی* [ī], it is written on its own “seat” identical in shape to the *ب* series.

(barber) [nāī] نائی

Note that in the preceding example, the *hamzah* is required by spelling convention. It would be incorrect to write the word without it (the erroneous نای). The rule is simple. Any time two vowels appear in succession, a *hamzah* must appear between them.

Hamzah likewise marks the separation of a short vowel and a long one. The following example combines the short vowel syllable *ک* [ka] with the long vowel *ی* [ī]:

(several) [katī] کئی

Hamzah may also *carry* a short vowel in the middle of a word. In the following example, it carries the short vowel *i*, since, without the *hamzah*, there would not be a way to represent the presence of both

Urdu Script and Pronunciation

by Gregory M. Bruce

vowels. Note that the symbol ʾ in the transliteration is simply used to distinguish the two successive short vowels *aʾi* from the single long vowel *ai*.

(satisfied) [mutmaʾin] مطمئن

When *hamzah* separates a vowel preceding و [au, ū, o], it is written above و without a separate seat. In the following example, *hamzah* is essential because it distinguishes the two-vowel combination *āo* from the vowel-consonant pair *āv* [او].

(come!) [āo] أو

Tashdīd. The superscript diacritic ّ, called *tashdīd* (“intensification”), generates a doubled letter. For example, if we were to place *tashdīd* above the letter ن [n] in غنّه [ghunnā] (nasal) the effect is similar to the pronunciation of the two *ns* in “unnatural.” We don’t pronounce the two English letters separately, but as a single, “intensified” sound. Likewise, we “intensify” the single *n* and pronounce it as two. Here is another example using an *s* sound.

(part, share) [hissā] حصّه

Tanwīn. This letter literally means, “*nūn*ization,” which is to say that it generates the pronunciation of the Urdu letter *nūn* [n]. It is written as an *alif* with two short superscript diagonal diacritical marks identical to *zabar* (see above), and is pronounced *an* (like the *on* in “won”) and is used mainly to create adverbs. Hence, it is best translated or understood as analogous to the English suffix -ly.

(legally) [qānūnan]	قانوناً	(law) [qānūn]	قانون
(certainly) [yaqīnan]	يقيناً	(certainty) [yaqīn]	يقين

The Dependent Variable Letter ں

The letter ں (called نون غنه [nūn ghunnā]) represents nasalization—that is, when a vowel is pronounced by pressing air through the nose rather than through the mouth. By definition, ں cannot appear at the beginning of a word, since it requires a preceding vowel to nasalize. It is an extraordinarily important letter because it often marks the difference between singularity and plurality. We will represent it with the sign *n̄*.

Urdu Script and Pronunciation

by Gregory M. Bruce

When ۷ appears at the end of a word, it is written in its full form. In the following example, the long vowel *e* is nasalized.

[karen̄] کریں

When attached to the succeeding letter, ۷ is orthographically identical to ن [n].

[karen̄ge] کرینگے

When ۷ appears in the middle of a word, its pronunciation is variable. In such cases it follows a pattern very similar to English. Just as *n* in English changes depending on what follows it in words like “sand” and “song,” and may even appear as *m* before letters like *b* and *p* as in “compare,” so too ۷ changes according to the letter that succeeds it. Here are some examples.

(egg) [an̄ḍā] انڈا

(war) [jan̄g] جنگ

(wonder, amazement) [achan̄bhā=achamb̄hā] اچنبھا

In some cases, when ۷ appears in the middle of a word, it may be written separately from the succeeding letter even though it is pronounced as if it were connected to it.

(I will do/make) [karūn̄gā] کروں گا

The Dependent Invariable Letter ڀ

The letter ڀ, called *do-chashmī he* (two-eyed ڀ), is only used in Modern Standard Urdu to aspirate consonants. That is why, like ۷, it is unpronounceable without a preceding letter. When ڀ follows a letter, the combination of the two is articulated as a single, aspirated sound. Consider the following pair. Note that both the single letter and the two-letter compound are pronounced as a single sound.

[bh] ڀ

[b] ب

In Urdu and other Indic languages, aspiration has semantic consequences. The following pair shows that the difference of a single aspirated letter can change the meaning of a word entirely.

Urdu Script and Pronunciation

by Gregory M. Bruce

(plate) [thālī] تہالی (clapping) [tālī] تالی

Do *chashmī he* (آ) and *gol* or *chhoṭī he* (و) are, historically, two styles of writing the same letter. Indeed, up until the early twentieth century, they were used interchangeably. However, in Modern Standard Urdu, they are clearly and universally distinguished from one another. Do not confuse them. Note the difference between the following pairs in pronunciation and meaning.

