

CHAPTER EIGHT

PERSIAN AND TAJIK

Gernot Windfuhr and John R. Perry

1 INTRODUCTION

1.1 Overview

The focus of this chapter is Modern Standard Persian and Modern Standard Tajik. Both evolved from Early New Persian. Western Persian has typologically shifted differently from modern Tajik which has retained a considerable number of Early Eastern Persian features, on the one hand, and has also assimilated a strong typologically Turkic component, on the other hand. In spite of their divergence, both languages continue to share much of their underlying features, and are discussed jointly in this chapter.

1.1.1 Historical background

Persian has been the dominant language of Iranian lands and adjacent regions for over a millennium. From the tenth century onward it was the language of literary culture, as well the *lingua franca* in large parts of West, South, and Central Asia until the mid-nineteenth century. It began with the political domination of these areas by Persian-speaking dynasties, first the Achaemenids (c. 558–330 BCE), then the Sassanids (224–651 CE), along with their complex political-cultural and ideological Perso-Iranianate constructs, and the establishment of Persian-speaking colonies throughout the empires and beyond. The advent of Islam (since 651 CE) represents a crucial shift in the history of Iran and thus of Persian. It resulted in the emergence of a double-focused Perso-Islamic construct, in which, after Arabic in the first Islamic centuries, Persian reasserted itself as the dominant high register linguistic medium, and extended its dominance into formerly non-Persian and non-Iranian-speaking territories in the East and Central Asia.

The writing system became that of the new dominant religion, and there occurred increasing infusion of Arabic features into the lexicon, phonology and grammar (comparable to the absorption of the Norman component into English). However, throughout the evolution of the literary standards from Early New Persian to Modern Standard Persian the considerable typological changes that Persian underwent are due to both internal Persian developments, including the leveling of regional features, and to the assimilation of expanding areal cross-linguistic typological isoglosses.

1.1.2 Persian and related groups

See Chapter 2, Map 2.6.

1.1.2.1 Varieties of Persian

Overall, Persian varieties are divided into a Western group mainly in Iran and an Eastern group in Afghanistan and Central Asia, with transitional varieties.

Caucasus

The northwestern outpost of Persian is Caucasian Tat Persian spoken in an Azeri Turkic, Caucasian and Armenian environment, with three varieties: (a) Muslim and (b) Jewish (Juhuri) in Azerbaijan and Dagestan, and (c) Christian Armeno-Tat in Armenia.

Iran

(1) Persian sociolinguistic registers include:

- (a) Modern Standard Persian, the written norm in Iran (Farsi) and Afghanistan (Dari), evolved during the last few centuries;
- (b) Colloquial Persian, specifically the normalized form of Colloquial Tehrani Persian, used for most polite spoken communication, which increasingly shows reflexes in the standard language;
- (c) *Xodemuni* ‘our own’, (< *xod-emān-i*), i.e. familiar speech, the non-normalized local variant such as in Tehran.

(2) Regional and local varieties in the urban centers throughout Iran in non-Persian dialect and language areas.

(3) Khorāsān Persian varieties, representing a major distinct regional subgroup and stretching from east of Tehran to the Afghan border. Tehran to the Afghan border.

Transitional group

Varieties straddling the Iran-Afghan border: (1) closely related Kohistāni in Iran and Afghan Fārsiwāni, and (2) Sistāni on both sides of the border.

Afghanistan

See Kieffer 1983, incl. map and detailed table.

- (1) Afghan Persian, officially called Dari, mostly close to literary Persian.
- (2) Kaboli (Kāboli), increasingly become the standard Afghan vernacular.
Large regional Persian varieties include:
- (3) Herati (Herāti) near the Iranian border.
- (4) Ayamaqi (Aymāqi) near Herat (in pockets also in Iranian Khorasan).
- (5) Hazaragi (Hazāragi), stretching northeast through the Afghan center, some in Iran.
- (6) Afghan Tajik in the NE, including Afghan Badakhshani, the latter with the outpost Madaglashti near Peshawar in Pakistan.

Note that the terms “Tajik” as well as “Dari” are sometimes erroneously used collectively for all Afghan varieties.

Central Asia

Tajik, in an Uzbek Turkic environment, with four broadly defined groups (see 1.2 below).

1.1.2.2 Related varieties and dialect groups

Colonial Persian

Persian was cultivated at the courts of the Anatolian Seljuk and Ottoman rulers (from ca. 1200–1922), several of whom are known for composing Persian poetry. Probably best known among their protégés is Rumi (d. 1273), the most cherished Persian mystic poet who had come to Konya from Wakhsh near Balkh in Afghanistan. Literary Ottoman Turkish is a virtual amalgam of Turkish and Persian (with all of the latter's Arabic loan elements).

In the East, Urdu developed under heavy Persian influence. Persian first entered India with the conquest of north-west India by Ghaznavid armies in the eleventh century. Four centuries later, Persian was chosen as the court language of the Mogul rulers (1530–1857), who were major patrons of Persian literature and poets from Iran, unlike the contemporary Safavids in Iran.

It was at the courts of India and Turkey where many of the major traditional dictionaries of Persian were compiled from the fifteenth to the eighteenth centuries, including grammatical treatises. Simultaneously, there developed in India a Persian vernacular, and it was from the Indian scribes and secretaries that the English officers of the East India Company, a number of whom wrote grammars of Persian, learned their Persian, with all its local idiosyncrasies. Persian was abolished in its last official bastion—the courts of law—in 1837 by the authorities of the East India Company.

Judeo-Persian

Judeo-Persian varieties, like other Judeo-Iranian varieties, are found throughout Iranian-speaking regions, such as Jewish Tat Persian and Bukhara Tajik. In fact, Jewish merchants and travellers have been the earliest speakers who wrote in Early New Persian, and left its earliest documents. These and later documents, in a Hebrew-writing tradition parallel to that in Arabic, have been crucial for the study of the diachrony and diatopy of Persian (e.g. Lazard 1968; Paul, ed., 2003a).

Luri-type and Fars dialects

In SW Iran there are two groups which can be recognized as “Perside”, i.e. they continue numerous features that evolved from Southern Early New Persian (see 7 Diachrony), though each evolved differently:

- (1) The Luri-type dialects (Luri proper, Bakhtiāri, Boyer-Ahmadi, Mamasani-Kohgeluye).
- (2) The Fars dialects stretching from the Gulf into western and central Fars.

1.1.2.3 Number of speakers

According to various sources, the totals of native speakers of Persian and its varieties (or rather ‘ethnic’ Persians) in the three Persian-speaking countries vs. the total population (who at least use or understand Persian as a second language) are: Iran 35/70 m; Afghanistan 17/33 m (Hazāra 2.8 m; Aymāq 380,000); Tajikistan 5~6/7.2 m. Overall, the numbers are, counting speakers outside these countries, a total of 60/110 m, and of these: 35/70 m for Western Persian vs. 23/40 m for Eastern Persian. The estimates for Caucasian Tat suggest about 26,000, now reduced from a much larger number due to (partially forced) assimilation and emigration.

1.1.3 Evolution of Standard Persian

1.1.3.1 Early linguistic groups

Sources from the ninth–eleventh centuries distinguished four linguistic groups:

- (1) Pārsi, the literary Middle Persian (mainly used by Zoroastrian priests).
- (2) Pārsi proper (*mutlaq*), the literary and spoken Southern Early New Persian used from Fars to Sistan, which had retained numerous Middle Persian features and vocabulary, with relatively few Arabic loans.
- (3) In the North, (*Pārsi-i*) *Dari*: The term originally referred to the administrative and spoken Persian that had developed at the Sassanian court in Ctesiphon and was administratively used throughout the Empire. In Khorasan it had already replaced the local Parthian dialects even before Islamization. As a ready vehicle for the Muslim administrations, besides Arabic, it became the vehicle of the Muslim mission into Central Asia and beyond, where a Persian variety had already been used as a *lingua franca*. This process not only led to the ultimate replacement of Sogdian, Bactrian, and Khwarezmian, but also inserted into this Dari an increasingly larger Arabic loan component as well as local eastern Persian and other Iranian vocabulary.
- (4) Pahlavi, Ar. *fahlavi*, lit. ‘Parthian’: The term implied the non-Persian Iranian languages, particularly in western and central Iran which was once part of the former Parthian Empire.

In addition, Early New Persian varieties must also have been spoken in Sassanian border garrisons east and west. One of these was probably the ancestor of Tat Persian in the Sassanian outpost at the Caspian gate to the Caucasus, Darband. The others, probably older, were the outpost in Central Asia out of which ultimately developed Afghan and Tajik Persian.

See also the detailed study by Perry (2009).

1.1.3.2 Arabic and Turkic

See Chapter 2, Map 2.7.

(1) Arabic and Turkic

A major contributing factor to the Arabization of Persian was the magnitude of Arab settlement in Greater Iran, and presumably the intense interaction and intermarriage between the immigrants with the local populations at both the highest and lowest social levels, and after large numbers of Arab tribes moved into the Fertile Crescent during Sassanian times, Shapur II (r. 309–379) settled some of them in Fars as well as in the hinterland of Bam and Kerman. After Islam, settlement occurred in various waves throughout, and was most extensive in eastern Iran, including Khorasan, Afghanistan, and Central Asia. At its height, the number of Arab immigrants may have totaled 250,000. While these Arabic-speaking populations were ultimately absorbed, except for isolated Central Asian Arabic pockets (eastern Iran, northern Afghanistan, Central Uzbekistan), Arabic continued as the high register literary language during the earlier centuries of New Persian, mainly as the dominant language of science and religion, and may have at least indirectly affected even syntax, particularly through extensive translation activities both from and into Arabic.

In contrast, the immigration of Turkic speakers has led to the Turkification of large regions, mainly in Azerbaijan (Azeri Turkic) and less so in Fars (Qashqā’i) and Khorasan (Turkmen). In addition, dynasties of Turkic or Turko-Mongol origin ruled in Iran from

the late tenth century to the twentieth century. Linguistically, there has been considerable Iranization of Turkic, but the intense Turko-Iranian symbiosis has also effected grammatical innovations in Persian, including the emergence of the evidential category in the verb system during the nineteenth century (for this section, see also Windfuhr 2006).

(2) Shift of cultural centers

Early New Persian Dari became the vehicle for the emerging New Persian literature. Even before the Mongol invasion in the thirteenth century, the literary center shifted westward, prominently to the city of Shiraz in Fars. This extended process led not only to the elimination of the Eastern linguistic shibboleths, but also to the demise of the Early Southern standard. With the rise of the Safavids at the beginning of the sixteenth century, the center shifted north to Turko-Persian bilingual Esfahan, and finally with the rise of the Qajars to Tehran at the beginning of the nineteenth century.

In overview, the sociolinguistic set-up from Late Sassanian times may be shown as follows: (see also section 8 Diachrony and dialectology)

	<i>Iranians</i>	<i>Persians/Non-Iranians</i>
	High	Low/Lingua franca
Late Sasanian	Middle Persian	Dari
Early Islamic	Arabic, Dari	Dari
Mongols, 13th cent.	Persian	Persian
Safavids, 16th cent.	Persian	Persian Turkic
Qajars, 19th cent.	Persian	Persian

1.2 Tajik

Tajik, or Tajik Persian (*zabon-i tojikī*, *zabon-i forsi-i tojik*) is a variety of New Persian used in Tajikistan and parts of Uzbekistan, including the cities of Bukhara and Samarkand. Since the 1920s it has been fostered as the national and literary language of the Soviet Socialist Republic (from 1991, the independent state) of Tajikistan. The Tajik and Iranian Persian speech areas are not contiguous, but lie at opposite ends of a continuum with Persian dialects of Afghanistan in between, and interrupted by areas of Turkic (Turkmen and Uzbek) speech. Spoken Tajik has been evolving independently of Persian of Iran since the sixteenth century, but the written language (which functioned as the common language of high culture, government, and diplomacy in Central Asia and India) maintained a virtually universal standard, based on Classical Persian, until the early decades of the twentieth century.

In the Soviet period, with the promotion of a more vernacular style and lexicon and the systematic introduction of Russian loanwords, language change was more rapid. The writing system was changed from Arabic to Latin in 1928, then to Cyrillic in 1939. The modern literary language (*zabon-i adabi-i hozira-i tojik*, as planned and exemplified by Soviet Tajik writers of the 1920s and 1930s, notably Sadriiddin Aini) was based loosely on the style of the old cultural center, Bukhara, and contained many Uzbek loanwords and some syntactic structures calqued on Uzbek usage. In recent decades writers from different regions of Tajikistan, and some who have traveled abroad, have introduced a more varied style, including features closer to the Persian of Iran.

Tajik dialects may be divided broadly into two groups: North-western and South-eastern, corresponding in rough topographical terms to the lowlands and highlands respectively of the Oxus basin. Several refinements of this scheme have been proposed, and much fieldwork remains to be done. The most comprehensive classification posits

four groups: (1) Northern, comprising Bukhara, Samarkand and Derbend in Uzbekistan, the Ferghana Valley (including Khujand, former Leninabad), and extending down the Varzob valley to the region of the capital, Dushanbe. (2) Central, comprising the upper Zarafshon (Zeravshan) valley. (3) Southern, stretching south and east of the capital, in Kulob and Qarotigin (Karategin) regions, including Gharm, as far as the Pamirs (Mountain Badakhshan). (4) South-eastern, in Darvoz region, on the bend of the Oxus near Rushan.

Dialects of the last two groups have strong affinities with those of the left bank of the upper Oxus in Afghanistan's province of Badakhshan. They include Tajik as the contact vernacular (called *forsi*) of Gorno-Badakhshan, extending into Afghanistan and Wakhan. In these regions the native language of a majority of the population is one of the Iranian dialects of the Pamir group (see Chapter 14a).

Northern dialects have been influenced to varying degrees by Uzbek, with which there is widespread bilingualism. A distinctive sub-variety of Northern Tajik speech, with a modest literature, is that of the Jews of Bukhara, most of whom have emigrated. Dushanbe was a small market town before its promotion to national capital in 1924 and the consequent influx of Tajiks from elsewhere in the region and of Russians and other Soviet nationalities. Though nominally included in the Northern dialect area, it is again in a state of demographic fluctuation since independence, and is best considered separately. There are approximately five million Tajik Persian speakers in Central Asia and Afghanistan, including minorities in Uzbekistan and Kirgizstan. The so-called Tajiks of south-west Xinjiang, in and around Taxkorgan (Tashqurghon) – an official nationality in China – are speakers of Pamir languages (mainly Sarikoli and Wakhi), not Persian.

The variety of Tajik described here is basically the ideal, and to an extent still transparently hybrid, Modern Literary Tajik (MLT), unless otherwise noted.

Notes

- (1) In the following Modern Standard Persian and Modern Standard Tajik are referred to as Persian and Tajik.
- (2) In examples, Persian and Tajik forms are separated by a slash: *xānd/xond* '(s)he read'.
- (3) In phonemic transcription, (a) initial glottal stop is not marked, and (b) stress is indicated by an underlined vowel.

1.3 Writing systems

1.3.1 Persian

The Persian writing system uses the Arabic alphabet, which is a consonantal system. Vowels are written as follows: the three long vowels are represented by the letter of the consonant nearest in pronunciation. Thus, the letter <y> represents both *y* and *i* [*ī*], <w> both *v* and *u* [*ū*], and <alef> both the glottal stop ' and *ā*. The three short vowels may be, but are usually not, represented by diacritics which ultimately derive from the same letters <w>, <y>, and <alef>. The main innovations in Persian are two: additional letters were created for the four Persian consonants *p*, *č*, *g*, *ž* by providing three dots to distinguish the "Arabic" letters , <j>, <k>, <z> (the dots merged into an oblique stroke in the

case of <g>). Unlike Arabic, in final position, short vowels are always represented by consonantal letters, final *o* by <w>, and both final *e* and *a* by <h>.

The Arabic orthography is retained, distinctly by the subset of letters marked A in Table 8.1a. They represent pharyngeals, a pair of non-strident interdental fricatives, and the voiceless stop *q*, which in Persian phonemically merged with the corresponding plain and strident consonants, respectively, except *z* > *z*, while *q* merged with voiced *ɣ*. Otherwise, the orthography is basically phonemic in the representation of the consonants and long vowels, but does not represent short vowels other than in final position. In rare

TABLE 8.1a: PERSO-ARABIC ALPHABET

Position in Letter Group				Name	Phonemic
Final	Medial	Initial	Separate		
ا			ا	alef	ʾ
آ			آ		ā
ب	ب	ب	ب	be	b
پ	پ	پ	پ	P pe	p
ت	ت	ت	ت	te	t
ث	ث	ث	ث	A se	s
ج	ج	ج	ج	jim	j
چ	چ	چ	چ	P če	č
ح	ح	ح	ح	A he-ye jimih	
خ	خ	خ	خ	xe	x
د			د	dāl	d
ذ			ذ	A zāl	z
ر			ر	re	r
ز			ز	ze	z
ژ			ژ	P že	ž
س	س	س	س	sin	s
ش	ش	ش	ش	inš	š
ص	ص	ص	ص	A sād	s
ض	ض	ض	ض	A zād	z
ط	ط	ط	ط	A tā	t
ظ	ظ	ظ	ظ	A zā	z
ع	ع	ع	ع	A eyn	ʿ
غ	غ	غ	غ	ɣeyn	ɣ
ف	ف	ف	ف	fe	f
ق	ق	ق	ق	A qāf	q
ک	ک	ک	ک	kāf	k
گ	گ	گ	گ	P gāf	g
ل	ل	ل	ل	lām	l
م	م	م	م	mim	m
ن	ن	ن	ن	nun	n
و			و	vāv	v/u/ow/final o
ه	ه	ه	ه	he	h/final e, a
ی	ی	ی	ی	ye	y/i/ey

A = letters occurring mostly in Arabic loanwords; P = letters found in Persian only.

The basic shapes و ا د ر و do not connect to the left.

instances, an Arabic pharyngeal letter is used, such as <š> in <šad> for *sad* ‘hundred’. In Turkic loans, fronted /k/ and backed /k/ are represented by <k> and <q>, respectively.

1.3.2 Tajik

The Tajik alphabet adds six supplementary letters to the Cyrillic inventory (distinguished by means of diacritics), and despite post-Soviet spelling reforms, the Cyrillic orthography established for Tajik in 1939 remains Russian-specific in significant ways. Since examples taken from MLT will be presented in transliteration, certain peculiarities of the standard orthography must be noted here.

The yotated vowels (я, е, ё, ю: in Russian, devices to indicate that a preceding consonant is palatalized) each serve to represent the combination of the consonant /y/ and a following vowel, though /y/ is sometimes represented by й: cf. гӯяд *gūy-ad* ‘says’, юндӣ *yundī* ‘dishwater’, хонае *xonae* /xona-yel/ ‘a house’, тайёр *tauyor* ‘ready’; in the last example, the geminate /yy/ is written with a sequence of two distinct graphs. Prefixed verb stems may likewise be disguised: ояд *o-yad* or биёяд *bi-o-yad* ‘let him come’. The letter <e> additionally represents simple /e/ after a consonant: дидем *did-em* ‘we saw’; word-initially, /e/ is written with non-yotated э: элак *elak* ‘sieve’. Russian и is quasi-yotated and in Tajik always represents /yi/ after a vowel: хонаи ман *xona-i man* /xona-yi man/ ‘my house’, наистон *naiston* /nayiston/ ‘reed bed’.