(the river flowed) [nadī bahī] ندی بہی (the river, too) [nadī bhī] ندی بھی

As the preceding example shows, the distinction between aspirated and unaspirated letters has semantic implications in Indic languages. The same is not true in Arabic, English, or Persian, for example, even though these languages have aspirated letters. The following chart surveys the letters in Urdu that can be combined with *h* to produce aspirated consonants by listing them in combination.

[th] تھ	[tʰ] ٹھ	[ph] پھ	[bh] بھ
		[chh] چھ	[jh] جھ
		[dh] ڈھ	[dh] دھ
			[rh] رھ
		[gh] گھ	[kh] کھ

Note that in some cases, the letter *و* or *ن*, too, may be followed by *h* and thus pronounced as an aspirated letter. For example, in the word *منہ* [munh] (face).

The aspiration in all these compounds is much stronger than the aspiration of consonants in English. Hold your palm in front of your mouth while you say the English word “pull.” Speakers of Standard American and British English will feel a puff of air against their palms. Now, keeping your hand in front of your mouth, pronounce the Urdu word *پھل* [pʰal] (fruit). The aspiration of the Urdu should be noticeably stronger than the English—that is, the air should be more forceful and there should be more of it. Finally, keeping your hand where it is, pronounce *پل* [pal] (moment). You should feel no air (or hardly any air) on your palm. Practice going back and forth among these words to appreciate the distinctions between aspirated and unaspirated consonants.

Variable letters.

Unlike the invariable letters surveyed above, the pronunciation of the following letters depends on their position in a word or the quality of the vowels around them.

Urdu Script and Pronunciation

by Gregory M. Bruce

Alif. When in the initial position, this letter may carry any one of the short vowel markers and is pronounced accordingly. For example:

(now) [ab]	أب
(over here) [idhar]	إدھر
(over there) [udhar]	أُدھر

When *alif* appears in the middle or at the end of a word, it is pronounced as a long ā (as in “father”).

(gate; chapter) [bāb]	باب
(with) [bā]	با

To produce the same long ā (as in “father”) at the beginning of a word, a curved diacritical mark is placed on top of *alif*. The resulting letter, *ā*, is called *alif madd*.

(water) [āb]	آب
--------------	----

Alif madd only very rarely appears in the middle of words. In those cases, it is pronounced by most speakers as a long ā, but technically indicates the presence of a short vowel preceding the long ā:

(center, repository) [ma'āb]	مآب
------------------------------	-----

A final variation of the letter *alif* involves suspending a superscript ^۱ over either *۱* or *۲*—for example, *۱* or *۲*. This letter is called *الف مقصوره* [alif maqsūrā] or “dagger alif,” and is only found in Arabic words. Up until the mid-twentieth century, *۲* was the preferred shape, but in recent decades the more Arabicized *۱* has become increasingly preferred.

(lit. the chosen one; Mustafa) [mustafā]	مصطفیٰ
(lower; lowest) [adnā]	ادنیٰ

The Letter *۱* (vāv). When this letter is followed by a vowel, it is pronounced like an English w.

(back, return) [vāpas]	واپس
(invitation) [bulāvā]	بلاوا

Urdu Script and Pronunciation

by Gregory M. Bruce

In medial or final position, و may have one of four pronunciations. What determines its pronunciation is the short vowel (or lack thereof) that precedes it. If it is preceded by َ(zabar), it is pronounced *au* as in “aural.” If it is preceded by no vowel, it is pronounced *o* as in “open.” If it is preceded by ُ[u] it is pronounced as a long *ū* as in the “oo” in “loose.”

(desire, passion, interest) [shauq]	شوق
(hat) [ṭopī]	ٹوپی
(shoe) [jūtā]	جوتا

If و appears at the end of a word following a long vowel, and there is no *hamzah* above it, then it is pronounced و. Compare the following.

(digging, searching) [kāv kāv]	کاوکاو
(eat!) [khāo]	کھاو

For و to represent a vowel sound at the beginning of words, an *alif* must be placed before it. The *alif* then carries the short vowel that determines the quality of the و. Consider the following:

(and) [aur]	اور
(cover, screen) [ot]	اوت
(above, on) [ūpar]	اوپر

The Letters ی (*chhoṭī ye*) and ے (*barī ye*)

These letters are, as their names suggest, two forms of the same letter, *ye*. Historically, both forms of *ye* were used interchangeably to represent the same set of sounds. Since the standardization of Urdu in the early 20th century, the two letters have come to represent distinct vowel sounds in their independent and final forms. In initial and medial position, however, they have the same shape.

When *ye* appears in initial or medial position, it has the shape of the ب series letters and is distinguished by two subscript dots. For example, the middle letter in بیت [bait or bīt].

When *ye* is succeeded by a vowel of any length, it is pronounced *y* as in “yes.”

(or) [yā]	یا
-----------	----

Urdu Script and Pronunciation

by Gregory M. Bruce

When *ye* appears in the medial position and is not succeeded by a vowel, it has three possible pronunciations. If preceded by \acute{o} [a], it is pronounced similarly to the *a* in “apple.” We will represent this with the combination *ai*. When preceded by no vowel, it is pronounced like the *ay* in “way.” We will represent this as *e*. And when it is preceded by و [i], it is pronounced just like the *ea* in “easy.” We will represent this as \bar{i} . Consider the following.

(house; couplet) [bait]	بیت
(apple) [seb]	سبب
(middle) [bāich]	بیچ

In their final forms, *ye* takes two different shapes. The letter ے at the end of a word is always pronounced \bar{i} . The letter ی in the final position may either be pronounced *ai* or *e* depending on the presence of the short vowel \acute{o} or the lack of a short vowel on the preceding letter.

(hat) [ṭopī]	ٹوپی
(before, in front) [āge]	آگے
(tune, melody) [lai]	لے

For *ye* to be pronounced as a vowel at the beginning of a word, we must place an *alif* before it, just as we did before و . The short vowel (or lack thereof) on *alif* determines the quality of the long vowel.

(like this) [aisā]	ایسا
(one) [ek]	ایک
(brick) [īnṭ]	اینٹ

The Letter ہ (*gol he* or *chhoṭī he*). In most positions and cases, this letter is pronounced just like the breathy *h* in “home.”

(to be) [honā]	ہونا
(world) [jahān]	جہان
(group) [guroh]	گروہ

However, when the letter appears at the end of a word and is preceded by the short vowel \acute{o} [a], its pronunciation changes to \bar{a} as in “father.”

(tabla) [tablā]	ٹبلہ
-----------------	------

Urdu Script and Pronunciation

by Gregory M. Bruce

(taste, flavor; pleasure, delight) [mazā] مزہ

The Letter ع (‘ain). This letter is only found in words that are borrowed directly or indirectly from Arabic. Its pronunciation is perhaps the most variable and complicated of any letter in the Urdu language.

In initial position, the letter is pronounced just like ا (alif).

(court) [adālat] عدالت
(respect) [izzat] عزت
(excellent) [umdā] عمدہ

When ع precedes a long vowel in initial position, ع likewise follows the same rules as *alif*. That is, it has no pronunciation and simply serves to “carry” the long vowel phonetically. Note that in the first case, the rule followed is similar to that of *alif madd* ا [ā].

(general, common) [ām] عام
(woman) [aurat] عورت
(lute) [ūd] غود
(glasses) [ainak] عینک
(Eid) [īd] عید

Despite the complexities of its pronunciation, ع can also give us clues about the quality of the long vowel that follows it. This is because it must carry a short vowel (ا [a] اِ [i] or اُ [u]) and, unlike ا, cannot carry *jazm* or *sukūn* (no vowel) in initial position. Hence, the combination of ain and *ye* can never yield the sound *e*, and likewise the combination of ع and *o* cannot yield *o*.

In medial position, the pronunciation of ع depends largely on the vowels that precede and succeed it. If it appears between two vowels of any length, then it has no pronunciation. Instead, it acts like a bridge between the vowels similar to *hamzah* and is elided.

(consciousness) [shuūr] شعور
(Shia) [shīā] شیعہ

Urdu Script and Pronunciation

by Gregory M. Bruce

When ɛ is preceded by a short vowel but not succeeded by any vowel, the quality of the short vowel changes and ɛ is pronounced as follows. When preceded by َ [a], the a is elongated and the ɛ is pronounced ā (as in “father”). When preceded by ِ [i], the i is elongated to e (as in “say”). When preceded by ُ [u], the u is elongated to o (as in “open”).

(education) [tālīm]	تَعْلِيم
(trust, confidence) [etibār]	اعتبار
(length) [bod]	بُعد

Note that in all three of the above cases, ɛ is not followed by a vowel. The same rule, as mentioned, holds when ɛ appears at the end of a word.

(crowd) [majmā]	مَجْمَع
(means, media) [zarāe]	ذرائع
(expectation) [tavaqqo]	تَوْقَع

When ɛ is preceded by a long vowel, but not succeed by a vowel, it is not pronounced. This is most common when it appears at the end of a word.