The letter ӣ (with macron) is a device to distinguish accented word-final -ī (one of two morphological formatives) from unstressed final -i of the syntactic *ezāfelizofat* (EZ) clitic (see next section): дӯсти ман ‘my friend’, but дӯстӣ ‘friendship’; the macron is dropped before an EZ: дӯстии халқо *dūsti-i xalq-ho* /dūsti-yi xalq-ho/ ‘friendship of peoples’. It is also used in monosyllables: кӣ *kī* ‘who?’ (as distinct from the unstressed subordinizer *ki*), чӣ *čī* ‘what?’, сӣ *sī* ‘thirty’, -мӣ *-mī?* (interrogative enclitic).

2 PHONOLOGY

A major factor in the development of the Persian phonology has been the impact of the Arabic loan component, which constitutes approximately 50 percent of the lexicon, with approximately 25 percent frequency of usage. Not as extensive was the impact of the Turkic loan component (and Persian-Turkic bilingualism). Also, the effects of social dynamics are difficult to determine. A particularly challenging aspect is the rather unique cultural setting in Iran, where until recently schooling involved the intensive study of classical Persian literature and prosody through all levels of education.

2.1 Inventory

The sound system of both Modern Standard Persian and Tajik are essentially the same and quite symmetrical. There are a total of 29 Persian and 30 Tajik segmental phonemes including four pairs of stops and four pairs of fricatives, two nasals, a liquid and a trill, three glides, and three pairs of vowels. The single difference is found in the diachronic development of the ENP velar fricative *ɣ* and loaned uvular *q*, which merged in Persian but not in Tajik, and that of the vowels.

TABLE 8.1b: TAJIK ALPHABETS

Cyrillic Letters	Sound	Letter name	ARABO-PERSIAN ALPHABET EQUIVA			
			Position in letter group			
			Final	Medial	Initial	Isolated
а	a				ا	ا
б	b	be	ب	ب	ب	ب
в	v	vov	و	و	و	و
г	g	gof	گ	گ	گ	گ
ғ	γ	γayn	غ	غ	غ	غ
д	d	dol	د	د	د	د
е	e					
ё	yo					
ж	ž	že	ژ	ژ	ژ	ژ
з	z	ze	ز	ز	ز	ز
з	z	zol	ذ	ذ	ذ	ذ
з	z	zod	ظ	ظ	ظ	ظ
з	z	zo	ظ	ظ	ظ	ظ
и	i, V-yi	ye				
ӣ	ī	i (-i zadanok)				
й	y		ي	ي	ي	ي
к	k	kof	ک	ک	ک	ک
қ	q	qof	ق	ق	ق	ق
л	l	lom	ل	ل	ل	ل
м	m	mim	م	م	م	م
н	n	nun	ن	ن	ن	ن
о	o		ا	ا	ا	ا
п	p	pe	پ	پ	پ	پ
р	r	re	ر	ر	ر	ر
с	s	sin	س	س	س	س
с	s	se (-i se nuqta)	س	س	س	س
с	s	sod	ص	ص	ص	ص
т	t	te	ت	ت	ت	ت
т	t	to	ط	ط	ط	ط
у	u					
ӯ	ū		و	و	و	و
ф	f	fe	ف	ف	ف	ف
х	x	xe	خ	خ	خ	خ
ҳ	h	he (-i hutti)	ح	ح	ح	ح
ҳ	h, -a	he (-i havvaz)	ح	ح	ح	ح
ч	č	če	چ	چ	چ	چ
ҷ	j	jim	ج	ج	ج	ج
ш	š	šin	ش	ش	ش	ش
ь	ʔ	alif			ا	ا
Ӣ	ʔ	ayn	ع	ع	ع	ع
э	e initial					
ю	yu					
я	ya					

TABLE 8.2: VOWEL SYSTEMS

Diachronic Development						
Modern Standard Persian	<i>e</i>	<i>i</i>	<i>o</i>	<i>u</i>	<i>ā</i>	<i>a</i>
		^		^		
Early New Persian	<i>i</i>	<i>ī ē</i>	<i>u</i>	<i>ū ō</i>	<i>ā</i>	<i>a</i>
	v		v			
Tajik Persian	<i>i</i>	<i>e</i>	<i>u</i>	<i>ū</i>	<i>o</i>	<i>a</i>
Qualitative Characteristics						
<i>Modern Standard Persian</i>			<i>Modern Standard Tajik</i>			
	Front	Central	Back	Front	Central	Back
High	<i>i</i>		<i>u</i>	High	<i>i</i>	<i>u</i>
High mid			<i>o</i>	High mid		<i>ū</i>
Mid	<i>e</i>		<i>ā</i>	Mid	<i>e</i>	<i>o</i>
Mid low		<i>a</i>		Mid low		<i>a</i>
Low				Low		
Tenseness				Tenseness		
tense/long (stable)	<i>i</i>	<i>u</i>	<i>ā</i>	stable	<i>e</i>	<i>ū</i> <i>o</i>
lax/short (unstable)	<i>e</i>	<i>o</i>	<i>a</i>	unstable	<i>i</i>	<i>u</i> <i>a</i>

2.1.1 Vowels

Both Modern Standard Persian and Tajik have reduced the eight-vowel inventory of Middle and Early New Persian to six, but in quite different ways. Whereas in the Persian of Iran the two long mid vowels *ē* and *ō* collapsed with high *ī* and *ū*, in some varieties of Tajik length was neutralized by the merger of the short and long high vowels and the rounding of long *ā*, as in the scheme in Table 8.2. The vowel systems of Persian and Tajik are thus as follows (see Windfuhr 1987: 543, and Ansarin 2008).

2.1.1a Persian

While the original opposition was one of length, the lowering of the short high vowels and the raising of the long mid vowels and merger with the old long high vowels resulted in the opposition of levels, a pair of high vowels and pair of mid vowels. Correspondingly, *ā* is lower than low mid *a*. The primary distinguishing feature, while inherently long and short, respectively, is stability, in part indicated by the lengthening of *e a o* (< *i a u*) when stressed and their shortening and partial conditioned assimilation when unstressed.

While length distinction is neutralized in Persian in most environments, it is usually found in initial stressed syllables of vocatives and imperatives (Windfuhr 1979: 136):

<i>i</i> vs. <i>e</i> :	<i>bī-adab!</i>	<i>bē-gu!</i>	‘impolite!’	‘say!’
<i>u</i> vs. <i>o</i> :	<i>hūšang!</i>	<i>hōseyn!</i>	‘Hushang!’	‘Hoseyn!’
<i>ā</i> vs. <i>a</i> :	<i>nā-kas!</i>	<i>na-gu!</i>	‘nobody!’	‘don’t say!’

Assimilation

In Persian colloquial speech the unstable vowels *e o a* are subject to assimilation in height and rounding: *kelid* > *kīlid* ‘key’, *xe yābān* > *xī yābun* ‘street’, *bolur* > *būlur* ‘crystal’, *bē-gu* > *bo-gu* ‘say!’ (*bē-* > *bo-* ~ *bū-* only in a few high-frequency stems with back vowel).

Similarly, *a* (1) tends to be raised (a) before sibilants: *madrāse* > *madrese* ‘school’, *čašm*

> *čes̄m* ‘eye’ (whose *a* is retained in the venerable expression *čes̄m* ‘all right, it will be done’ abbreviated from the phrase meaning ‘by the light of my eyes’), and (b) generally in the 2s and 3s personal clitics: *ketāb-at, -as̄* > *ketāb-et, -es̄* ‘your, his book’, and (2) there tends to be length assimilation: *nahār* > *nāhār* ‘lunch’.

The two diphthongal sequences *ey* and *ow* (< *ai, au*) increasingly tend to be contracted to the long vowels [e:], and *ow* > [o:] even in higher registers: *key* > [ke:] ‘when’, *dow* > [do:] ‘run(ning), race’, *če-towr* > [četo:r] ‘how’, contrasting with short *e* and *o*: *ke* [ke] ‘general subordinating conjunction’, *do* [do] ‘two’.

2.1.1b Tajik

The Tajik central and back vowels are rounded: *ū* (orthographically <ū>) is more close than *u*, and slightly lower; *o* is more open than the Russian *o*, and without any *w*-glide.

The older phonemic contrast of length is now replaced by a contrast between stable and unstable vowels. The stable vowels, which are phonetically invariant, are *e, ū, and o*. The unstable vowels, in which the length and quality of articulation vary according to the phonetic environment, are *i, a, and u*. Thus in stressed position and unstressed closed syllables they are equivalent in length to the three stable vowels; in unstressed open syllables they may be shortened and reduced to *ə* or elided, e.g. *did* ‘(s)he saw’, *d’gar* ‘other, else’; *dud* ‘smoke’, *g^udoz* ‘melting’; *bad* ‘bad’, *b^odan* ‘body’. (In some dialects, e.g. of Bukhara, phonemic length is preserved in *ilī* and *ulū*.) The vowel *i* is lowered to *e*, and *u* to *ū*, before *h* or the glottal stop and a consonant: *istehsol* ‘production’, *mone* ‘hindrance’ (cf. *istiloh* ‘idiom, term’, *mohir* ‘skilled’); *mūhtaram* ‘respected’ (cf. *mustaqil* ‘independent’).

Note that in the following *ū* is transcribed as *ū*, and finally stressed *i* by *ī*, which reflects the Cyrillic-Tajik orthography.

2.1.2 Consonants

TABLE 8.3: CONSONANTS

	<i>Labial</i>	<i>Dental</i>	<i>Palatal</i>	<i>Velar</i>	<i>Uvular</i>	<i>Glottal</i>
Stops/Affricates						
tense/voiceless	<i>p</i>	<i>t</i>	<i>č</i> [tsh]	<i>k</i>	<i>q</i> (TAJ)	ʔ
lax/voiced	<i>b</i>	<i>d</i>	<i>j</i> [dzh]	<i>g</i>	<i>q~</i> (PERS)	
Fricatives						
tense/voiceless	<i>f</i>	<i>s</i>	<i>š</i>	<i>x</i>	<i>~γ</i> (PERS)	<i>h</i>
lax/voiced	<i>v</i> (PERS) <i>v~w</i> (TAJ)	<i>z</i>	<i>ž</i>	<i>γ</i> (TAJ)		
Nasals	<i>m</i>	<i>n</i>				
Liquids/Glide		<i>l, r</i>	<i>y</i>			

~ = conditioned variation

2.1.2.1 Stops and fricatives

The distinctive feature of the pairs of stops and fricatives in contemporary Persian is still being debated. It may be identified either as voice or as tenseness, as in this essay.

Tense *p t k* are slightly aspirated word-initially [k^hār] ‘work’.

Fronting of *k* and *g* before *i e a* is typical for Persian speakers of Azeri Turkish background, as it is in Tajik under the influence of Uzbek: Pers./Taj. [k’ar] ‘deaf’.

2.1.2.2 Velar γ and uvular q

This phoneme has triple origins: (1) the inherited lax fricative γ ; (2) the Arabic voiceless stop q and the voiced fricative γ ; and (3) the Turkic voice-neutral velar stop k before back vowels. – In Tajik, the opposition between q and γ is retained.

In Persian, q and γ have merged into a lax consonant: while systemically a lax fricative in the system, its peculiar articulation appears like a virtual compromise of its origins: (1) intervocalically it is a lax velar fricative, [γ]; (2) in initial and final position it is a lax uvular stop [q], which often has an affricate-like release before vowels (varying with the speaker and speech-situation); (3) in spoken Persian, it tends to be assimilated to a tense stop or fricative, *eqtesād* [extesād] ‘economy’, *raqs* [raxs] ‘dance’, *vaqt* [vaxt] ‘time’.

2.1.2.3 Glottal stop

The glottal stop has a triple origin: (1) the inherited automatic feature before initial vowels and in hiatus, [ʔinʔin] ‘this’, [pāʔiz/poʔiz] ‘autumn’, affixal [xānē-ʔi/xona-ʔe] ‘a house; (2) in Arabic loans, (a) the phonemic glottal stop: *soʔāl/suʔol* ‘question’ and (b) the pharyngeal stop *ayn*: *maʔnālmaʔno* ‘meaning’.

Note that in the transcription here the initial glottal is not marked.

2.1.2.4 Palatal \ddot{z}

This is of limited distribution, occurring in a few Persian words such as *žāležola* ‘dew’, *moželmiža* ‘eyelash’. – In Tajik, it also occurs in Russian loans; in spoken Tajik it tends to be replaced by j . (This is also a feature of Uzbek, standardized under a single grapheme; it is noteworthy that the Tajik and Uzbek sound systems are virtually identical.)

2.1.2.5 $v \sim w$

In Persian this labiodental is bilabial only in the diphthong, *au* > *ow* before consonant and pause. – In Tajik there is conditioned variation: intervocalically v is realized as bilabial w in (1) final position, (2) between rounded vowels, and (3) pre-consonantly after a : *vara \dot{q}* ‘page’, *beva* ‘widow’, but *gow* ‘cow’, *suwol* ‘question’, *qawl* ‘speech’ (Taj. orthographically, <qavl>), corresponding to Pers. *gāv*, *soʔāl*, *qowl*. Afghan Persian has retained bilateral w in all environments (note that Persian and Tajik v < w).

2.1.2.6 Lateral l

Arabic rounding of l may be found in *allāh* and *yallā* ‘let’s go’.

2.1.2.7 Final position

Lax stops tend to lose release in final position, while in some dialects of Tajik b d g tend to be devoiced.

Similarly nasals and laterals tend to lack release in the environment CC-pause: [*satL*] ‘pail’, [*hokM*] ‘order, judgment’.

2.1.3 Syllable structure

Given the phonemization of the glottal stop, the syllable structure is predictable: CV, CVC, CVCC: *tāto* ‘till’, *darldar* ‘in’, *dast/dast* ‘hand’. Accordingly, the syllable boundary

is any consonant before a vowel: *dast-am* [das-tam] ‘my hand’, and diphthongs have the structure VC: *eyvānlayvon* [CVCCVC] ‘porch’, *qowmlqawm* CVCC ‘extended family, clan’.

2.1.4 Clusters

In clusters, tenseness tends to be neutralized: [moʃkel] ‘difficult’.

2.1.4.1 Initial clusters

The inherited initial clusters have been resolved by prothetic or epenthetic vowels, either of which could become standardized, e.g. *st-*: *stār* ‘star’ > *setārelsitora*, *br-*: *brādar* ‘brother’ > *barādar/barodar*, but also *brū* ‘brow’ > *abrulabru*. Borrowings of an aberrant structure are assimilated by the introduction of epenthetic or prothetic vowels (to resolve initial consonant clusters), e.g. Pers. *estudiyo* ‘studio’, *teren* ‘train’, Taj. *istansa* < Russ. *stantsiya* ‘station’.

In Tajik, Russian borrowings have added alien initial clusters to the corpus, as *zveno* ‘team, unit’, *Stalinobod* (former name of Dushanbe). Such loanwords have been reproduced in their original orthographies (regarded as normative) in successive writing systems; the degree to which speakers assimilate them to native phonotactic and phonetic norms depends partly on non-linguistic factors such as their level of education and cultural predilections.

2.1.4.2 Final clusters

The Arabic impact brought a large number of final clusters that contravened Persian phonotactics, especially those involving obstruents and their clusters, notably in final position, e.g. *qotblqub* ‘pole’, *majdlmajd* ‘glory’. These were integrated into the Persian system.

In Tajik one finds a common “Eastern” feature whereby medial or final consonant clusters may be epenthized: *šukrullo* ‘thank God’, but *šukʳ* ‘thanks’, or be reduced.

2.1.4.3 Gemination

This is a distinctive characteristic of Arabic, whereas in inherited Persian items it is a marginal feature. While retained in high registers, it is eliminated in contemporary standard pronunciation, e.g. *korre* ‘foal’ vs. *kore* ‘globe’ (Ar.), *mat(t)e* ‘(mechanical) drill’, *banā* ‘building’ (Ar.) vs. *bannā* ‘architect’ (Ar.)

2.1.4.4 Glottal stop, *h* and *r*

Other than in high register or in slow articulation the glottal stop and *h* tend to result in the compensatory lengthening of the preceding short vowel: Pers. *baʳd* > [ba:d] ‘after’, or in occasional compensatory germination of continuants: Pers. *jomʳé* > [jom(:)e] ‘Friday’.

Similarly, *h* tends to be lost in postconsonantal and final position: Pers. *sobh* > [sob] ‘morning’, [sob-e zud] ‘early morning’.

In colloquial Persian, final *r* tends to be elided in final position in some high-frequency items such as *agar* > [age] ‘if’, *digar* > [dige] ‘other; moreover’.

In Tajik, particularly in southern dialects, there is a tendency to drop *h* in all positions, and to drop *r* before *d* in a few common words, especially in the past stem of the common auxiliary *kardan*, *kad-um* ‘I did’, etc.

This loss Tajik has in common with Afghan Persian, where the /a/ in this environment is additionally lowered to /a/: *ba’d* > *lba:dl* ‘after’, *šahr lša:r l* ‘town’.

2.1.5 Hiatus

2.1.5.1 Standard Persian and Tajik

No two adjacent vowels are allowed Hiatus breakers are (1) *-y-*, or (2) a glottal stop with the copula and with both indefinite and derivational *-i-*:

(1) Pers. *xāne-ye tolxona-yi to* ‘your house’, *qāli-ye bozorg/qoli-yi buzurg* ‘large carpet’; *na-y-ālna-y-o* ‘don’t come!’, less distinct when combined with *i*: *mi-ā-y-andlme-^o-o-yand* ‘they come’, *mi-ā-ⁱimlme-^o-o-em* ‘we come’, *bi-ālbj-^o-o* ‘come!’

(2) *xastelxasta -am, -i, -ast*, etc. ‘I, you, (s)he am, are, is, tired’, etc.; indefinite *-il-e*: *jā-iljo-e* ‘some place’; and derivative *-i*: *kojā-ilkujo-i* ‘where from’.

For intervocalic *-g-* see section 2.3 Morphophonemic alternations.

2.1.5.2 Colloquial absence of hiatus

The absence of the hiatus breakers *-y-* and glottal stop, coupled with contraction and loss of postconsonantal *h* and *-rā* > *-ro* > *-o*, is a major feature that distinguishes less formal registers of Persian (see section 2.1.5.2):

<i>ketāb-hā-yam</i> >	<i>hame-y-aš</i> >	<i>hame-y-aš-rā</i> >	<i>raft-e-i</i> >
<i>ketāb-ā-m</i>	<i>hama-š</i>	<i>hama-š-o</i>	<i>raft-i</i>
‘my books’;	‘all of it’;	‘all of it (dir. obj.)’;	‘you have left’.

2.1.5.3 Persian final *-e* as a dialect feature

Where Persian has final stressed *-e* all other varieties of Persian have *-a*, including Tajik and Dari and Kaboli of Afghanistan: Pers. *hame*, Taj. *hama*, Kab. *hama*. The examples above show that in suffixation Persian colloquial has retained the common Persian feature and not been subject to this dialectal feature of Standard Persian.

2.2 Non-segmental features

(1) Stress is word-final in nominals, including nominal verb forms: *baččelbačā* ‘child’, *did-dan/did-an* ‘to see’. Trisyllabics have secondary initial stress: *bačče-g-ilbačā-g-i* ‘childhood’, *did-dan-il did-an-i* ‘worth seeing’. Transparent compounds may exhibit one or more secondary accents: *kam-sāllkam-sol* ‘juvenile’, Pers. *bonyād-gozāri* ‘founding’, trisyllabic *mājerā-ju* ‘adventure seeker, adventurous’; Taj. *kam-bayal* ‘poor (person)’ (*kam* ‘little’ + *sāllsol* ‘year’, *bayal* ‘armful’), Taj. *no-tavon-binī* ‘envy’ (*no-tavon* ‘not-able’ + *binī* ‘seeing’).

A number of common adverbials and interjections have initial or penultimate stress: *balelbale*, *ārelore* ‘yes’, *balkelbalki* ‘but (rather)’, *xeylilxele* ‘very, a lot’, *ammālammo*, *valilvale* ‘but’, *ya’nilya’ne* ‘i.e., ‘I mean’, *yāloyo* (interrog.), *zīrālzero* ‘because’, *hattā/hatto* ‘even’, but Pers. *albatte* vs. Taj. *albatta* ‘of course’, Pers. *hālā* vs. Taj. *holo*, *hole* ‘now’; further, Taj. *māylāš* ‘Good, OK’, Taj. *kānī* ‘well, then’; and conjunctions compounded

with the complementizer *kelki*: *har-kelhar-ki*. Adpositions and enclitics of all kinds are unstressed.

(2) Stress in verb forms is basically regressive, though less so in Tajik than in Persian. The subjunctive-imperative prefix *be-lbi-*, the imperfective prefix *mi-lme-*, and the negative prefix *na-*, are always stressed, the latter assuming primary stress: *mī-rav-am/lmē-rav-am* 'I won't go', *ne-mi-rav-am/na-me-rav-am* 'I won't go'. With *be-lbi-*, Pers. *be-gir-id* 'take-2p!', Tajik differs in having primary stress on the final syllable with secondary stress on the prefix: *bi-gir-éd*. Similarly, stem-initial past tense forms have stress on the stem syllable in Persian, but in Tajik may stress either stem or the (personal) ending according to context: Pers. *gofit-im/Taj. guft-em ~ guft-em* 'we said (it)'. The original stress on the personal ending in un-prefixed present forms is retained in the definite future: *xāh-am raft* 'I will go'.

2.3 Morphophonemic alternations

Synchronically there is a morphophonemic alternation between word-final *-el-a* and *-egl-ag* before the stressed suffixes *-ān/-on* (alternate plural marker), *-ānel-ona* (adjectival) and *-il-ī* (nominal and adjectival): *baččelbača* 'child' > *bačče-g-ān lbačag-on* 'children', *bačče-g-ānel/bača-g-ona* 'childish, children's', *bačče-g-ilbača-g-ī* 'childhood'. Unlike Persian, Tajik has also developed a derived participle: *kard-a* > *kard-a-g-ī* 'done' (see section 3.2.2.2).

In Tajik, this *g* is now intuited as euphonic, and is supplied even for words in *-a* of non-Persian origin (*talaba* 'student', pl. *talaba-g-on*). Note also *yak-to* 'single', *yak-to-g-ī* 'singly', while in Persian such derivation from words of Arabic origin is rare): *qā'ede* 'rule; menstruating' > *qā'ede-g-i* 'menstruation'.

Related, but irregular and rare, is suffixation of the generalizing plural suffix *-j-āt* to final vowel: *ruz-nāme* 'newspaper' > *ruz-nāme-j-āt* 'the printing media', *sabzi* 'greens, vegetable' > *sabzi-j-āt* 'vegetables' in the generic sense.

Diachronically, *-eg* originates in the Middle Persian suffix *-ag*, while the alternation *-i ~ -i-j-āt* originates in Middle Persian suffix *-īg*, which was loaned early into Arabic, with Pers. *j* > *g*, and, with the abstract feminine plural marker *-āt*, was later re-loaned into Persian, where it is moderately productive, where *g* thus still underlies the stem form.

In fact, most conspicuous is the complex morphophonological alternation that is inherent in the morphology of the massive Arabic loan component (see section 7.2.1).

Less evident, and frozen, is the complex Indo-European type morphophonology of Old Persian which is still evident in ablaut series like *bār* 'load' > *bar-* 'bear, carry' > *bor-d* 'carried' (< *bṛ-tá*, from the OIr. verb *bar-* 'bear, carry'), but is largely obscured (see section 3.2.1.1).

3 MORPHOLOGY

3.1 Nominal morphology

Substantives are weakly divided between nouns and adjectives; there are no characteristic stems and no case inflections. Many adjectives and primitive nouns are identical (*javānl javon* 'young; a youth', *bālābolo* 'high; top' – also an adverb, 'up'), and derived adjectives may often function as nouns (*Buxoro-ī* 'Bukharan; a Bukharan', *honar-mand/hunar-mand* 'skillful; a craftsman').

Gender is not marked in nouns or pronouns. Sex may be distinguished lexically : *morgl mury* 'fowl, hen', *xorus/xurüs* 'rooster'; more generally by selective use of a qualifier: *boz/buz* 'goat', *nar-boz/nar-buz* 'billy goat', *māde-boz/moda-buz* 'nanny goat'; *xar-e nar/xar-i nar* 'male donkey', *xar-e mādel/xar-i moda* 'female donkey'. Borrowed Arabic and Russian feminine nouns are used in context: *raqqāselraqqos-a* '(female) dancer', Taj. *studentka* '(female) student', including the feminine marking in abstract terms such as *xārej-e*, in terms like *vezārat-e omur-e xārej-e* 'ministry of foreign affairs', and female first names: m. *Tāher*, f. *Tāhere*.

Animate vs. inanimate distinction is found in subsets of interrogative-indefinite pronouns (see section 3.1.3.7a)

3.1.1 Nouns

Nouns are distinguished by having the categories of number and of definiteness ~ indefiniteness, and by syntactic status, i.e. their ability to combine in an *ezāfelizofat* (EZ) construction as both heads and modifiers by means of the particle *-el-i* and to take adpositions. Since this construction is so fundamental to the elucidation of nominal categories in Persian, a short preliminary definition is offered here (see section 4.2 Ezāfe):

(1) In a nominal EZ the modifier is a noun, pronoun or NP designating the possessor, matrix, source, agency or other relationship with the head: *dar-e xānel/dar-i xona* 'the door of the room', *xolāse-ye dāstān/xulosa-i doston* 'the gist of the tale', *raftan-e ul raftan-i vay* 'his departure'.

(2) In an adjectival EZ the modifier is an adjective or adjectival phrase qualifying the head: *havā-ye xub/havo-i nayz* 'nice weather', *čašm-ān-e por az askl/čašm-on-i pur az ask* 'eyes full of tears'.

3.1.1.1 Number

3.1.1.1a *-hāl-ho* and *-ānl-on*

Singular nouns are zero-marked, plurals by the suffixes *-hāl-ho* and *-ānl-on*. The former is universally applicable: *čiz-hāl/čiz-ho* 'things', *pesar-hāl/pisar-ho* 'boys', *asb-hāl/asb-ho* 'horses', *ma'nā-hāl/ma'no-ho* 'ideas'; *-ānl-on* is more limited in application and less frequent, denoting (1) animates: *pesar-ānl/pisar-on*, *asb-on*, *mo'allem-hālmuallim-on* 'teachers'; (2) paired parts of body: *lab-ānl/lab-on* 'lips', *dast-ānl/dast-on* 'hands', etc. – a relic of the ancient dual number; (3) a few other words: *deraxt-ānl/daraxt-on* 'trees'; *setārel/sitora* 'star', *setāre-g-ānl/sitora-g-on* 'stars'; it is stylistically marked as literary in register, and relatively more frequent in Tajik than in Persian.

Morphophonemic variants occur after *-āl-o* (*dānā-yānl/dono-yon* 'sages'). There is a frozen plural of *niyālniyo* 'forefather' in *niyā-k-ānl/niyo-k-on*.

3.1.1.1b Arabic loan plurals

The Arabic loan component includes a large number of plural forms:

(1) Arabicate suffixes in *-āt/-ot*, *-e-j-āt/-a-j-ot*, *-in/-in*: *xiyāl-āt/xayol-ot* 'dreams, fantasies'; *qal'e-j-āt/qal'a-j-ot* 'fortresses'; *mo'allem-in/mu'allim-in* 'teachers'.

(2) Most distinct are the Arabic "broken plurals": *olamālulamo* 'religious scholars', sg. *ālem/olim*; *axbār/laxbor* 'news', sg. *xabar/xabar* 'information, news'. For lexicalization of such plurals, see section 7.2.1 Loanwords.

3.1.1.2 Indefinite *-il-e* and *yeklyak*

Indefiniteness may be marked by the (unstressed) enclitic *-il-e*. The preposed numeral *yeklyak* ‘one’ may also function to mark indefiniteness. Both may combine, mostly in colloquial speech.

<i>mard-i āmadl</i>	<i>yek mard āmadl</i>	<i>yek mard-i āmadl</i>
<i>mard-e omad</i>	<i>yak mard omad</i>	<i>yak mard-e omad</i>
‘some man came’;	‘one/a man came’;	‘some/a man came’.

While most frequently found with singular nouns, it likewise occurs with plural nouns (see also section 5.2.1 Definiteness and specificity). Tajik also allows plural without enclitic *-i*:

<i>yek mard-hā-i āmad-andl</i>	
<i>yak mard-ho-e omad-and</i>	<i>yak piyola-ho</i>
‘some (or a certain) men came’;	‘some cups’.

Note that the admission of the plural marker distinguishes this function of *yeklyak* from its role as a numeral.

3.1.1.3 Tajik indefinite determiners

In Tajik, there are also two explicitly indefinite determiners, *yag-on* and *kadom (yak)*. These determiners may also be applied to plural NPs. (Note that this function of *kadom* is distinct from its function as interrogative-indefinite ‘which’, which determines a specific NP):

<i>yag-on rūz ba xona-i</i>	<i>mo marhamat kun-ed</i>	<i>yag-on odam-ho-e</i>
some dayto house-EZ us kindness do-p		some persons-INDEF
‘please come and visit us some day’;		‘some persons or other’;
<i>kadom yak zan-e</i>	<i>did-am</i>	
some one woman-INDEF I saw		
‘I saw some woman or other’.		

3.1.1.4 Persian referential *-e*

Colloquial Persian stressed *-e* (*-h-e* after vowel) has referential function to a person or an item spoken about by the speaker (similar to colloquial English ‘this ... here, that ... there’);

<i>pesar-e</i>	<i>čiz-i</i>	<i>na-goft</i>	<i>bačče-h-e</i>	<i>če</i>	<i>mi-xā-d</i>
boy-REF	thing-INDEF	did not say	child-REF-what	wants	
‘this/that boy didn’t say a thing’;			‘what does this/that child want?’.		

3.1.2 Adjectives

Adjectives may be used attributively, predicatively, and adverbially. In no case do they vary for number, or show agreement with any other of the properties of the head noun, unless themselves substantivized:

in gol-hā sefid astlin gul-ho safed-ast
 ‘these flowers are white’.

3.1.2.1 Subsets

3.1.2.1a Substantivized adjectives

Substantivized adjectives behave like nouns, but show animacy distinction in the plural, animate *-ān* vs. inanimate *-hā*: *digar-ān/degar-on* ‘the other people’ vs. *digar-hā/degar-ho* ‘the other things’.

<i>zibā-yān-e</i>	<i>hāliwudl</i>	<i>[sefid-hā-yaš]</i>	<i>zibā-stl</i>
<i>zebo-yon-i</i>	<i>holivud</i>	<i>[safed-ho-š]</i>	<i>zebo-st</i>

‘the beauties-EZ Hollywood’; ‘the white ones are beautiful’, lit.
 ‘the white [ones] of it’; note the sg. verb.

3.1.2.1b Morphological subsets

- (1) Several morphologically distinct classes of adjectives, often substantivized, are made up of verbal participles (see section 3.2.2 Nominal forms).
- (2) Another large notable subset is represented by Arabic loan adjectives as well as active and passive participles.

3.1.2.2 Comparative

The comparative degree is marked by the stressed suffix *-tar/-tar*: *bozorg-tar/buzurg-tar* ‘greater, bigger’. Suppletive are (1) Pers. *xub* > *beh-tar* ‘good > better’, but Taj. *xub-tar*; (2) *xeylil/xele* > *biš(-tar)/beš(-tar)* ‘much, very > more’.

The standard of comparison is introduced by the preposition *az/az* ‘from’: *az ān bozorg-tar/az on buzurg-tar* ‘bigger than that’.

In attributive position the comparative is connected with the noun in the EZ construction like the positive; in predicative position, the adjective usually precedes the copula or verbs of becoming (but may also follow the verb):

Attributive EZ	Predicative
<i>ketāb-e bozorg-tar [az ān]</i>	<i>in boluz az ān [beh-tar] astl</i>
<i>kitob-i [az on] buzurg-tar</i>	<i>in kurta az on [xub-tar]-ast</i>
book-EZ from that bigger	this blouse from that better is
‘the bigger book’;	‘this blouse/tunic is better than that’.

In colloquial Tajik the positive may also be used in a comparative function in predicative position, similar to Turkic:

in kurta az on [xub] as
 this tunic from that good is
 ‘this tunic is better’.

3.1.2.3 Superlative

The superlative degree is marked by *-tar-in/-tar-in*.

In an attributive NP the superlative usually precedes the noun in both Persian and Tajik, but may also follow in the EZ construction in which case the superlative is

expressed as a universal comparative, using the phrase *az hamelaz hama* ‘than all’. In Tajik, the superlative may follow the head similar to the positive and comparative:

boland-tar-in kuhl kuh-e az hame boland-tarl
baland-tar-in kūh kūh-i az hama baland-tar kūh-i baland-tar-in
 ‘the highest mountain’.

In predicative position the superlative is similarly expressed as a universal comparative:

in kuh az hame boland-tar astl
in kūh az hama baland-tar ast
 ‘this mountain is higher than all = the highest’.

ELLIPTIC SUPERLATIVE

The elliptic substantivized superlative may head a partitive EZ construction, or a pronominal clitic, denoting the class of things being compared:

[boland-tar-in]-e kuh-hā-ye donyāl bozorg-tar-in-ešānl
[baland-tar-in]-i kūh-ho-i dunyo buzurg-tar-in-ašon
 highest-EZ mountains-EZ world’; biggest their
 ‘the highest mountains of the world’; ‘the biggest one of them’.

3.1.2.4 Intensive reduplicated and echoic superlative

Adjectives may form a reduplicated EZ construction to express superlative grade, ‘exceedingly’, etc. *xub-e xub* ‘totally good’, *sabz-e sabz* ‘very, thoroughly green’.

In Tajik, similar to Turkic, some qualitative adjectives (esp. of color) form an intensive by addition of a stressed pre-echoic syllable, i.e. a duplicate of the adjective’s first syllable plus a labial, voiced or unvoiced as appropriate: Taj. *sap-safed* ‘snow-white’, *sip-siyoh* ‘jet black’, *top-torik* ‘dark as dark (could be)’.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns and deixis

Personal pronouns are of two kinds: independent and enclitic (colloq. = colloquial):

TABLE 8.4: PERSONAL AND DEICTIC PRONOUNS

	1s	2s	3s	1p	2p	3p
INDEP.	<i>manl</i>	<i>tol</i>	<i>ul</i>	<i>veyl</i>	<i>māl</i>	<i>išānl</i>
	<i>man</i>	<i>tu</i>	<i>ū</i>	<i>vay</i>	<i>mo</i>	<i>ešon</i> <i>vay-o</i>
Colloq.				<i>mā-hāl</i>	<i>šomā-hāl</i>	
				<i>mo-ho(n)</i> ,	<i>šumo-ho</i> ,	
				<i>mo-yon</i>	<i>šumo-yon</i>	
Far			<i>ānl</i>			<i>ān-hāl</i>
			<i>on</i>	<i>vay</i>		<i>on-ho</i>
Near			<i>inl</i>			<i>in-hāl</i>
			<i>in</i>			<i>in-ho</i>
ENCLITIC	1s	2s	3s	1p	2p	3p
	<i>-aml</i>	<i>-atl</i>	<i>-ašl</i>	<i>-emānl</i>	<i>-etānl</i>	<i>-ešānl</i>
	<i>-am</i>	<i>-at</i>	<i>-aš</i>	<i>-amon</i>	<i>-aton</i>	<i>-ašon</i>

3.1.3.1a Register and animacy

1p *mālmo* may refer to the speaker, both as the plural of majesty and as a member of a social group. In polite or formal speech, 2p *šumāšumo* is also used to address an individual. In turn, the colloquial pluralized forms in *-hāl-ho*, Taj. also with alternate *-on*, function to refer to more than one person.

The 3rd person personal pronouns are marked for human animacy, but not gender: *ulū* ‘he, she’, *išānlešon* ‘they (human)’, in Persian also polite for an individual.

In Tajik, the 3rd person is augmented by the demonstratives *in* and *on*. *On-ho* may also refer in respectful usage to one person, and has replaced earlier *ešon* ‘they’ (human), which became an honorific and ultimately a common noun referring to religious dignitaries (*on-ho* may thus refer in respectful usage to one person).

Vej/vay. In Persian the pronoun *vej* is a highly marked polite 3s human alternate, mostly literary, and lacks a plural form, while in Tajik *vay* is both animate ‘he, she’ and inanimate ‘it’. It has the plural form *vay-o*, and may also function as an unmarked demonstrative adjective: *vay kor* ‘that matter’, *vay mard-ho* ‘those men’.

(Note also the animacy distinction in the indefinite-interrogatives *kilkī* ‘who’ vs. *čelčī* ‘what’, and *kas/kas* ‘person’ vs. *čiz/čiz* ‘thing’, and compound forms with them).

3.1.3.1b Attachment of object marker

The direct object marker *-rāl-ro* (see section 3.1.4.2) is affixed regularly in the literary language, except for the standard contraction 1s *ma-rālma-ro* (< *man-rā*). Spoken Persian and Tajik exhibit a variety of forms, notably postconsonantal *-rāl-ro* > *-ol-a*, including *man-olman-a*, *to-rolto-ra*, Taj. *vay-a*.

3.1.3.2 Possessive construction

Possession ‘mine, yours, his’, etc., is expressed by EZ constructions: Persian *māl-e*, lit. ‘possession of’, Tajik *az on-i* ‘from that of’ followed by an independent pronoun. *Az ān-i* and similar constructions are also found in local Persian dialects and earlier stages of Persian.

<i>in pul</i>	<i>māl-e</i>	<i>ki-st</i>	– <i>māl-e</i>	<i>mā-stl</i>
<i>in pul</i>	<i>az on-i</i>	<i>kī ast?</i>	– <i>az on-i</i>	<i>mo-st</i>
this money	property-EZ/ from that -EZ	who COP.3s	– property EZ/ from that-EZ	we COP.3s
	‘whose money is this?’		– it’s ours’.	

3.1.3.3 Enclitic pronouns and oblique

These interpose a euphonic *-y-* after vowels (*šahr-hā-yemān/šahr-ho-yamon* ‘our cities’), but in colloquial speech the initial short vowel of the personal clitics contracts with preceding vowels, as shown in the following example (combined with contracted *-rāl-ro*):

<i>dast-hā-yat-rā</i>	<i>be-deh-id!!</i>	>	<i>dast-ā-t-o be-d-id!!</i>
<i>dast-ho-yat-ro</i>	<i>dih-ed!!</i>	>	<i>dast-o-t-at-it!</i>
hand-p-your.s-DO	IMP.give-2p		
‘give me your hands!’.			

These enclitics function as oblique cases, including possession as in the example above, and as direct and indirect objects.

3.1.3.4 Demonstratives and quasi-articles

The demonstrative pronouns and demonstrative adjectives are identical in form; as adjectives they are preposed. In Persian there is a two-term deictic system, *inlin* ‘this’, *ānlon* ‘that’.

In Tajik there is a three-term system, the basic forms being *in* ‘this’, *on* ‘that’, as well as *vay* ‘that’; the third term is generally interchangeable with *on*, but may be more a rhetorical anaphoric word:

qin bud-agī-st vay kor
tough GERV-COP3s that work
‘must be a tough job, that’.

Inlin and *ānlon* are fairly weak demonstratives, and function virtually as definite articles.