(beginning) [shurū]	شروع
---------------------	------

Sometimes, ɛ appears at the end of a word and is not preceded by a vowel. In such cases, it is typically pronounced ā.

(addition) [jamā]	جَمْع
(candle) [shamā]	شَمْع

When words in which ɛ appears at the end of a word and is not preceded by a vowel are the first part of a Persian اضافت [izāfat] construction, ɛ loses its vowel pronunciation (ā). Instead, the preceding consonant is doubled:

(the candle of the gathering) [shamm-e mehfil]	شَمْعِ مَحْفَلِ
--	-----------------

Letter Combinations and Word Formation

Urdu Script and Pronunciation

by Gregory M. Bruce

The letters in the Urdu alphabet can be divided into two categories: those that connect to the letter that follows them and those that do not. We will call connecting letters “friendly” and non-connecting letters “unfriendly.” Don’t be confused. All letters in Urdu can be run into by the preceding letter. When we talk about “unfriendly” letters, we are only referring to the letters that do not connect to the letter that *follows* them.

Friendly Letters

Most letters in the Urdu alphabet are friendly, that is, they connect to the letter that follows them. Friendly letters may have different shapes in their initial, medial, final, and independent forms. Here is a chart of all the friendly letters in their independent forms.

ب پ ت ث
ج چ ح خ
س ش
ص ض
ط ظ
ع غ
ف ق
ک گ
ل
م
ن ن
ہ ہ
ی ی
ء

In the initial and medial positions, most friendly letters are shortened by dropping the final $\frac{2}{3}$ or $\frac{3}{4}$ of their independent shape. The shape of most friendly letters in final position (that is, at the end of a word) is identical to their independent forms. Compare the shortened length of ت in its initial position shape with the final-position shape of ب in the word تب [tab] (then).

ت + ب = تب

The shape of friendly letters in their connecting form (initial and medial position) will be the same for all the letters in a given family. The only difference among them will be the number and position of the dots.

Urdu Script and Pronunciation

by Gregory M. Bruce

The following examples illustrate the shape of each friendly family in the initial, medial, and final positions.

The *be* family

(then) [tab] تب
(affixing, stamping) [sabt] ثبت

The *jīm* family

(when) [jab] جب
(savings) [bachat] بچت
(ice) [yakh] یخ

Note that the letter preceding any letter in the *jīm* family connects at the left corner of the *jīm* series letter, creating a zig-zag like pattern (see یخ and بحث above).

The *sīn* family.

(all) [sab] سب
(eye) [chashm] چشم
(enough) [bas] بس

Note that *sīn* and *shīn* may also have an extended, smooth shape in addition to the toothed shape above. Although rare in typewritten texts, this extended form is commonly used when two members of the family appear in succession. For example:

(inexpensive) [sastā] سستا

The *svād* family. These letters lose their final ۛ shape in the initial and medial positions.

(face, form) [sūrat] صورت
(beautiful) [khūbsūrat] خوبصورت
(some) [bāz] بعض

The *toe* series. These letters do not change in the initial, medial, or final positions.

Urdu Script and Pronunciation

by Gregory M. Bruce

(demand, request) [talab] طلب

(moment) [lehzā] لحظه

(letter) [khatt] خط

The *‘ain* series. In the initial position, ع loses the final 2/3 of its shape. In the medial and final positions, it is written as a triangular shape above the line. Note that this is a rare case in which the final-position shape of a letter differs from its independent form.

(respect) [izzat] عزت

(side, armpit) [baḡhal] بغل

(gathered) [jamā] جمع

The *fe* series. The shape of *fe* and *qāf* is the same in initial and medial position, but they differ in appearance in their final and independent forms (*fe* is horizontal and *qāf* is deep and rounded).

(only) [faqat] فقط

(lock) [qufl] قفل

(rank, row) [saff] صف

(research) [tehqīq] تحقیق

The *kāf* family. The shape of both letters is identical in all positions. However, when *kāf* and *gāf* are followed by *alif* or *lām*, their typically long neck changes to a small circle, as in گلاب and کا.

(when?) [kab] کب

(rose) [gulāb] گلاب

(of) [kā] کا

(kind, ethical) [nek] نیک

Lām. This letter drops its ۷ ending in initial and medial positions.

(lassi) [lassī] لسی

(demand, request) [talab] طلب

(solution) [hall] حل

Urdu Script and Pronunciation

by Gregory M. Bruce

Mīm. The precise shape of this letter in initial position depends on the shape of the succeeding letter. In all cases, however, it appears as a black dot. In the medial position, it will always appear as a black dot beneath the line. Likewise, in the final position, the dot that begins the letter in independent form will appear beneath the line (e.g., in ہم [ham]).