3.1.3.4 Emphatic *ham-*

The demonstrative function is often intensified by prefixation of the emphatic particle *ham-* ‘same’, *ham-inl ham-in*, *ham-ānlham-on*:

ham-in rāh dorost astl rāh-e dorost ham-in astl
ham-in roh durust-ast roh-i durust ham-in-ast
this very road right-COP.3s road-EZ right this.very-COP.3s
‘this road is the right one’; ‘the right way is this one’.

In Tajik dialect also occurs *havay* (<*ham* + *vay*) ‘that very (one)’; other colloquial variants include *ī*, *amī*, *amu* (for *in*, *ham-in*, *ham-on*).

Compounds include: (*ham*)*in-jāl*(*ham*-)*in-jo* ‘right here’, (*ham*-)*ān vaxtl*(*ham*-)*on vaqt* ‘just then’; (*ham*-)*čon-inl*(*ham*-)*čun-in* ‘this very –, just like this, just so’, (*ham*-)*čon-ānl*(*ham*-)*čun-on* ‘that very one, etc.’; and idioms such as Pers. (*ham*-)*injur* ‘just this way, manner’, etc., Taj. (*ham*-)*in xel* ‘this sort’, (*ham*-)*on guna* ‘that kind’.

3.1.3.5 Reflexive-emphatic pronoun

3.1.3.5a Reflexive

The pronoun *xodlxud* is used possessively, ‘own’, reflexively, ‘self’, and emphatically, ‘-self’. When possessive and reflexive, it refers to the subject; when emphatic, there is no such restriction.

(1) In a possessive EZ phrase, it combines with a specifying enclitic pronoun:

medād-e xod-am šekastl
qalam-i xud-am šikast
pen-EZ own-my brake.PT.3s
‘my own pen broke’.

In the 3rd person, possessive *xodlxud* and *xod-aš*, *xod-ešānlxud-aš*, *xud-ašon* may substitute for enclitic *-aš*, *-ešānl-aš*, *-ašon* to resolve a possible conflict of scope:

barādar [-aš] -rā koštl
barodar [-aš] -ro kušt
brother his DO kill.PT.3s ‘he (A) killed his (A’s or B’s) brother’;

barādar-e [xod]-aš-rā košt
barodar-i [xud]-aš-ro kušt
 brother-EZ self-his-DO kill.PT.3s ‘he (A) killed his (A’s) brother’;

barādar-e [u]-rā košt
barodar-i [ū]-ro kušt
 Brother-EZ he-DO kill.PT.3s ‘he (A) killed his (B’s) brother’.

In this disambiguative use, *xod/xud* is not necessarily emphatic.

(2) As a reflexive pronoun it combines with the appropriate enclitic pronoun and adposition:

<i>xod-am-rā mi-šenās-aml</i>	<i>az xod-ešān mi-tars-andl</i>
<i>xud-am-ro me-šinos-am</i>	<i>az xud-ašon me-tars-and</i>
self-my-DO IPFV.know.PR-1s	from self-their IPFV.fear.PR-3p
‘I know myself’;	‘they are afraid of themselves’.

3.1.3.5b Emphatic

(1) When used emphatically, it precedes a noun or pronoun attribute in an EZ construction: *xod-e tol/xud-i to* ‘you yourself’, or with the addition of an enclitic pronoun: *xod-at/xud-at* ‘you yourself’:

Subject	Object
<i>xod-e mo’allem goftl</i>	<i>xod-e mo’allem-rā košt-andl</i>
<i>xud-i mu’allim guft</i>	<i>xud-i mu’allim-ro kušt-and</i>
self-EZ teacher say.PT.3s	self-EZ teacher-DO kill.PT-3p
‘the teacher himself said so’;	‘they killed the teacher himself’.

(2) In topical constructions, the noun or independent pronoun is placed at the head:

<i>mā xod-emān mi-dān-impl</i>	<i>mo’allem xod-aš goftl</i>
<i>mo xud-amon me-don-em</i>	<i>mu’allim xud-aš guft</i>
we self-our IPFV.know.PR-1p	teacher self-his say.PT.3s
‘we ourselves know’;	‘the teacher, he himself said so’.

3.1.3.6 Reciprocals

(1) The main reciprocal pronoun is *ye—digar/yak—digar* ‘one—another’ or *ham-digar/ham-digar* ‘each other’ (often with pronominal enclitic):

<i>yek-i pas az digar-i</i>	<i>ham-digar-ešān-rā mi-šenās-andl</i>
<i>yak-e pas az digar-e</i>	<i>ham-digar-ašon-ro me-šinos-and</i>
one-INDEF after from other-INDEF	each other-their-DO IPFV.know.PR-3p
‘one after the other/another’;	‘they know each other’; also <i>pas az yak digar</i> .

(2) The particle *ham/ham* ‘same, together’ may also function as a reciprocal pronoun synonymous with *yek digar/yak digar*, here with prepositions *bālbo* ‘with’ and *azlaz* ‘from’:

<i>[bā ham] raft-impl</i>	<i>[az ham] jodā</i>	<i>šod-andl</i>
<i>[bo ham] raft-em</i>	<i>[az ham] judo</i>	<i>šud-and</i>
together go.PT-1p	from each other separated	become.PT-3p
‘we went together’;	‘they parted’.	

3.1.3.7 Interrogatives, indefinites, and related adverbials

Animacy distinction is partially retained lexically in the basic interrogative and indefinite pronouns, and by the plural distinction animate *-ānl-on* vs. inanimate *-hāl-ho* of the indefinite pronominal determiners and of ‘other’.

3.1.3.7a Basic interrogatives and indefinite determiners

(1) Basic interrogatives

- (a) pronominal: animate *kilkī* ‘who?’, pl. *ki-hālki-ho*, inanimate *čelčē* ‘what?’, pl. *če-hālčē-ho*, *če-lča-* in compounds;
- (b) demonstrative: *kodāmlkadam* ‘which?’, pl. *kodām-hālkadam-ho*;
- (c) quantitative: *čandlčand* ‘how many?’;
- (d) adverbial: *keylkay* ‘when?’, *kojāлкуjo* ‘where’, and *kulku* ‘where is, are’ (lacking the copula);
- (e) causal: *čerālčaro* ‘why?’;
- (f) yes-no interrogative: *āyāloyo* ‘whether’ (see section 5.1.3 Questions).

(2) Basic indefinite determiners and compounds

- (a) pronominals: (aa) animate *kaslkas* ‘person, somebody’ (pl. *kas-ānlkas-on*, contextually with a connotation of persons of good or noble standing) and (ab) inanimate *čizlčiz* ‘thing, something’; (b) quantifiers: (ba) distributive *harlhar* ‘each’, (bb) *hičlheč* ‘any (at all)’ (interrogative or with negative *na-*), and (bc) collective *hamelhama* ‘all’, (bd) *čandlčand* ‘several, a few’, also *čandinlčandin*.

Examples:

ki budlkībud ‘who was (it)?’; Pers. *če šode* ‘what happened?’, Taj. *čī gap-ast?* ‘what’s the matter?’; *kodām, har, hič, hame ruzlhar, heč, hama ruz* ‘which, every, no (not any), all day’; *čand ketāb dār-ilčand kitob dor-i* ‘how many books do you have?’, *čand(-in) bār/čand(-in) bor* ‘several times’; *key, kojā raft-ilkay, kajo raft-i* ‘when, where did you go?’; *hasan kul hasan ku* ‘where is Hasan?’; *čerā na-y-āmad-ilčaro na-omad-i* ‘why didn’t you come?’, *āyā dorost nist?loyo durust nest?* ‘isn’t that correct?’.

3.1.3.7b Pronominal quantifiers and demonstratives

The interrogatives and quantifiers function as pronouns under the following conditions:

(1) *hičlheč* functions as a pronoun by itself: *hič na-goft-amlheč na-guft-am* ‘I didn’t say anything’.

(2) *hamelhama* similarly functions as a pronoun: *hame āmad-andlhama omad-and* ‘all came’. It is also often used in EZ with a pronoun or noun, singular or plural, to express totality: *hame-yemānlhama-amon* ‘all of us’, *hame-ye donyālhama-i dunyo* ‘the whole world’, *hame-ye zan-hā-ye dehlhama-i zan-ho-i qišloq* ‘all the women of the village’ (It likewise functions as an adverb, see section 3.1.5 Adverbs).

(3) *kodāmlkadam* requires suppletion by a partitive EZ phrase or an enclitic: *kodām-e in-hālkadam-i in-ho* ‘which (one) of these/them?’; *kodām-ešān āmadlkadam-ašon omad?* ‘which (one) of them came?’ (human, pl. clitic); *kodām-aš-rā mi-deh-i?lkadam-aš-ro me-dih-i?* ‘which (of them) will you give?’ (non-human, sg. clitic); plural, *kodām-hā-yemānlkadam-ho-yamon?* ‘which (ones) of us?’.

(4) *harlhar* ‘every’ minimally requires a nominal when used pronominally: *har yeklhar yak* ‘every one’.

(5) *čandlčand* both as interrogative and non-interrogative determinative minimally requires a classifier: *čand tālčand to* ‘how many/several (items)’.

3.1.3.7c Compound indefinite forms

The basic interrogatives and basic indefinite determiners combine to form indefinites:

- (1) *hič-kodāmlheč-kadom* ‘not anyone, none’: *hič kas nistlheč kas nest* ‘there’s no one (here)’, *hič kodām-e ān-hālheč kadom-i on-ho* ‘none of them’;
- (2) *har-kas, -čizlhar-kas, -čiz* ‘each one (person, thing), whoever, whatever’; *hič-kas, hič-čiz* (colloq. *hiččī*)/*heč-kas, -čiz* ‘anyone, anything’ (with negative *na-*);
- (3) *hame-kas, -čizlhama-kas, -čiz* ‘everyone (all persons), everything (all things)’; *har-kodāmlhar-kadom* ‘every (single) one’;
- (4) *harlhar* with the pronominal interrogatives function as heads of generalized relative clauses: *har(-ān-)kelhar(-on-)kī . . .* ‘who(so)ever . . .’, *har-čelhar-čī . . .* ‘whatever . . .’; *har kojālhar kujo . . .* ‘wherever . . .’, etc.

In addition, there are compounds with *yeklyak* ‘one’: *har-yeklhar-yak* ‘everyone’, *hič-yeklheč-yak* ‘no one’, *kodām-yeklkadom-yak* ‘which one?’, *yek-čandlyak-čand* ‘a few’: *yek-čand nafarlyak-čand nafar* ‘a few people’ (with regular singular after numbers).

The indefinite clitic *-il-e*, in colloquial Persian combined with *yek*, is frequently added to the pronominals: *kodām kas-ilkas-e* ‘which one (person)’; (*yek*) *kas-ilkas-e*, (*yek*) *čiz-ilčiz-e* ‘something’. It is similarly added to the numeral *yeklyak* ‘one’ and several common nouns which thereby may function as indefinite pronominals: *yek-ilyak-e*, *šaxs-ilšaxs-e* ‘someone’ (‘person, individual’).

In terms of specificity, while the indefinite series *kas-ilkas-e*, *čiz-ilčiz-e*, etc., may be either non-specific (‘someone or other’) or specific (‘a certain person’), *kodāmlkadom* in its function as an interrogative determines a specific NP, with obligatory *-rāl-ro*:

<i>kodām zan-rā</i>	<i>did-id?!/</i>
<i>kadom zan-ro</i>	<i>did-ed?</i>
which woman-INDEF-DO see.PT-2p	
‘which woman did you see?’	

3.1.3.7d Other indefinites

- (1) Other pronominal indefinites include: *folān/falon* used for a person or thing not specifically named, ‘such-and-such, so-and-so’: *folān ruzl falon rūz* ‘on such-and-such day’, *āqā-ye folānloqo-yi falon* ‘Mr. so-and-so’, *ey folān! lay falon!* ‘hey, you there’, *folān-kasl falon-kas* ‘such-and-such a one, what’s his name?’, and low-respect *yāru* ‘that fellow’, lit. ‘friend’.
- (2) Other qualitative indefinites include (mostly literary) *čen-in, -ān ~ čon-in, -ānlcun-in, -on* ‘such’: *čen-ān kārčun-on kor* ‘such work, doing’, also used adverbially: *čon-ān kardlčunon kard* ‘he did thus’.
- (3) Other quantifiers include *xeylilxele* ‘much, many’, *kam* ‘little, few’, *kam-tar* ‘less, fewer’: *kam-tar kas-ilkam-tar kas-e* ‘only a few, fewer (persons)’, etc.
- (4) The determiner ‘some’ is expressed by the Arabic loan *ba’z-ilba’z-e* preposed to a plural noun: *ba’z-i Tājik-hālba’z-e Tojik-on* ‘some Tajiks’ (note animate plural *-on*).

3.1.3.7e *Compounds with če-lča-*

Other adverbial determiners are compounds with interrogative *čēlčī* plus semantically appropriate terms, here exemplified with interrogative *če-lčī* (literal meaning of the second component in parentheses):

- (1) time and location: *če-vaqt/lčī vaqt* ‘when’ (‘time’); *če-jā’ilčī jo-e* ‘where’ (‘place’);
- (2) quality and manner: (a) *če now’lčī nav’* + N ‘what sort of’ (‘sort, species’), (b) *če-gunelčī guna* ‘how’, + N ‘what kind of’ (‘kind, color’), (c) *če-towr/lčī tavr* ‘how, what manner’ (‘manner, kind’), (d) Pers. *če-jur* + N ‘how, kind of’ (‘kind’), (e) Taj. *čī xel: če-gune hekāyat/lčī guna hikoya* ‘what sort of tale(s)?’, *če now’ lebāsl čī nav’ libos* ‘what kind of clothing?’, etc.;
- (3) amount expressed by *-qad(a)rlqadar* ‘measure’: *če-qadr mi-šav-ad l čī qadar me-šav-ad* ‘how much will it be?’, *čeqadr nānlčī qadar non* ‘how much bread?’.

3.1.3.7f *Digarldigar* ‘other, else’

‘Other, else’ is *digarldigar*, used as a noun and adjective both independently and in various constructions, including regular and inverted (preposed) EZ, the latter confined to fixed phrases in the modern standards:

- (1) independent use with obligatory animacy distinction *-ān* vs. *-hā*: *digar-ān* ‘the others (people)’, *digar-hā* ‘the others’;
- (2) regular construction, *mardom-e digarl mardum-i digar* ‘the other people’, *kas-i digarl kas-e digar* ‘someone else’, *yek-i digarl yak-e digar* ‘another one, someone else’;
- (3) preposed, *digar ki?ldigar ki?* ‘who else?’, *digar čī?ldigar čī?* ‘what else, what next?’.

Particular idioms include *ruz-e digarl ruz-i digar* ‘the following day’, *bār-e digarl bor-i digar*, *digar bārl digar bor* ‘once more, again; next time’.

3.1.3.7g *Tajik hybrids*(1) Turkic *kim-*

Assimilated into the Tajik literary language is a series of Uzbek-Tajik hybrids formed from Uzbek *kim* ‘who?’: *kim-kī* ‘someone (or other), anyone’, *kim-čī* ‘something, anything’, *kim-kadom* ‘some-N or other’: *vay az kim-čī no-rozi ast* ‘she’s unhappy about something,’ *kim-kadom vaqt* ‘sometime or other, whenever’, *dar kim-kujo-ho* ‘somewhere or other’. As opposed to the Persian series *kas-ilkas-e*, *jā’iljo-e*, *čiz-ilčiz-e*, etc., the *kim* series is unambiguously non-specific indefinite.

(2) Tajik indefinite determiners

In Tajik, there are also two explicitly non-specific indefinite determiners, *yag-on* and *kadom (yak)*:

[*yag-on*] *rūz ba xona-i mo marhamat kun-ed*
 some day to house-EZ us kindness IMP.do-2p
 ‘please come and visit us some day’.

These determiners may also be applied to plural NPs.

3.1.4 *Adpositions and case relations*

Case relations may be expressed by means of prepositions, postpositions, or ambipositions.

3.1.4.1 Prepositions

3.1.4.1a Primary prepositions

There are nine primary prepositions, one of which forms an EZ constructions (6), while one is partially (7) and another (9) nearly fully replaced by such construction in colloquial Persian.

(1) *belba* ‘to, in; with, by’ (dative, direction, manner, means)

This is the most frequent preposition with a wide range of meaning:

Dative: *be to mi-deh-amlba tu me-dih-am* ‘I’ll give it to you’; direction: *be šahr/ba šahr* ‘to the city’, manner: *be āsānilba osoni* ‘with ease’, lit. ‘easiness’; means: *be (zabān-e) fārsi harf be-zan!/ba (zabon-i) tojikī gap zan!* ‘speak in (the language of) Persian/Tajik!’.

(2) *dar/dar* ‘in(to)’ (locative)

In Persian colloquial speech this preposition in the sense of “inside” is frequently often replaced by *tu(-ye)*, lit. ‘the inside’, in an EZ construction; in Tajik dialects it tends to be reduced to *da*: *dar ānjāldar on jo* ‘in that place, there’, *dar tābestāndar tobiston* ‘in the summer’, *dar (~ tu-ye) sanduq gozāštdar sanduq guzošt* ‘he put into the box’; idiomatic: *dar in surat/dar in surat* ‘in this case’, *do dar doldu dar du* ‘two by two’.

Omission of *be* and *dar* in Persian. In the locational and directional sense these two may be omitted: (*dar*) *tābestān xeyli garm mi-šav-ad* ‘it gets quite warm in the summer’, (*dar*) *tehrān zendegi mi-kon-ad* ‘he lives in Tehran’; *man (be) sinemā mi-rav-am ~ mi-rav-am (be) sinemā* ‘I’m going to the movies’.

(3) *az/laz* (*a* in Taj. dialects) ‘from, than; along’ (ablative, source, partitive; passage)

Ablative: *az injālaz injo* ‘from here’, *az emruz/laz imrūz* ‘from today’; than: *qeyr az tolyayr az tu* ‘other than you’, *bozorg-tar az tolbuzurg-tar az tu* ‘older than you’; source: *az tarslaz tars* ‘out of fear’; partitive: *por az talālpur az talo* ‘full of gold’; passage: *az pelle bālā raftlaz zina bolo raft* ‘he climbed the stairs’; *az rāh-e Tehrānlaz roh-i Dušanbe* ‘by way of Tehran/Dushanba’.

(4) *tālto* (up)to, until’ (goal), ‘for, as far/long as’ (extent, duration)

tā Tehrānlto Dušanba ‘till Tehrān/Dushanba’, *tā hālālto holo* ‘until now’; *tā andāze-il to andoza-e* ‘to an extent’; *tā dah ruz xabar-i na-dāšt-amlto dah ruz xabar-e na-došt-am* ‘for (a period of) 10 days I didn’t get any news’.

Note that with human objects the locational-directional prepositions require the adverbial noun *piš/peš* ‘front, before’ in the sense of “presence”: *dar, be az piš-e mādar/dar, ba, az, to pēš-i mādar* ‘(being) with, (going) to, (coming) from mother’.

(5) *bālbo* ‘with’ (comitative, instrumental)

Comitative: *bā tolbo tu* ‘with you’, *bā ajalelbo ajala* ‘with, in haste’; instrumental: *bā čakošlbo čakuš* ‘with the hammer’, *bā asb raftlbo asb raft* ‘he left on horseback’.

(6) *barā-yelbaro-i* (benefactive, purposive, causative, in origin *ba rāy-e* ‘for the sake of’)

Benefactive: *in barā-ye to-stlin baro-i tu ast* ‘this is for you’, *barā-ye man tā’rif kon!! baro-i man naql kun!* ‘tell me!’, lit. ‘narrate for me’; purposive: *barā-ye čelbaro-i čī* ‘what for?’; causative: *barā-ye garmā* ‘because of the heat’, but not so used in Tajik.

Pers. colloquial also has *vāse-ye* (benefactive, purposive).

(7) *bilbe* ‘without’

While still active in Tajik, in Persian this preposition has been largely reduced to a privative preformative (see section 7.1.3 Word formation) and replaced by *be-dun-e* in an EZ construction: *bi ~ be-dun-e pedar o mādar/be pidar-u modar* ‘without father and mother’.

(8) *barlbar* 'on(to)' (locative)

This preposition occurs mostly in fixed phrasing:

banā bar inlbano bar, 'based on this, accordingly'. *bar akslbar aks* 'on the contrary'.

In Persian it has been mostly replaced by *ru-ye*, lit. '(sur)face of', in an EZ construction, *ru-ye mizlrū-yi mez* 'on(to) the table'. In Tajik it is still used in basic locative and directional phrases: *bar kūh barf bud* 'there was snow on the mountains'; *bar po istod* 'she stood up, got to her feet'.

(9) *čunlčün* 'like, similar to'

pesar-i čun tolčun tu pisar-e 'a son like you'; *čun man o tolčun man u tu* 'like me and you'. This preposition is mostly replaced by the EZ construction.

3.1.4.1b Secondary prepositions

There are many more secondary prepositions.

(1) Common compound prepositions of the structure N + Preposition. These include *hamrāh bālhamroh bo* '(together) with (person)', *joz az* 'except' (also without *az/az*), *piš az/peš az* 'before', *pas az/pas az* 'after', and prominently those based on a borrowing from Arabic such as *qabl az/qabl az* 'before', *ba'd az/ba'd az* 'after', *ebārat az/liborat az* 'consisting of'; Pers. *rāje* be/Taj. *oid ba* (or *oid-i*) 'about, concerning', *qeyr az/ya'yr az* 'other than, except'.

The preposition *az/az* may alternate with the EZ without and with a difference in spatial relationship; thus both *birun-elberun-i* and *birun azlberon az* 'outside (of)', *birun-e ~ az šahr/berun-i ~ az šahr* 'outsidel/out of town'; Pers. *ba'd az* 'after', but Taj. *ba'd az ~ ba'd-i*. However, *piš/peš* 'front, before' when locational requires an EZ construction: *dar, be, az piš-e xāneldar, ba, az peš-i xona* 'at, to, from the front of the house', but *az/az* when temporal: *piš az zohrl/peš az zuhr* 'before noon'.

(2) There are numerous noun-adverbs or adjective-adverbs forming EZ constructions with the focal noun or phrase, in combination with a primary preposition, mainly *dar, be, az* (of which the first two tend to be omitted, particularly in colloquial):

rulrū 'on' ('face, surface'), *dar, be, az ru-ye divār/rū-yi devor* 'on, onto, (from) off the wall'; others include *miyānlmiyon* 'between' ('middle, waist'), *miyān-e man o tolmiyon-i man-u tu* 'between you and me'; (*dar*) *miyān-e kuh-hāldar miyon-i kūh-ho* 'amid the mountains'; Pers. *tu, darun/Taj. darun* 'in' ('interior'), *tu-ye, darun-e xāneldarun-i xona* 'inside the house'; *tah/tağ* 'under(neath), base of, end of' ('bottom'), *tah-e deraxl/tağ-i daraxt* 'underneath the tree'; *zirl/zer* 'under(neath)', *bālā/bolo* 'above' ('upper part') and *zirl/zer* 'below' ('lower part'), *bālā-ye, zir-e mizl bolo-yi, zer-i mez* 'above, under the table'; *mānand-elmonand-i* 'like', lit. 'resembling', commonly *mesl-elmišl-i N* ('likeness of'), *mānand-e, mesl-e tolmonand-i, mišl-i tu* 'like you'.

3.1.4.2 Postposition *-rāl-ro*

3.1.4.2a Specific direct object

The principal function of *-rāl-ro* is to mark a definite and any otherwise specific noun or noun phrase as the direct object:

<i>ketāb-rā</i>	<i>xarid-aml</i>	<i>doxtar-e</i>	<i>dust-eš-rā</i>	<i>didl</i>
<i>kitob-ro</i>	<i>xarid-am</i>	<i>duxtar-i</i>	<i>dost-aš-ro</i>	<i>did</i>
book-DO	buy.PT-1s	daughter-EZ	friend-his-DO	see.PT.3s
'I bought the book'; 'he saw the daughter of his friend'.				

In Persian colloquial, *-rā* > *-ro* > postconsonantal *-o*: *u-rā gereft-and* > *u-ro gereft-and* ‘they caught him’, *xāne-rā xarid-and* > *xuna-ro xarid-and* ‘they bought the house’ (with *xāne-* > *xuna-*); *in-rā be-gir* > *in-o be-gir* ‘pick up this one’.

In Tajik, *-ro* has the colloquial reflexes *-ra*, *-a*, *-ya*, this last after a vowel: *kūza-ya ovard-am* ‘I brought the jug’.

3.1.4.2b Temporal and local extension

In addition, *-rā* may mark focused temporal and local extension:

<i>[emšab]-rā</i>	<i>injā</i>	<i>bāš-idl</i>	<i>[in hame rāh]-rā</i>	<i>raft-andl</i>
<i>[imšab-ro]</i>	<i>injo</i>	<i>boš</i>	<i>[in hama roh]-ro</i>	<i>raft-and</i>
tonight-DO	here	be.IMP-2s/p	this all	way-DO
‘stay here for the night’;			‘they walked all this way’;	

<i>ru-ye</i>	<i>yax-rā</i>	<i>āb</i>	<i>rixt-and</i>
face-EZ	ice-DO	water	pour.PT-3p
Pers. ‘they poured water on the ice’.			

3.1.4.2c Tajik circumpositional *-ro*

In addition to marking the direct object, *-ro* (or rather its colloquial reflexes) can be found in several other uses. Thus, it may form circumpositions with nouns governed by prepositions:

<i>baro-i kī</i> >	<i>baro-i man</i> >	<i>az</i>	<i>xandidan</i>	<i>murd-im</i> >
<i>[baro-i] ·kī[-ra]?</i>	<i>man[-a]</i>	<i>[a (= az)]</i>	<i>xandidan[-a]</i>	<i>murd-em</i>
sake-EZ	who-for	sake-EZ	I-for	from
‘for whom?’	– for me’;			laugh.INF-for
				die.PT-1p
				‘we died from laughing’.

3.1.4.2d Northern Tajik EZ with *-ro*

In Northern dialects of Tajik, a construction using *-ro* widely replaces the Persian type of EZ: The word order is that of the equivalent Uzbek NP *muallim[-ning] kitob-i*, lit. ‘of-the-teacher his book’:

<i>pisar-i</i>	<i>man</i> >	<i>man[-a]</i>	<i>pisar-am</i>	<i>kitob-i</i>	<i>muallim</i> >	<i>muallim [-a]</i>	<i>kitob-aš</i>
son-EZ	me	I-for	son-my	book-EZ	teacher	teacher-for	book-his
‘my son’;				‘the teacher’s book’.			

Note that with this construction the definite direct object is not marked with *-ro*.

<i>Zaydullo-ra</i>	<i>palink-o-š</i>	<i>peš-i</i>	<i>usto</i>	<i>ovard-am</i>
Z.-for	shoe.p-his [no DO]	before-EZ	cobbler	bring.PT-1s
‘I took Zaydullo’s shoes to the cobbler’s’.				

The use of *-ro* in the preceding two constructions is reminiscent of the earlier stages of Persian (predicative) dative-possessive construction: *ma[-rā] pesar-i hast* ‘I have a son’, lit. ‘to-me a son is’.

3.1.4.3 Other Tajik postpositions

The use of postpositions is a feature that distinguishes Tajik from Persian.

(1) MLT is relatively conservative, admitting formally only *barin* ‘like, resembling’: *man [barin] odam* ‘a person like me’ (vs. Pers. *ādam-i mesl-e, mānand-e man*, with indefinite *-i*).

(2) Other postpositions are:

(a) *qatī* ‘with’ (found in Southern dialects and Afghanistan as a preposition): *tu [qatī]* ‘with you’, *[bo] qošuq [qatī]* ‘with a spoon’ (here as an ambiposition with the synonymous preposition *bo*); and (b) *da* (< *dar*) ‘in, at, to’: *ow[-da] raft* ‘she’s gone to (fetch) water’.

(3) Other postpositions used widely in the Northern dialects are direct borrowings from Uzbek, e.g. *-dan* ‘from’.

3.1.5 Adverbs

3.1.5.1 Nouns

Nouns with inherent locative and temporal meaning are typically found in adverbial function, frequently without adpositions, as mentioned above: *bāzār raft/bozor raft* ‘she went to (the) market’, *xāne nist-and/xona nest-and* ‘they’re not (at) home’.

Iterative-approximative *-hāl-ho*: Such nouns may be marked by plural *-hā* in its function to express iteration or approximation: *šab-[hā] kār mi-kon-ad/šab[-ho] kor me-kun-ad* ‘he works nights’, *in-jā[-hā]/in jo[-ho]* ‘around here somewhere, hereabouts’.

3.1.5.2 Adverbs proper

Adverbs proper include the following:

- (1) A small set of invariable, unmodified adverbs, including *hanuz/hanūz* ‘still, yet’, *hamišelhameša* ‘always’, *hargez/hargiz* ‘ever’ (with negated verb, ‘never’, in Pers. coll. *hič vaxt* ‘any time’), *faqat* ‘only’, Pers. coll. *hey* ‘continuously, repeatedly’.
- (2) A small set of those which selectively admit of limited qualification includes the words for ‘now’: *hālā/holo, hole, aknun ~ konun* (literary) *laknun*, Taj. *hozir*, modified *ham-in hālā, ham-aknun* ‘right now’, *al’ān* ‘now, this moment’. *az, tā hālā/az, to holo* ‘from now on, until now’; *tā aknun/to aknun* ‘up till now’, *az al’ān* ‘from right now’; further, *bas/bas* ‘enough’ (*bas-ā/bas-o* ‘many, much’, literary): *az bas kelaz bas ki . . .* ‘so much that . . .’

3.1.5.3 Adjective-adverbs and noun-adverbs

Essentially all adjectives may function as adverbs. In addition, a large number of nouns function as noun-adverbs. Accordingly, the majority of adverbs are identical with or derived from adjectives or nouns.

- (1) Common adjective-adverbs are: time, *zud/zud* ‘early, soon’, *dirdler* ‘long, late’ (*dir bāz/der boz* ‘long ago, for a long time’; Taj. *dar-rav* ‘immediately’; location, *dur/dur* ‘far, distant’, *nazdik/nazdik* ‘near (to), at hand’; manner, *tanhāl/tanho* ‘alone, only’, *āhestel/ohīsta* ‘slow, slowly’, *yavāš* ‘slow, slowly, restrained’, *tond/tez* ‘quick, quickly’. Adjectival adverbs admit of the comparative: *zud-tar/zud-tar* ‘sooner, quicker’, *har če zud-tar* ‘as quickly as possible’.
- (2) Common noun-adverbs are: *bālā/bolo* ‘up, above, upstairs’, *pā’in/poyon* ‘below, down(stairs)’; *em-ruz/lim-rūz* ‘today’, *em-sāllim-sol* ‘this year’ and related time

- relational terms; *nāgāhān/nogahon* ‘suddenly’ (< *nā-gāh* ‘inopportune time’), *gāh-gāhi*/Taj. *goho* ‘sometimes’ (in origin a plural, *goh-ho*).
- (3) Abstract suffix *N-īl-ī*. Adverbs may also be formed by the periphrastic structure: (a) *be N-ilba N-ī* for manner and (b) *dar, az N-ildar N-ī* for location, the derived quality of the nominal assuming the stressed abstract suffix *-il-ī* (see section 7.1.2 Word formation): *be zud-ilba zud-ī* ‘soon, quickly’, lit. ‘with quickness’, *be sefid-i-ye barf/ba safed-ī-yi barf* ‘as white as snow’, *dar bist kilometri-ye Tehrān/dar bist kilometrī-yi Dušanbe* ‘(at a distances of) 20km from Tehran/Dushanbe’, (*be*) *hamegil(ba) hamaḡī* ‘altogether’ (*hame* ‘all’).
- (4) Reduplications of substantives, adjectives, and participles are also specifically adverbial: *gāh-gāh(i)/goh-goh(ī)* ‘from time to time, sometimes’ (abstract *-i/-ī*), *tond tond gozašt/tez guzašt* ‘it raced by’, *raft-e-raft-elraft-a-raft-a* ‘gradually’.
- (5) Derivational are: (a) a few adjectives with suffix *-ānel-ona* (see section 7.1.2 Word formation) form a few adverbs of time: *šab-ānelšab-ona* ‘at night, by night’, *xoš-baxt-ānelxoš-baxt-ona* ‘luckily, fortunately’; *mota* ‘assef-ānelmutaassif-ona’ ‘regretfully’; (b) a few active participles in *-ānl-on*, e.g. : *davān-davān/dav-on-dav-on* ‘at a run’.
- (6) Arabic borrowings: (a) Adverbs in *-an*. These constitute the largest class of dedicated, morphologically marked adverbs: *moḡqat-an/muḡqat-an* ‘temporarily’, *ettefāq-an/littifoq-an* ‘incidentally’, *ba’d-an/bā’d-an* ‘later’, *fowr-an/fawr-an* ‘immediately’, *asl-an/asl-an* ‘at all, ever’, with neg: ‘not at all, never’. (b) Common adverbials modifying adjectives, adverbs and measure such as *nesbat-an/nisbat-an* ‘relatively’, *kāmel-an/komil-an* ‘completely’, Pers. *taqrib-an*/Taj. *taxmin-an* ‘approximately’. (c) Common Arabic adverbial constructs include *lā-bodllo-bud* ‘undoubtedly’, *be-lā-fāsele/bi-lo-fosila* ‘without interruption, immediately’, *belāxarel biloxira* ‘finally’.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

TABLE 8.5: CARDINAL NUMBERS

1–9		11–19		10–90		100–900	
Pers.	Taj.	Pers.	Taj.	Pers.	Taj.	Pers.	Taj.
<i>yek</i>	<i>yak</i>	<i>yāzdah</i>	<i>yozdah</i>	<i>dah</i>	<i>dah</i>	<i>(yek)sad</i>	<i>(yak)sad</i>
<i>do</i>	<i>du</i>	<i>davāzdah</i>	<i>duvozdah</i>	<i>bist</i>	<i>bist</i>	<i>devist</i>	<i>dusad</i>
<i>se</i>	<i>se</i>	<i>sizdah</i>	<i>sezdah</i>	<i>si</i>	<i>sī</i>	<i>sisad</i>	<i>sesad</i>
<i>čahār</i>	<i>čor, *čahor</i>	<i>čahārdah</i>	<i>čordah</i>	<i>čehel</i>	<i>čil, *čihil</i>	<i>čahārsad</i>	<i>čorsad</i>
<i>panj</i>	<i>panj</i>	<i>pānzdah</i>	<i>ponzdah</i>	<i>panjāh</i>	<i>panjoh</i>	<i>pānsad</i>	<i>panjsad</i>
<i>šeš</i>	<i>šaš</i>	<i>šānzdah</i>	<i>šonzdah</i>	<i>šast</i>	<i>šast</i>	<i>šešsad</i>	<i>lšašsad</i>
<i>haft</i>	<i>haft</i>	<i>hefdah</i>	<i>hafdah</i>	<i>haftād</i>	<i>haftod</i>	<i>haftsad</i>	<i>haftsad</i>
<i>hašt</i>	<i>hašt</i>	<i>heždah</i>	<i>haždah</i>	<i>haštād</i>	<i>haštod</i>	<i>haštsad</i>	<i>haštsad</i>
<i>noh</i>	<i>nuh</i>	<i>nuzdah</i>	<i>nuzdah</i>	<i>navad</i>	<i>navad</i>	<i>nohsad</i>	<i>nuhsad</i>

* literary, archaic

Numbers are construed in order from highest to lowest, with the connective enclitic *-ol-u* ‘and’ after each range: *bist-o yek/bist-u yak* ‘21’, *si-o yek/sī-u yak* ‘31’, etc.; 1000: *hazār/hazor*, or *yek hazār/ryak hazor*; 2000: *do hazār/du hazor*, etc.; 4963: *čahār hazār-o nohsad-o*

šast-o selčor hazor-u nuh-sad-u šast-u se; ‘zero’ is *sefr/sifr* (< Arabic), in Taj. also or *no* ‘nil’ (< Russian).

SINGULAR AND PLURAL MARKING

Numerical quantifiers precede the numerand and canonically require ‘the singular, including cardinal numbers as well as indefinite and interrogative quantifiers (*čand/čand* ‘several; how many’: *čahār, čand darviščor, čand darveš* ‘four, several dervishes’). The exception is the indefinite *ba’z-ilba’z-e* ‘some, several’ which obligatorily requires the plural: *ba’z-i deraxt-hālba’z-e daraxt-ho* ‘several trees’.

3.1.6.2 Ordinal numbers

Ordinal numbers are formed with the suffix *-oml-ūm, -yūm*; after vowels: Pers. *avval*, Taj. *yak-ūm, do-v-oml-du-yūm, se-v-oml-se-yūm, si-yoml-si-yūm* ‘first, second, third, thirtieth’; note that Persian has *avval* (< Arabic), but regular *yek-om* with decades: *bist-o-yek-om* ‘21st’.

Like regular adjectives, ordinals follow the noun qualified in EZ construction: *sāl-e se-v-oml-sinf-i se-yūm* ‘third year/grade’.

Ordinals themselves may be substantivized and constitute the head noun in an EZ phrase to express a date: *avval-e xordādlyak-ūm-i may* ‘the first of Xordād/the first of May’, and may be joined to a pronominal enclitic: *čahār-om-aš-rā na-did-amlčor-ūm-aš-ro na-did-am* ‘I didn’t see the fourth one (of them)’.

3.1.6.3 Fractions

In Persian fractions are expressed by cardinal numerator + ordinal denominator: *yek čahār-om, do se-v-om* ‘one-fourth, two-thirds’. Quite differently, Tajik uses a partitive collocation in the inverse order, ‘from [cardinal denominator] [cardinal numerator]’: *az se du*, lit. ‘from three (equal parts) two’ = ‘two-thirds’, etc., omitting the preposition if the numerator is ‘one’: *panj-yak* ‘one-fifth’, *dah-yak* ‘one-tenth; ten percent’.

3.1.6.4 Distributive adverbs

Distributive adverbs are formed by simple repetition of the number, with or without the classifier *-tāl-ta*: *čahār čahār/čor čor* ‘four by four, in fours’; *do tā do tāldu-ta du-ta (du-to du-to)* ‘two by two, in twos’; or by suffixing stressed *-i* in Persian: *yek-i yek-i* ‘one by one’, *yek-tā-i*, ‘singly’, and in Tajik by suffixing *-g-ī* to the number + classifier stem: *yak-ta-g-ī (yak-to-g-ī)* ‘singly, one by one’.

3.2 Verb morphology

Verbs are inflected for person, number, tense, aspect, and mood. Further distinctions are made by aspectual-modal particles and periphrastic constructions.

3.2.1 Stem formation

3.2.1.1 Present and past stems

Every verb has two stems: Stem I (traditionally called the “present stem”, from which are formed present tenses, and present and active participles) and Stem II (the “past stem”)

from which are formed past tenses, past and passive participles, and other nominal forms. Stem II always ends in a dental, *-d* or *-t*.