(love) [mahabbat/muhabbat] محبت

(mom, mum) [mān] ماں

(long, tall) [lambā] لمبا

(body) [jism] جسم

Nūn. In the initial and medial positions, this letter has the same shape as the *be* family.

(kind) [nek] نیک

(idol; beloved) [sanam] صنم

(passion, love, enthusiasm) [lagan] لگن

He. *Gol* or *chhoṭī he* in medial position takes a subscript diacritical mark resembling an inverted *pesh* (the name of the letter used to mark the short vowel *u*). The shape of *gol* or *chhoṭī he* changes depending on the letter that follows it. If it is succeeded by an upward stroke (e.g., the letters *alif*, *lām*, or *dāl*) its shape resembles the *be* series (e.g., ہل [hal]). Otherwise, it takes the shape of a small tooth (e.g., ہر [har]) or a downward-sloping curve (e.g., مہک [mehek]). In final position when it is run into by a preceding letter, its shape differs from its independent form. In such cases, it appears as a wavy line (as in حصہ [hissā]).

(plow) [hal] ہل

(every) [har] ہر

(fragrance) [mehek] مہک

(part, share) [hissā] حصہ

(gaze, sight) [nigāh] نگاہ

Do-chashmī he can never appear at the beginning of a word since it only serves to aspirate the preceding letter. In medial and final position, its shape is identical to its independent form.

(also) [bhī] بھی

(understanding) [samajh] سمجھ

Urdu Script and Pronunciation

by Gregory M. Bruce

Ye. In initial and medial position, *ye* takes the form of the *be* series and has two subscript dots.

(this) [ye] یہ

(glasses) [ain] عینک

In final positions, Modern Standard Urdu distinguishes between *chhoṭī* and *baṛī ye*.

(lassi) [lassī] لسی

(children) [bachche] بچے

(fixed, decided; covered) [tai] طے

Hamzah. This letter is never written in initial position, though technically any time a vowel appears at the beginning of a word, there is an implicit *hamzah*, since the glottal stop represented by *hamzah* is required to pronounce the vowel. Hence, in ا there is an implied hamzah above ا. In medial position, *hamzah* is typically written on top of a seat identical in shape to the *be* series. In this form, it may either serve simply to mark the separation of two vowels, as in the following:

(barber) [nāī] نائی

(Gabriel) [jibraīl] جبرائیل

When *hamzah* marks the separation of a vowel from *baṛī ye* or *vāv*, it is simply written on top of the latter without its own seat.

(boat) [nāo] ناو

(chai) [chāe] چائے

Hamzah may also carry a short vowel when the vowel is second in a series of vowels.

(satisfied) [mutma'in] مطمئن

(permissible) [jāiz] جائز

The Unfriendly Letters

Four letter families are called “unfriendly” because they do not connect to the letters that follow them. Fortunately, it is easy to remember which letter families they are, since all four appear in the word اردو (Urdu). But don’t be confused. All letters in the Urdu script can be run into by the preceding letter. What we mean when we say that that these four letter families are “unfriendly” is that they do not connect to the letter that *follows* them.

Urdu Script and Pronunciation

by Gregory M. Bruce

Alif. The shape of *alif* in the initial, medial, and final positions is the same. Note the space that appears after *alif*. This is because it is an unfriendly letter.

(now) [ab] اب

(water) [āb] آب

(portal, gate; chapter) [bāb] باب

(with) [bā] با

The *dāl* family. Note that when run into by a preceding letter, the shape of the *dāl* family changes by reclining slightly. Hence, the independent form ذ takes a slightly more reclined shape in جذبہ [jazbā] and تشدید [tashdīd]).

(style) [ḍhanḡ] ڈھنگ

(emotion) [jazbā] جذبہ

(aim, purpose) [maqṣad] مقصد

The *re* family. The shape of *re* in its independent and initial form is similar to that of *dāl* in its medial and final form. The difference is that when *re* is in the medial or final position, it always slopes downward, whereas *dāl* requires an upward stroke.

(lord) [rab] رب

(writhing) [taṛap] تڑپ

(dear) [azīz] عزیز

Vāv. The shape of this letter does not change, but, as noted above, its pronunciation changes depending on its position in a word. When succeeded by a vowel, it is pronounced like an English w. When preceded by a short vowel and not succeeded by a vowel, it is pronounced one of three ways (*au*, *o*, or *ū*).

(epidemic) [vabā] وبا

(above) [ūpar] اوپر

(one hundred) [sau] سو

Urdu Script and Pronunciation
by Gregory M. Bruce