Verbs are divided into two form classes: (1) regular (weak), in which Stem II is derived from Stem I by suffixation of *-id* (as *ras-lras-* > *rasid-lrasid-* ‘arrive’), and (2) irregular (strong), in which the form of one stem is not predictable from that of the other (reflecting Old Iranian ablaut patterns, obscured and mixed since Middle Persian). There are a few pairs of verbs with both irregular and regularized Stem II forms, such as *gard-*, *gašt- ~ gardid-* ‘turn, become’. A good number of irregular verbs have regularized forms in *-id*, side by side with the irregular Stem II retained in the higher registers. Note that “regular” and “irregular” refer only to stem relationships; except for the copula, all conjugations and other verb forms are regularly derived from these two stems.

Similar to other Indo-European languages, there are small subsets which have irregular formants and those where the original morphophonemic change is still partially recognizable. The following lists some subsets with select members (omitting Tajik forms; parentheses = higher register; * = final consonant of Old Iranian verbal root + principal marker *-tá-*):

<i>-id</i>	<i>ras-</i> > <i>ras-id-</i> ‘reach, arrive’, <i>keš-</i> > <i>keš-id-</i> ‘stretch, pull’;
<i>-d ~ -t</i>	<i>xān-</i> > <i>xān-d-</i> ‘read, call, sing’, <i>koš-</i> > <i>koš-t-</i> ‘kill’;
<i>-ād</i>	<i>ist-</i> > <i>ist-ād-</i> ‘stand (up)’, <i>oft-</i> > <i>oft-ād-</i> ‘fall’;
<i>-est</i>	<i>dān-</i> > <i>dān-est-</i> ‘know’; <i>tavān-</i> > <i>tavān-est-</i> ‘can’;
<i>-ā > -ud</i>	<i>farm-ā-</i> > <i>farm-ud-</i> ‘command, order’, <i>rob-ā-</i> > <i>rob-ud-</i> ‘rob, steal’;
<i>ar > or-d</i>	<i>bar-</i> > <i>bor-d-</i> ‘bear, carry’, <i>šomār-</i> > <i>šomor-d-</i> ‘count’ (< * <i>ār-tá-</i>);
<i>b > f</i>	<i>yāb-</i> > <i>yāf-t-</i> ‘find’, <i>kub-</i> > (<i>kuf-t-</i>) > <i>kub-id-</i> ‘pound’ (< * <i>p-tá-</i> > <i>-ftá-</i>);
(<i>y</i>) > <i>s</i>	<i>šu(y)-</i> > <i>šos-t-</i> ‘wash’, <i>ru(y)-</i> > (<i>ros-t-</i>) > <i>ru(y)-id-</i> ‘grow’ (< * <i>d-tá-</i> > <i>-stá-</i>);
<i>s > š</i>	<i>nevis-</i> > <i>neveš-t-</i> ‘write’, <i>ris-</i> > (<i>reš-t-</i>) > <i>ris-id-</i> ‘spin’ (< * <i>ā-tá-</i>);
<i>z > x</i>	<i>sāz-</i> > <i>sāxt-</i> ‘build’, <i>riz-</i> > <i>rixt-</i> ‘drop, pour’ (< * <i>k-tá-</i> > <i>xtá-</i>).

Typically, most of the statistically frequent verbs are also most irregular, including:

<i>kon-</i> > <i>kar-d-</i> ‘do, make’;	<i>zan-</i> > <i>za-d-</i> ‘hit’;
<i>šav-</i> > <i>šo-d-</i> ‘become’;	<i>band-</i> > <i>bas-t-</i> ‘bind, tie, close’;
<i>deh-</i> > <i>dā-d-</i> ‘give’;	<i>gu-</i> > <i>gof-t-</i> ‘say’;
<i>gir-</i> > <i>geref-t-</i> ‘take’;	<i>rav-</i> > <i>raf-t-</i> ‘go, leave’.

Suppletive

<i>ā-</i> vs <i>āma-d-</i> ‘come’;	<i>bin-</i> vs <i>di-d-</i> ‘see’.
------------------------------------	------------------------------------

3.2.1.2 Denominal verbs

Verbs may be formed by suffixing regular *-id* to the noun or nominal stem: *nām-*, *nāmid-/nom-*, *nom-id-* ‘name’ (< *nom* ‘name’), Pers., Taj. *fahm-*, *fahm-id-* ‘understand’ (< *fahm* ‘understanding’), Taj. *favt-lfavtid-* ‘pass away’ (< *favt* ‘death’; in origin Arabic action nouns). In both Persian and Tajik this procedure is no longer very productive.

3.2.1.3 Derived causative verbs

Intransitive verbs may be transitivized, and both intransitive and transitive verbs may be made causative by suffixation of *-ān/-on* to Stem I and *-ān-(i)d/-on-(i)d-* to Stem II: *ras-/ras-*, *rasid-/irasid-* ‘arrive’ > *ras-ān-/ras-on-*, *ras-ān-(i)d-/ras-on-(i)d-* ‘make arrive, take to’, *xor-/xūr-*, *xord-/xūrd-* ‘eat’ or ‘drink’ > *xor-ān-/xūr-on-*, *xor-ān-(i)d-/xūr-on-(i)d-* ‘make eat or drink, feed’. Irregular causatives include Pers. *nešin-*, *nešast-* ‘sit (down)’ > *nešān-*, *nešānd-* ‘set (down)’, but Taj. *šin-*, *šišt-* ‘sit’ > *šin-on-*, *šin-on-id-* ‘set, plant’.

A few causative-type derivatives are commonly used with little or no modification of the meaning of the basic verb:

Pers. *suz-/sūz-*, *suxt-/sūxt-* ‘burn (tr., intr.)’ > *suz-ān-*, *suz-ān-d-* ‘(make) burn’; Taj. *firist-*, *firistod-* > *firist-on-*, *firist-on-id-* ‘send’. Others represent a different specialization of the base verb: *keš-/kaš-*, *kešid-/kašid-* ‘pull, draw (tr.), stretch (intr.)’ vs. *keš-ān-/kaš-on-*, *keš-ān-d-/kaš-on-id-* ‘take away, carry off’.

TAJIK

While in Persian derived causativation is only partially productive, in Tajik it is fully so, including:

- (1) make-other-causatives: *dūz-/dūxt-* ‘sew’, *dūz-on-/dūz-on-id-* ‘have something sewn’;
- (2) transitive denominal and deadjectival verbs: *mukofot-on-*, *mukofot-on-id-* ‘reward’ (< *mukofot* ‘reward’), *elektr-on-*, *elektr-on-id-* ‘electrify, power’ (< *elektrifika* ‘electric’), *xušk-on-*, *xušk-on-id-* ‘dry’ (< *xušk* ‘dry’);
- (3) causativation of transitive compound verbs with *kun-*, *kard-* ‘do, make’: *remont kun-on-*, *kun-on-id-* ‘have (something) repaired’.

3.2.1.4 Preverbs

The three most frequently occurring preverbs are *bar/bar* ‘up’ (Pers. colloq. *var*), *dar/dar* ‘in; out’, *foru(d)/furū(d)* ‘down’; further, *bāz/boz* ‘re-, again’ (mostly colloq. *vālvo*), and literary also *farā* ‘forth’.

In Tajik, but not in Persian, the first three preverbs have become inseparably attached to the stem of the most common verbs of motion, and are preceded by the negation *na-* and imperfective marker *me-*:

Pers. <i>bar ā-</i> , <i>āmadl</i>	<i>dar āvar-</i> , <i>āvord</i>	<i>foru ā-</i> , <i>foru āmad-</i>
Taj. <i>bar o-</i> , <i>omad-</i>	<i>dar or-</i> , <i>ovard-</i>	<i>furo-</i> , <i>furomad-</i>
‘go up, out’;	‘bring in, out’;	‘come down, descend’;

but

Pers. <i>bar mi-ā-yam</i>	<i>dar mi-āvar-am</i>	<i>foru mi-āy-am</i>
Taj. <i>me-[bar]-o-yam</i>	<i>me-[dar]-or-am</i>	<i>me-[furo]-yam.</i>
IPFV-out-come.PR-1s	IPFV-out-bring.PR-1s	IPFV-down.come.PR-1s
‘I go up, out’;	‘I bring in, out’;	‘I come down’.

An exception to the Tajik coalescence is *dar gir-*, *girift-* ‘catch, take (fire, etc.)’; *dar na-me-gir-ad* ‘it isn’t catching’.

For light verb constructions, see section 5.5.6.

3.2.2 Nominal forms

While forms derived from Stem I are only minimally productive, and function mainly as adjectives and nouns, those derived from Stem II are fully productive.

3.2.2.1 Forms derived from Stem I

3.2.2.1a Present participle, PR-*andel-anda*

This participle is derived by the suffix *-andel-anda*: *ā-yandelo-yanda* ‘coming; future’. The latter is both an adjective and a noun: *hafte-ye āyandelhafta-i o-yanda* ‘next week’, *dar āyandeldar o-yanda* ‘in (the) future’, whereas *nevisandel/navis-anda* ‘writer’ and most others have evolved into agent nouns.

3.2.2.1b Active participle, PR-*ānl-on*

This participle is derived by the suffix *-ānl-on*: *rav-ānlrav-on* ‘going’.

3.2.2.2 Forms derived from Stem II

3.2.2.2a Infinitive PT-*an*

The infinitive is formed by addition of *-an*: *gereft-an/girift-an* ‘to take, taking’, *šod-an/šud-an* ‘to become, becoming’ (primarily an action noun, which does not function in the verbal system).

3.2.2.2b Short infinitive PT

The “short infinitive” is identical with the past stem (Stem II) of limited application in the verbal system.

3.2.2.2c Past (perfect) participle PT-*a*

The past, or perfect, participle (active or passive) is derived by the suffix *-el-a*: *gereft-el/girift-a* ‘(having been) taken’, *šod-el/šud-a* ‘(having) become’. They may also function as adjectives or nouns: *gozaštel/gozašta* ‘past, (the) past’.

3.2.2.2d Tajik participle PT-*agī*

From the past participle is derived a second “past participle” by the addition of stressed *-g-ī* to the base: *šud-agī* ‘having become’, *girift-agī* ‘having (been) taken’, *guzašt-agī* ‘having passed’.

3.2.2.2e Future participle or gerundive PT-*anil-anī*

This participle is derived by stressed *-il-ī* from the infinitive, or verbal noun: intransitive, *raft-an-i/raft-an-ī* ‘about to go’, transitive, *did-an-ī/did-an-ī* ‘worth seeing’, *bāvar-na-kard-an-ī/ibovar-na-kard-an-ī* ‘unbelievable’.

The latter participles play several specialized roles in verb formation and NP syntax (see sections 3.2.6b.4 Tajik Conjectural Mood, 5.5.4 Conject verbs, and 5.5.2 Modal constructions).

3.2.3 Person marking and 'be'

The usual present tense of the verb 'be' comprises an affirmative independent and an enclitic copula paradigm. The personal endings differ from the latter only in the 3s. The other tense and modal forms are supplied by *bāš-*, *bud-/ boš-*, *bud-*).

TABLE 8.6: 'BE' AND PERSONAL ENDINGS

	1s	2s	3s	1p	2p	3p
Independent	<i>hast-aml</i> <i>hast-am</i>	<i>hast-il</i> <i>hast-ī</i>	<i>hast, astl</i> <i>hast, ast</i>	<i>hast-impl</i> <i>hast-em</i>	<i>hast-idl</i> <i>hast-ed</i>	<i>hast-andl</i> <i>hast-and</i>
Copula	<i>-aml</i> <i>-am</i>	<i>-il</i> <i>-ī</i>	<i>ast, -stl</i> <i>ast, -st</i>	<i>-impl</i> <i>-em</i>	<i>-idl</i> <i>-ed</i>	<i>-andl</i> <i>-and</i>
<i>Personal Endings</i>						
PR	<i>-aml</i> <i>-am</i>	<i>-il</i> <i>-ī</i>	<i>-adl</i> <i>-ad</i>	<i>-impl</i> <i>-em</i>	<i>-idl</i> <i>-ed- ~ eton</i>	<i>-andl</i> <i>-and</i>
PT	<i>-aml</i> <i>-am</i>	<i>-il</i> <i>-ī</i>	<i>-Øl</i> <i>-Ø</i>	<i>-impl</i> <i>-em</i>	<i>-idl</i> <i>-ed ~ -eton</i>	<i>-andl</i> <i>-and</i>
Imperative		<i>-Øl</i> <i>-Ø</i>			<i>-idl</i> <i>-ed</i>	

The endings, but not the copula, insert *-y-* after vowels: Pers. *mi-gu-yam* 'I say', but *emrikā'i and* 'they are American', *xaste im* 'we are tired'; Tajik *me-gū-yam* 'I say', but *dono-yand* 'they are knowledgeable'. *Ast* contracts with final vowels in formal speech: *kojā astlkujo ast* > *kojā-stlkujo-st* 'where is it?'.
hast- functions both as the existential verb and as an alternate of the copula in persons other than 3s:

<i>man</i>	<i>hast-am, to</i>	<i>bo-row!</i>	<i>xub-aml/xub-am ~ xub hast-aml</i>
<i>man xub</i>	<i>hast-am, tu</i>	<i>raw!</i>	<i>xub hast-am</i>
I well EX-1s	you.s go.IMP.2s		well-EX-1s
'I will stay', lit. 'am here, you go!';			'I am well, fine';

but 3s *hast/hast* is only used in existential sentences:

dar in otāq do panjere hast!
dar in xona du tireza hast
in this room two window EX.3s
'there are two windows in this room'.

NEGATION

The negative form of both independent and enclitic forms of 'be' are based on *nist-/nest-*, inflected like *hast-*, *nist-am*, *nist-i*, *nist*, etc. *nest-am*, *nest-ī*, *nest*, etc.

Dialect and vernacular variants include:

	Pers.	Taj.
3s <i>ast/last</i>	> <i>-e</i>	= <i>ast, as</i>
3s <i>-adl-ad</i>	> <i>-e</i>	<i>-a</i>
2p <i>-idl-ed</i>	> <i>-in</i>	<i>-et, -e</i>
3p <i>-andl-and</i>	> <i>-an</i>	<i>-an</i>

Pers. *xub ast* > *xub-e* ‘it’s good’, *mi-ras-adlme-ras-ad* > *mi-res-elme-ras-a* ‘arrives’.

Honorific use: 2p *-idl-ed* and 3p *-andl-and* may be used to address or refer to a single person in polite or formal usage, while Tajik 2p *-eton* (cf. personal clitics) is used only for the plural (see section 3.1.3.1 Personal pronouns).

TAJIK *-ak*

In some Tajik dialects, such as Varzobi, occur forms with an apparent reflex of the nominal diminutive affix *-ak*, with affective connotations:

in-aš *raft[-ak]*, *dil-am* *sūxt*
 this-REF left-AK heart-my burn.PT.3s
 ‘she [left], and I was devastated’.

3.2.4 Aspectual-modal markers and negation

3.2.4.1 Prefix *mi-lme-*

Aspecto-temporal distinctions are expressed primarily by the presence or absence of the stressed imperfective marker *mi-lme-* added to Stem I and Stem II forms. With Stem I forms, *mi-lme-* distinguishes present/future indicative from subjunctive/optative: *mi-rav-aml me-rav-am* ‘I go, am going, will go’ vs. *be-rav-am/rav-am* ‘let me go’. For the modal use of *mi-lme-* with Stem II forms, see sections 3.2.5.9–10 Counterfactual forms.

3.2.4.2 Prefix *be-lbi-*

Stressed *be-* marks the subjunctive and imperative in Persian *be-rav-am* ‘(that) I go’, *be-rav-id* ‘go-2p!’

- (1) *be-* > *bi-* before initial stem vowel: *bi-ā-* ‘come’, *bi-andāz-* ‘throw’; the vowel may be lost: *bi-andāz-and* > *b-endāz-an(d)* ‘let them throw’.
- (2) In colloquial there is assimilation: *be-gozār* > *bo-gzār* ~ *be-zār* ‘let (me . . .)’, *be-rav* > *bo-row* ‘go-2p!’, *be-kon* > *bo-kon* ‘2s do!’

be- is mutually exclusive with the preverb *bar-*: *mi-xāh-am bar gard-am* ‘I want to return’, lit. ‘that I return’, but is optional in compound verbs, particularly in dependent constructions:

ejāze *be-deh-id* *xod-am-rā* *mo'arrefi* *(be-)kon-am*
 permission IMP.give-2p self-my-DO introduced SBJ.make.PR-1s
 ‘allow me to introduce myself’, ‘lit. ‘that I introduce’.

In Tajik, however, *bi-* is vestigial. It occurs regularly only as a morphological suppletive in Stem I forms of the two common verbs *o-lomad-* ‘come’ and *or-lovard-* ‘bring’: *me-bi-o-yam* ‘I come, am coming’; *bi-or*, *bi-or-ed* ‘bring (it)’.

It may be used with Stem I to mark the subjunctive: *agar bi-gū-yad* . . . ‘if he says . . .’, but this usage is characterized as poetic or archaizing (for *agar gū-yad*). As an optional prefix of the imperative it occurs with some common verbs in polite spoken Tajik: *bi-šin(-ed)* ‘sit-2s(2p) down!’ (cf. 2p *šin-ed* ~ *šin-eton*), *bi-don(-ed)* ‘know-2s(2p)!’. Before a stem beginning with *b-* the vowel is modified to *u-*: *bu-baxš-ed-2p* ‘excuse 2p’, *bu-bin-ed* ‘see-2p’.

3.2.4.3 Negation *na-/na-*

The negative prefix *na-*, Pers. *ne-* before *mi-*, precedes other prefixes and takes primary stress: *ne-mi-gu-yam/na-me-gū-yam* ‘I won’t say’, *ne-mi-āvar-aml na-me-bi-or-em* ‘we won’t bring (it)’, *na-raft/na-raft* ‘3s did not go’. It likewise precedes the periphrastic verb forms, e.g. *na-raft-e ast/na-rafta ast* ‘has not gone’.

3.2.5 Basic system of tense, aspect, and mood

Persian and Tajik, like other varieties of Persian, share a core set of basic forms of the verb, and especially in their personal inflection offer analogies with other Indo-European verbal paradigms. The basic aspect system has a triple distinction of imperfective, perfective-aorist, and resultative-stative forms.

The following comparative tables show the basic verb forms and their functions shared by Persian and Tajik, with the example of the verb for “go, leave”, present stem *rav-*, past stem *raft-*; perfect participle *raft-elraft-a*, all with 3p marking, except 2p imperative, first in list form, second in terms of their categorical vectors.

While there are nine basic finite forms, three of those forms have more than one function, of which two forms have both indicative and non-indicative functions: (1) the imperfective past *mi-raft-lme-raft-*, also functions as imperfective counterfactual; (2) the resultative-stative, *raft-e bud-lraft-a bud-*, also functions as perfective counterfactual. (3) one form, *raft-*, the simple past, often also referred to preterite, functions not only in past, but also in present and future contexts, as well as in potential conditions, for which reason the term “aorist” has been suggested by Windfuhr.

Future. All indicative forms may function in future contexts, as there is no paradigmatic future formation. (For the discussion of the Persian verb system, cf. also Windfuhr 1979: 83–126; 1982; 2006).

3.2.5.1 Present-future

mi-rav-ad/lme-rav-ad ‘goes, is going, will go’

The present indicative marked by *mi-lme-* is the general present imperfective form with a wide range of functions. Depending on the context and adverbial cues, it may refer to actions that are:

- (1) habitual-iterative and generic, ‘they (always, generally) go, leave’;
- (2) progressive, ‘they are going, leaving’;
- (3) intentional, ‘they are about to go, leave’;
- (4) future, ‘they will go, leave’.

3.2.5.2 Imperfect

mi-raft/lme-raft ‘was going, used to go, would go’

This imperfective form has both indicative and non-indicative functions.

As an indicative, the imperfect is the general past imperfective form with a wide range similar to the present indicative. Depending on the context and adverbial cues, it may refer to events that are:

- (1) habitual-iterative and generic, ‘they (always, generally) went, left’ ~ ‘would go, leave’ ~ ‘used to go, leave’;

TABLE 8.7: PERSIAN AND TAJIK CONJUGATIONS

	IMPERFECTIVE	PERFECTIVE	RESULTATIVE-STATIVE		
Indicative					
	1	2	3	4	5
	<i>Present</i>	<i>Imperfect</i>	<i>Preteritel Aorist</i>	<i>Present Perfect</i>	<i>Past Perfect</i>
1s	<i>mi-rav-am</i> <i>me-rav-am</i>	<i>mi-raft-am</i> <i>me-raft-am</i>	<i>raft-am</i> <i>raft-am</i>	<i>raft-e am</i> <i>raft-a am</i>	<i>raft-e bud-aml</i> <i>raft-a bud-am</i>
2s	<i>mi-rav-i</i> <i>me-rav-ī</i>	<i>mi-raft-i</i> <i>me-raft-ī</i>	<i>raft-i</i> <i>raft-ī</i>	<i>raft-e i</i> <i>raft-a ī</i>	<i>raft-e bud-il</i> <i>raft-a bud-ī</i>
3s	<i>mi-rav-ad</i> <i>me-rav-ad</i>	<i>mi-raft</i> <i>me-raft</i>	<i>raft</i> <i>raft</i>	<i>raft-e ast</i> <i>raft-a ast</i>	<i>raft-e budl</i> <i>raft-a bud</i>
1p	<i>mi-rav-im</i> <i>me-rav-em</i>	<i>mi-raft-im</i> <i>me-raft-em</i>	<i>raft-im</i> <i>raft-em</i>	<i>raft-e im</i> <i>raft-a em</i>	<i>raft-e bud-impl</i> <i>raft-a bud-em</i>
2p	<i>mi-rav-id</i> <i>me-rav-ed</i>	<i>mi-raft-id</i> <i>me-raft-ed</i>	<i>raft-id</i> <i>raft-ed</i>	<i>raft-e id</i> <i>raft-a ed</i>	<i>raft-e bud-idl</i> <i>raft-a bud-ed</i>
3p	<i>mi-rav-and</i> <i>me-rav-and</i>	<i>mi-raft-and</i> <i>me-raft-and</i>	<i>raft-and</i> <i>raft-and</i>	<i>raft-e and</i> <i>raft-a and</i>	<i>raft-e bud-andl</i> <i>raft-a bud-and</i>
Neg.	<i>ne-mi-rav-</i> <i>na-me-rav-</i> + PR Ending	<i>ne-mi-raft-</i> <i>na-me-raft-</i> + PT Ending	<i>na-raft-</i> <i>na-raft-</i> + PT Ending	<i>na-raft-e</i> <i>na-raft-a</i> + COP	<i>na-raft-e bud-l</i> <i>na-raft-a bud-</i> + PT Ending
Non-Indicative					
	6	2a	3a	7	5a
	<i>Subjunctive</i>	<i>Counterfactual</i>	<i>Aorist</i>	<i>Subjunctive</i>	<i>Counterfactual</i>
	<i>be-rav-</i> <i>rav-</i> + PR Ending	<i>mi-raft-</i> <i>me-raft-</i> + PT Ending	<i>raft-</i> <i>raft-</i> + PT Ending	<i>raft-e bāš-</i> <i>raft-a boš-</i> + PR Ending	<i>raft-e bud-l</i> <i>raft-a bud-</i> + PT Ending
Neg.	<i>na-rav-</i> <i>na-rav-</i> + PR Ending	<i>ne-mi-raft-</i> <i>na-me-raft-</i> + PT Ending	<i>na-raft-</i> <i>na-raft-</i> + PT Ending	<i>na-raft-e bāš-</i> <i>na-raft-a boš-</i> + PR Ending	<i>na-raft-e bud-l</i> <i>na-raft-a bud-</i> + PT Ending
	8			9	
	<i>Imperative present 2p</i>			<i>Imperative perfect/past 2p</i>	
	<i>be-row, be-rav-idl</i> <i>(bi-)raw, (bi-)rav-ed</i>			<i>raft-e bāš-idl</i> <i>raft-a boš-ed</i> <i>(na-raft-e bāš-id)</i> <i>(na-raft-a boš-ed)</i>	
Neg.	<i>na-row, na-rav -idl</i> <i>na-raw, na-rav-ed</i>				

- (2) progressive, less commonly, ‘they were going, leaving’;
 (3) intentional ‘they were about to go, leave’;
 (4) past future ‘they would go, leave (the next day, etc.)’.

Examples for intentional and past future contexts:

to ke čiz-i [ne-mi-dād-i]
tu ki čiz-e [na-me-dod-ī],
 you.s that thing-INDEF not-IPFV-give.PT-2s,
čerā hamān dam-e dar na-goft-il
čaro hamon dam-i dar na-guft-ī
 why that-same at-EZ door not-say.PT-2s
 ‘since you [were not going to ~ would not give] me anything, why didn’t you say so
 right at the door?’

fardā piš-e mādār-aš mi-raft va be u mi-goft
fardo peš-i modār-aš me-raft-u ba ū me-guft
 tomorrow before-EZ mother-his would go and to her IPFV-say.PT.3s,
 ‘tomorrow, he would go to his mother and would tell her . . .’

For the counterfactual function of the imperfect see section 3.2.5.9.

3.2.5.3 Preterite, or “aorist”

raftraft ‘he went’, ‘there he goes’

This is the perfective form of the verb system and used mainly in past, but also in present and future contexts. Due to the renderings of the predominantly past occurrences into English and other languages, it is traditionally often referred to as (simple) past.

(1) It states that an action was performed and (by implication) completed, usually in the past, irrespective of its duration, and is the usual tense of narrative:

mā xeyli dur peyāde raft-impl
mo xeledur piyoda raft-em
 we very far on foot go.PT-1p
 ‘we walked a long way’.

(2) It may also designate an action that is presented as completed at the moment of speech:

<i>to bāš-i,</i>	<i>man raft-aml</i>	<i>hasan kū?</i>	–	<i>āh, āmad!</i>
<i>tu boš-ī,</i>	<i>man raft-am</i>	<i>hasan ku?</i>	–	<i>oh, omad!</i>
you.s be.IMP/-2s	I go.PT-1s	Hasan where.is		aha come.PT.3s
‘you stay here, I am on my way’,		‘where is Hasan – there he is!’,		
lit. ‘I went’;		lit. ‘he came’.		

(3) It may also assert the completion to an action or event in the future, or refer to action or state that will necessarily have been completed by the time of the action proposed:

šāyad mā ham raft-impl
šoyad mo ham raft-im
 perhaps we also go.PT-1p
 ‘we will most likely go, too’, lit. ‘perhaps we went, too’.

Such use is typically found as a precondition for the action in the main clause:

<i>havā ke sard šod . . . /</i>	<i>vaqti āmad, be-gul</i>
<i>havo ki xunuk šud . . .</i>	<i>vaqt-e ki omad, bi-gū</i>
weather that cold become.PT.3s	time-L that come.PT.3s say.IMP.2s
‘when/once the weather becomes cold, . . .’;	‘as soon as he comes, tell me’.

(4) As such, the aorist also may have subjunctive function:

agar na-bud, zang be-zanl
agar na-bud, zang zan
 if not be.PT.3s bell ring.IMP.2s
 ‘if he is not there, call.2s me!’,

where the assumed condition is stated as a completed action or state.

3.2.5.4 Present perfect

raft-e ast/raft-a ast 'has gone, is gone'

(1) The present perfect tense is formed from the past participle plus the enclitic copula (see Table 8.6). In its "traditional" function as a resultative, the perfect designates an action viewed as having been completed, and its consequences still in force, at the time of speaking:

barnāme-rā be kolli ejrā nemud-e andl
plan-ro tamoman ijro namud-a-and
 plan-DO completely execution make.PART make-3p
 'they have fully implemented the plan'.

(2) The perfect may refer not only to the result, but also to the process. This is readily shown with change of state verbs, where English makes overt distinctions:

<i>pušid-e and ~ bud-andl</i>	<i>bačče-rā dast gereft-e budl</i>
<i>pošid-a and ~ bud-and</i>	<i>bača-ro dast girift-a bud</i>
put on.PART be.3p be.PT-3p may imply	child-DO hand take.PART be.PT.3s
'they have/had put on (a coat)',	'(s)he was holding the child by the hand',
or	or
'they are/were wearing (a coat)';	'(s)he had taken the child by the hand'.

The perfect has an additional function which will be illustrated under section 3.2.6b.1 Evidential mode.

3.2.5.5 Pluperfect or distant past

raft-e bud/raft-a bud 'had gone, was gone'

The pluperfect, or distant past, is formed from the past participle and the simple past of *bud-an* 'be'.

(1) Anterior past action

It refers to an action or state in the past already completed by a certain time in the past, often that of the start of another past action (often rendered by the simple past in English):

ketāb-i rā, ke be man dāde bud-id, xānd-aml
kitob-e-ro, ki ba man dod-a bud-ed, xond-a tamom kard-am
 the book-DO [that to me give.PART be.PT.-2p] read.PT-1s/read.PART finished
 make.PT-1s
 'I read/have finished reading the book that you gave me'

(for Tajik, *xond-a tamom kard-am* see section 5.5.3 Aktionsart constructions).

Sometimes a past action is not mentioned, but is understood from the context:

be šomā ādat kard-e bud-impl
ba šumo odat kard-a bud-em
 to you.p adjustment make.PART be.PT-1p
 'we are used to you', lit. 'we had habituated to you';

this is a tag explaining the speaker's reaction to an action just completed.

The pluperfect form, in the context of the time frame mentioned, locates the action as definitively prior to any recent discussion:

u pārsāl āmade budl
ū porsol omad-a bud
 he last year come.PART be.PT.3s
 ‘as you know, etc. he came last year’.

(2) Remoteness

This tense may also be used to underline the remoteness of an action:

devist sāl-e piš az in
dusad sol-i peš az in
 two-hundred year-EZ before from this
pedar-ān-e mā be injā kučid-e bud-andl
bobo-yon-i mo ba injo kūčid-a omad-a bud-and
 fathers-EZ us to here migrate.PART come.PART be.PT-3p
 ‘two hundred years ago our forefathers migrated here’

(for Tajik *kučid-a omad-a bud-and* see section 5.5.4 Conjunct verbs).

(3) Change-of-state verbs

The pluperfect tense of change of state verbs such as *istād-anlistod-an* ‘stand up, stand’, *nešast-anl šišt-an* or *šist-an* ‘sit down, sit’, *xābid-an/xobid-an* ‘lie down, lie’ and ‘go to sleep, sleep’, *pušid-an/pošid-an* ‘put on, wear’, designates past durative time without a sense of prior action or remoteness:

Piruz dar bālāxāne-yaš nešast- budl
Afandī dar boloxona-aš nišast-a bud
 P./A. in upper story-his sit/PART be.PT.3s
 ‘Piruz/Afandi was sitting in the upper story of his house’.

For the counterfactual function of the pluperfect see section 3.2.5.10.

3.2.5.6 Imperative

(1) present imperative: *be-rav-id/rav-ed* ‘go-2p’

The imperative is based on Stem I: 2s -Ø, 2p -*idl-ed*: *be-gir, be-gir-id! gir-, gir-ed* ‘take, take!’.

(2) perfective-resultative imperative: *raft-e bāš-id/raft-a boš-ed* ‘be gone-2p’.

The perfective-resultative imperative is as elsewhere rarely found and both semantically and contextually restricted, mostly found with change-of-state verbs: *nešast-e bāš-idl/šist-a boš-ed* ‘stay seated’, *raft-e bāš/raft-a boš* ‘be gone’.

3.2.5.7 Present-future subjunctive

be-rav-adl (bi-)rav-ad ‘he may go, (that) he go’

The present-future subjunctive (also known as the “aorist subjunctive”, following mainly Turkological nomenclature), is formed on Stem I + personal endings. Persian has the prefix *be-*, while Tajik mainly the plain stem (see Prefix *be-/bi-* section 3.2.4.2).

(1) Independent use

Independently the subjunctive is used as optative, or hortative, and as suggestion:

boland be-šav-iml
xez-em!
 tall SBJ-become.PR-1p/(SBJ).rise.PR
 'let's get up!';

<i>na-tars-iml</i>	<i>be-rav-adl</i>	<i>be-rav-am?l</i>
<i>na-tars-em</i>	<i>rav-ad</i>	<i>rav-am?</i>
not.SBJ-fear.PR-1p	SBJ.go.PR-3s	SBJ.go.PR-1s
'let's not be afraid';	'let him/her go';	'shall I go?'

bāš-adlboš-ad 'let it be, OK', Pers. colloquially *bāš-e*.

(2) Dependent use

Given its function to refer to potential action or state, the subjunctive typically follows modal verbs and expressions:

<i>mi-xāst-am</i>	<i>xāne</i>	<i>be-rav-aml</i>
<i>me-xost-am</i>	<i>xona</i>	<i>rav-am</i>
IPFV-want.PT-1s	house	SBJ-go.PR-1s
'I wanted to go home', lit. 'that I go'.		

As such it occurs in modal constructions and subordinate clauses implying potential actions or states.

Time reference is usually irrelevant, the action expressed by the subjunctive being potentially future in relation to the time of the main verb.

3.2.5.8 Past (perfect) subjunctive

raft-e bāš-adlraft-a boš-ad 'he may have gone, be gone', in past contexts 'might have gone, may be gone'

The past, or perfect, subjunctive is constructed with a past participle + the subjunctive of 'be'. It most commonly occurs in doubting or questioning, desiring or regretting a past action (accordingly rendered in English by a variety of modal formations, depending on the context):

<i>man</i>	<i>boyad</i>	<i>šaš-sola</i>	<i>šud-a</i>	<i>boš-am</i>	
I	must	6-year-old	become.PART	SBJ.be-1s	
<i>ki</i>	<i>ma-ro</i>	<i>padar-am</i>	<i>ba maktab . . .</i>	<i>burd-a</i>	<i>mond</i>
that	I-DO	father-my	to school	take.PART	do.PT.3s
'I may, must have been six years old when my father took me to school . . .'					

for *mond* see section 5.5.4 Conjunct verbs.

<i>bāvar</i>	<i>ne-mi-kard</i>	<i>ke</i>	<i>do</i>	<i>gusfand</i>	<i>gom</i>	<i>šod-e</i>	<i>bāš-andl</i>
belief	not-IPFV.do.PT.3s	that	two	sheep	lost	become.PART	(SBJ).be-3p
'he still didn't believe that two sheep were, may have, could have been lost'.							

Thus it often follows precative particles such as Pers. (*xodā*) *na-kon-ad* 'may (God) not do', Tajik *na-xod* 'God forbid', *ma-b-ād-ā kelma-b-od-o* (*ki*) 'let it not be', hopefully not', *kāškelkoški* 'would that', and the modal *bāyad/boyad* 'must' in its resumptive sense:

na-kon-ad ke šomā harf-e ma-rāl
na-xod šumo gap-i ma-ro
 God forbid that you.p word-EZ I-DO
jeddi gereft-e bāš-id?!?
jiddī fahmid-a boš-ed?!
 seriously take.understand.PART (SBJ).be.PR-2p
 ‘did you really take me seriously?’

3.2.5.9 Imperfect counterfactual

mi-raft/me-raft ‘might, would go, might, would have gone’, ‘if he went, had gone’.

In its counterfactual function the imperfect is an irrealis and tense neutral. As such it is used in both present-future and past contexts and may imply a present or past irreal action:

(1) to express an unrealized desire, after *koški* ‘would that’:

kāške ne-mi-āmad!
koški na-me-omad!
 would that not-IPFV-come.PT.3s
 ‘if only she had (not) come!’

(2) in both clauses of a counterfactual conditional sentence (see section 6.5.3.7 Conditional clauses):

agar mi-āmad, xod-etān mi-did-id!
agar me-omad-ed, xud-aton me-did-ed
 if IPFV-come.PT-2p self-your IPFV-see.PT-2p
 ‘if you came, you would see for yourself’.

In Persian it may also follow *bāyad* for which see section 5.5.2 Modal constructions.

3.2.5.10 Pluperfect counterfactual

raft-e bud/raft-a bud ‘would have gone, if he had gone’

The pluperfect may also express unfulfilled past conditions in either protasis or apodosis, or both, of a counterfactual conditional sentence (see also section 6.5.3.7 Conditional clauses).

kāške zud-tar āmad-e bud-and!
koški zud-tar omad-a bud-and
 would that early-COMP come.PART be.PT-3p
 ‘would that they had arrived sooner’;

agar zud-tar āmad-e bud-i, mi-did-il
agar zud-tar omad-a bud-ī me-did-ī
 if early-COMP come.PART be.PT-2s IPFV-see.PT-2s
 ‘if you had come earlier, you would have seen it’.

In Tajik, but not in Persian, it may add the prefix *me-* in pluperfect conditional function: *raft-a [me-]bud-am*, etc.

The following two tables provide an overview of the basic categorical vectors and of the conjugations of the Persian and Tajik verb systems.

TABLE 8.8: BASIC PERSIAN AND TAJIK VERB SYSTEM, 3s

	Imperfective	Perfective	Resultative-stative
Present	<i>mi-rav-ad</i> <i>me-rav-ad</i>		<i>raft-e astl</i> <i>raft-a ast</i>
Past	<i>mi-raft</i> <i>me-raft</i>		<i>raft-e budl</i> <i>raft-a bud</i>
Preterite/Aorist		<i>raft</i>	
Subjunctive	<i>be-rav-ad</i> <i>rav-ad</i>		<i>raft-e bāš-adl</i> <i>raft-a boš-ad</i>
Counterfactual	<i>mi-raft</i> <i>me-raft</i>		<i>raft-e budl</i> <i>raft-a bud</i>
Imperative 2p	<i>be-rav-id</i> <i>rav-ed</i>		<i>(me-) rafta bud</i> <i>raft-a boš-ed</i>

3.2.5.11 Defective verbs

The verbs *bāš-*, *bud-/boš-*, *bud-* ‘be’ and *dār-*, *dāšt-/dor-*, *došt-* ‘have’ are by nature stative and as elsewhere are irregular and partially defective. They do not take *mi-/me-*, but may do so as part of compound verbs.

3.2.5.11a ‘be’

This verb also lacks the perfect subjunctive and past perfect, **bud-e bāš-*, **bud-e bud-*, while the present and past imperfective *mi-/me-* forms have limited use.

(1) Stem I form *bāš-/boš-* supplies the present subjunctive and imperative of the copula, and is also found in a good number of frozen idiomatic expressions, including: subjunctive in optative function: *zende bāš-idl/zinda boš-ed* ‘may you live! Bravo!’, imperative:

<i>mard bāš!</i>	<i>xaste</i>	<i>na-bāš-idl</i>
<i>mard boš!</i>	<i>monda</i>	<i>na-boš-ed</i>
man IMP.be.2s	tired	not.IMP.be-2p
‘be a man!’;	‘don’t be tired’, i.e. ‘take it easy’, a casual greeting.	

<i>Persian</i>	<i>Tajik</i>
<i>bāš-ad</i>	<i>na-boš-ad</i>
(SBJ.)be-3s	not(SBJ.)-be-3s
‘let it be, OK;	‘in that case’, lit. ‘may it not be’;

<i>na-bāš-ad</i>	<i>ke . . .</i>	<i>heč na-boš-ad</i>
not(SBJ.)-be-3s	that	any not(SBJ.)-be-3s
‘I hope not that . . .’;		‘at least, in any case’, lit. ‘may it not be at all’.

TAJIK

<i>na-boš-ad, pagoh</i>	<i>me-o-yed</i>
well then tomorrow	IPFV-come.PR-2p
‘well then, why don’t you come tomorrow?’;	

heč na-boš-ad, yak rūz mehmon-i mo šav-ed
 any not(SBJ.)-be-3s one day guest-EZ we (SBJ.)become-2p
 'at least stay with us one day'.

PRECATIVE

na-boš-ad man kasal šav-am
 not(SBJ.)-be-3s I ill (SBJ.)become.PR-1s
 'God forbid I should fall ill'.

(2) The imperfective form, *mi-bāš-/me-boš-*, may substitute for the existential *hast-* under felicity conditions, e.g. typically literary, but also to lend weight to a statement or when the verb is widely separated from its antecedent:

salim, ke hame-ye šomā u-rā mi-šenās-id,
Salim, ki hamā-i šumo ū-ro me-šinās-ed,
 S. that all-EZ you.p he-DO IPFV.know.PR-2p
ozv-e ān hey'at [mi-bāš-ad]!
a'zo-i on komsomol [me-boš-ad]
 member-EZ that committee IPFV-be.PR-3s
 'Salim, whom you all know, is a member of that committee/the Komsomol'.

(3) In Tajik, but not in Persian, *me-* in its counterfactual function may be added to Stem II: *bud-am*, etc. ~ *me-bud-am*, etc.

3.2.5.11b 'have'

The meaning 'have' of this verb is derived from its basic meaning 'keep, hold'. When used in its primary sense (which implies an imperfective-durative state), this verb does not admit the prefix *mi-/me-* with either stem, while the resultative-stative form *dāšte bāš-/došt-a boš-* functions also as the present subjunctive:

ketāb dār-adl pul na-dāšt-idl
kitob dor-ad pul na-došt-ed
 'he has books'; 'you had no money';

momken ast dāšt-e bāš-andl ān-rā dāšt-e bāšid!
mumkin ast došt-a boš-and on-ro došt-a boš-ed!
 'it is possible they have it' or 'had it'; 'have, keep it!'

It may take *mi-/me-* like any other verb when forming part of an idiomatic compound verb with a nominal or adjectival component:

u-rā dust mi-dāšt-am
vay-ro dūst me-došt-am
 he-DO friend IPFV.have.PT-1s
 'I used to like/love him';

pošt-e bām-rā tir-hā-ye mohkam negāh mi-dār-andl
bom-ro bolor-ho-i yafs nigoh me-dor-and
 roof-DO beam-p-EZ stout support IPFV.keep.PR-3p
 'the roof is held up by stout beams'.

3.2.6a *Extended verb system of Modern Standard Persian*

3.2.6a.1 Evidential or non-witnessed mode

While not found in Early New Persian, Modern Standard Persian has a sub-system of evidential forms. (Windfuhr 1979: 90 [“reported” speech, Table Verb system of contemporary literary Persian], 1982, 1985, 1987; Lazard 1985, 2000; and Jahani 2000a). It is morphologically signaled not by a single marker, but its forms are derived from the perfect participle. The evidential subset is categorically parallel to the sub-system of the past. This results in the double function of the form *raft-e ast*, as both present resultative-stative and evidential perfective past.

TABLE 8.9: PERSIAN EVIDENTIAL FORMS, 3s

	Imperfective	Perfective	Resultative-stative
Present	<i>mi-rav-ad</i>	<i>raft</i>	<i>raft-e ast</i>
Evidential	<i>mi-[raft-e] ast</i>	<i>[raft-e] ast</i>	<i>[raft-e bud-e] ast</i>
Past	<i>mi-raft</i>	<i>raft</i>	<i>raft-e bud</i>

This evidential mode is confined to the past. Its function and range are similar to those described for Turkish and in general by Johanson (2000). That is, it marks an action or state as non-manifest. This generally includes such qualifications as second-hand knowledge, conclusion, the latter often encountered in historical surveys and personal reminiscence, among others, and is rendered into Western languages by various tags, such as “they say, evidently, apparently”.

be zabān-e Māzandarāni še’r mi-[goft-e] and
 in language-EZ Māzandarāni poetry IPFV.say.PART COP-3p
 ‘(the records suggest that) they used to write poetry in the Mazandarani language’;

yād-am mi-ā-yad
 to memory-my IPFV-come.PR-3s
sāl-hā piš dar bāre-ye in xeyli motāle’e mi-[kard-e] am
 year-p before about-EZ this much research IPFV-do.PART COP.1s
 ‘I remember now that years ago I used to study this very intensively’;

mi-guy-and diruz dar yek tasādoḡ dah nafar [košt-e
 IPFV-say.PR-3p yesterday in one accident ten person kill.PART
šod-e] and
 become.PART COP.3p
 ‘they say yesterday ten people got killed in a single accident’;

ma’lum šod zud-tar [āmad-e bud-e] ast
 evident become.PT.3s soon-COMP come.PART be.PART COP.3s
 ‘it became evident from what he said, that he had come earlier’.

3.2.6a.2 Progressive

The progressive is not yet fully integrated into literary Persian. The means for this innovation is the verb *dār-/dāšt-* ‘to keep, have’ and refers to an event as ongoing

or imminent. It is preposed to the imperfective forms of the main verb, including the evidential imperfect.

TABLE 8.10: PERSIAN PROGRESSIVE, 3s

	Progressive		Imperfective	Perfective	Resultative-stative
PR	<i>dār-ad</i>	<i>mi-rav-ad</i>	mi-rav-ad	raft	raft-e ast
PT	<i>dāšt</i>	<i>mi-raft</i>	mi-raft	raft	raft-e bud
Evidential	<i>dāšt-e ast</i>	<i>mi-raft-e ast</i>	mi-raft-e ast	raft-e ast	raft-e bud-e ast

Functionally, it disambiguates the progressive and imminent-future functions of the imperfective, but has not yet emptied the latter of that function. Moreover, it is indicative only and cannot be negated. It precedes the main verb and may be separated from the latter. Significantly, both auxiliary and main verb are inflected, but may be separated:

al'ān dār-ad *howsele-am* *sar* *mi-rav-ad* (standard)
al'ān dar-e *howsela-m* *sar* *mi-r-e* (colloquial)
 now PROG.PR-3s patience-my head IPFV-go.PR-3s
 'my patience is running/about to run out now';

xorus *dāšt* *dāne* *mi-čid* *ke . . .*
 rooster PROG.PT.3s grain IPFV-pick.PT.3s when
 'the rooster was picking up grains, when . . .';

zāheran *rānande* *dāšt-e* *dowr mi-zad-e* *ke . . .*
 apparently driver PROG.PART-EV turn IPFV-hit.PT-EV.3s when
 'apparently the driver was about to make a turn when . . .'

3.2.6b Extended verb system of Modern Standard Tajik

Tajik has considerably expanded the common Persian system both in number of forms and range of functions, in addition to the inherited system of aspect, tense, and mood. There are three innovative subsystem which intersect with the basic system, and with each other.

Like Persian, Tajik has developed a sub-system of the evidential mode, based on the perfect participle, e.g. *rafta*, but expanded it to include also present tense forms and non-indicative forms.

Likewise similar to Persian, it developed a progressive sub-system which, however, is not based on an inflected converb, but on a doubly participial compound consisting of the perfect participle of the main verb plus the perfect participle of the verb 'stand', reflecting the pervasive participialization of Tajik. Here again, unlike Persian this new system is expanded to the non-indicative and, in one case, to the evidential mode.

In addition, Tajik has a less developed conjectural sub-system (see sections 5.5.3 Aktionsart constructions and 5.5.4 (Tajik conjunct verbs).

3.2.6b.1 Evidential mode

raft-a

There are four complex tense forms all structured on the form of the perfect participle *raft-a*. Similar to their function in Persian, they denote an action or occurrence known to

the speaker not by direct observation but from collateral sources – hearsay or logical inference (quotative, inferential); or a situation unexpectedly discovered, or a fact only just appreciated as relevant (mirative). The indirect nature of the experience may (or may not) be signaled by an epistemic verb such as *šunav-šunid-* ‘hear’ or *xabar gir-ğirift-* ‘be informed’, or a phrase such as *ma’lum šud* ‘it became known’.

3.2.6b.1a Perfect as evidential form

[raft-a] ast ‘(evidently, etc.) went, left’ and ‘has, is (evidently) gone’

The perfect tense, apart from its role as resultative, also functions as an evidential past (preterite, aorist) as well as evidential perfect (resultative):

mi-gū-yand ki kulliyot-i in navisanda [bar omad-a] ast
 IPFV-say.PR-3p that collected works-EZ this writer out-come.PART-EV COP.3s
 ‘they say that collected works of this writer’s have appeared, are out’;

rūz-i raft-an-aš kasal šud-a boš-ī,
 day-EZ go.INF-his ill become.PART (SBJ.)be.PR-2s

dušanbe [raft-a]-y
 Monday go.PART-EV COP.3s
 ‘if you fell ill on the day of his departure, then it was Monday that he left’
 (inference); *[raft-a]-y < rafta-ast*.

Unlike Persian, the perfect of ‘be’ and ‘have’, *bud-a-* and *došt-a-* refers to present time when used in evidential contexts (generally, as miratives):

odam-i xub [bud-a]-ast či zan-i xub-e
 person-EZ good be.PART-COP3s what woman-EZ good-INDEF
[došt-a]-ed!
 have.PART-EV COP.2p
 ‘he’s a good man as it turns out’; ‘what a good wife you have!’;

pul-i mayda-am na-[bud-a]-ast – šumo pul-i
 money-EZ small-to me not be.PART-EV.COP.3s you.2p money-EZ
non-ro dili-ed!
 bread-DO give.IMP-2p
 ‘I don’t have change – you pay for the bread’

(this sentence also exemplifies an idiom of ‘be’ with the sense ‘have’).

3.2.6b.1b Evidential pluperfect

[raft-a bud-a]-ast ‘(evidently) had gone’

Similar to its function in Persian, this tense corresponds to the regular pluperfect, while connoting quotation, inference or sudden realization:

loiha-i on peš az jang tayyor [šud-a bud-a]-ast
 projectEZ that before from war prepared become.PART be.PART.EV-COP.3s
 ‘his project was evidently ready before the war’, lit. ‘had been readied’.

3.2.6b.1c *Evidential durative*

me-[raft-a]-ast 'is (evidently) going, goes, will go, be going' and 'was (evidently) going, used to go, would go, be going'.

This form appears to be indifferent to tense, and may designate past or present habitual or progressive action, as well as future (or intended) action:

(1) Present:

in navozanda-g-ī xudd-i mullo-g-ī barin [bud-a]-ast-u
 this playing like-EZ mullah-ing like be.PART-EV-COP.3s-and
na-navoz-ed, az xotir me-[bar-omad-a]-ast
 not.SBJ-play-2pfrom memory IPFV.out.come.PART.EV COP.3s
 'playing a musical instrument is like reading-and-writing: if you don't play, you forget how to';

mā'lum-ast ki ū pagoh me-[raft-a]-ast
 known be.PT.3s that he tomorrow IPFV-go.PART.EV-COP.3s
 'it's known that he is going-EV tomorrow'.

(2) Past:

me-gū-yand ki peš-tar boy-ho-i qišloq
 IPFV-say.PR-3p that before-COMP richman-p-EZ village
sol-e yak-du bor ba šikor me-[bar-omad-a]-and
 year-INDEF one-two times to hunt IPFV-out.come.PART.EV COP-3p
 'they say that formerly the rich men of the village used to go hunting once or twice a year'.

3.2.6b.1d *Evidential progressive*

[raft-a istod-a bud-a]-ast 'is (evidently) going' and 'was (evidently) going'.

This is equivalent in tense function to two evidential tenses, the present progressive (*raft-a istoda-am*) and the past progressive (*raft-a istod-a bud-am*):

(1) Present:

Šodi dar xona, ba xob kard-an tayyor-ī [did-a istod-a
 Shadi at home to sleep make.INF preparing see.PART PROG.PART
bud-a] ast
 be.PART.EV COP.3s
 'Shadi is now at home; he is preparing to go to bed'.

šumo yak asar-i naw [nivišt-a istod-a bud-a] ed
 you.p one opus-EZ new write.PART PROG.PART be.PART.EV COP.2p
 '(I see) you are, have been writing a new opus'.

(2) Past:

vay kitob [xond-a istod-a bud-a] ast
 he book read.PART PROG.PART be.PART.EV COP.3s
ki man dar-ro taq taq [kard-a] am
 that I door-DO knocking make.PART COP.1s
 'he was evidently reading a book when I knocked at the door'.

3.2.6b.2 Progressive forms

The progressive forms are marked by the perfect participle of the main verb and the perfect participle *ist-*, *istod-* ‘stand’: *raft-a istoda* followed by the various tenses of ‘be’. By their semantics, the stative verbs, *dor-ldoš-* ‘have’ and *boš-/bud-* ‘be’ do not have progressive tenses. For evidential, subjunctive and conjectural forms of the progressive, see the respective subheadings.

3.2.6b.2a Present progressive

[*raft-a istod-a*]-*ast* ‘is going’

This form refers to an action now in progress, sometimes to habitual action:

man maktub [nivišt-a istod-a]-am
I letter write.PART PROG.PART-COP.1s
‘I am writing a letter’;

bača-ho bozi [kard-a istod-a] and
children play do.PART OROG.PART COP.3p
‘the children are playing’;

mo dar institut fann-ho-i gunogun-ro [omūxt-a istod-a] im
we in institute subject-EZ various-DO learn.PART.PROG.PART COP.1p
‘we are learning various subjects at the institute’.

3.2.6b.2b Past progressive

[*raft-a istod-a*] *bud* ‘was going’

This tense designates an action observed in progress in the past, often at the start of a second action:

Nuralī asb-aš-ro ba mex [bast-a istod-a] bud,
N. horse-hisDO to nail tie.PART PROG.PART COP.1s
ki Hoji-umar namo-yon šud
when H.-U. appear.PR.PART become.PT.3s
‘Nurali was tethering his horse to the stake when Hojjumar appeared’.

3.2.6b.3 Subjunctive forms

3.2.6b.3a Durative perfect subjunctive

[*me-raft-a*] *boš-* ‘may have gone’ and ‘might have gone’

The durative subjunctive is indifferent to present and past distinction, and expresses conjecture, doubt, apprehension, regret, etc. in relation not only to progressive, habitual or iterated actions in the past, but also to current or potential actions of a durative nature:

(1) Present

ba kujo [me-raft-a] boš-ad?
Ro where IPFV.go.PART (SBJ)be.PR-3s
‘where might she be going (I wonder)?’

(2) Past

boyad šodī ham az in kor-ho-i modar-i mehrubon-aš
 must Sh. also from these doing-p-EZ mother-EZ kind-his
zavq [me-girift-a] boš-ad
 pleasure IPFV.take.PART (SBJ.)be.PR-3s
 ‘Shodi must have derived pleasure from everything his dear mother did’.

3.2.6b.3b Present progressive subjunctive

[raft-a istod-a] boš-ad ‘may be doing’

This tense is rarely used in MLT, but is widespread in Northern dialects. It has similar modal connotations to the past (perfect) subjunctive and the durative subjunctive, in respect of a progressive action in past or present-future time, and occurs in concessive clauses:

on-ho az mašina xele dur kor [kard-a-istod-a] boš-and ham,
 they from machine very far work do.PART PROG.PART (SBJ.)be.PR-3p also
šavqun-i xomūš-na-šav-anda-i on na-me-mond
 noiseEZ silent-not-becoming-EZ that not-allow.PT.3s
ki ovoz-i hamsoya-i xud-ro šunav-and
 that saying-EZ neighbor-EZ self-DO (SBJ.)hear.PR-3p
 ‘although they were working quite far from the machine,
 its constant noise did not allow them to hear what the person next to them was
 saying’.

3.2.6b.4 Conjectural mood

raft-agī-st- ~ raft-agī-y-

This mood is used to express an unsubstantiated conjecture or assumption. It is a peculiarity of Northern dialects that has been assimilated into MLT. It is constructed upon the past, or perfect participle, in *-ag-ī* and forms of ‘be’. Only present tense forms are used. These are either formed with the contracted existential verb: *raftagī-st-* (*-st < hast-*), or with the copula: *raftagī-y-*. Colloquially, these forms are contracted (*-agī + personal endings beginning with a- and e- > -agī- > -age-*, and further all *> -agi-*):

<i>raftagī-st-am</i>	<i>raftagī-st-ī</i>	<i>raftagī-st</i> (zero)	<i>raftagī-st-em</i>	<i>raftagī-st-ed</i>	<i>raftagī-st-and</i>
<i>raftagī-yam</i>	<i>raftagī-yī</i>	<i>raftagī-st</i> (<ast)	<i>raftagī-yem</i>	<i>raftagē-yed</i>	<i>raftagī-yand</i>
<i>raftagem</i>	<i>raftagī</i>	<i>raftage</i> (<a(st))	<i>raftagem</i>	<i>raftag-et(on)</i>	<i>raftagen</i>
<i>raftagim</i>	<i>raftagi</i>	<i>raftagi</i>	<i>raftagim</i>	<i>raftagid,-it(on)</i>	<i>raftagin</i>

While theoretically the conjectural mood may have all tense, modal, and aspectual forms, only four forms are used in Tajik.

3.2.6b.4a Past conjectural

[raft-agī]-st ~ raft-age ‘I suppose he went’

The tense connotes various degrees of supposition or conjecture, and may be translated with the help of an adverb such as “probably” or modal idioms “may, might, must have”

<i>ū az šahr [omad-agī]-st</i>	<i>faromūš kard-aged (< [kardagī]-yed)</i>
he from city come.CNJECT-COP	forgotten make.CNJECT-COP.2p
‘he may have come from the city’;	‘you’ve probably forgotten’.

However, the tense of the verb ‘be’ invariably has present reference:

agar hamon faranjidor Pūlod na-boš-ad, mayonarav-aš [bud-agī]-st
 if that-same veiled figure Pulād may not be go-between-his has been-EV
 ‘if that veiled figure is not Pulad, it must be his go-between’.

3.2.6b.4b Present-future conjectural

[me-raft-agī]-st ‘might be going, about to go’

Constructed with the imperfective prefix *me-*, this form expresses a conjecture about a potential or a current (habitual or iterated) action:

pagoh [me-omad-agī]-st
 tomorrow IPFV.come.CNJECT-COP.3s
 ‘he’ll probably come tomorrow’;

gazeta-i muallim-on-ro ū har rīz [me-xond-agī]-st
 journal-EZ teacher-p-DO he each day IPFV.read.CNJECT-COP.3s
 ‘I’ll bet he reads the Teachers’ Gazette every day’.

3.2.6b.4c Present progressive conjectural

[raft-a istod-a]-gī-st- ‘might be going’

This progressive form corresponds to the indicative present progressive:

balki vaq dar yagon čariš ba xun-u xok-aš oyušt-a šud-a,
 perhaps he in some meadow in blood-and dirt its welter.PART become.PART
jon [kand-a istod-agī]-st
 soul tear.PART PROG.CNJECT-COP.3s
 ‘quite likely, in some meadow, weltering in gore and grime, he lies dying’.

Dialect reflexes of these forms are subject to various contractions, e.g. *kar-soda-y < kard-a istod-a ast* ‘might be doing’ *raf-sod-ag-em < raft-a istod-ag-ī-am* ‘might be going’.

3.2.6b.4d Pluperfect counterfactual conjectural

[raft-agī] me-bud ‘(if) he might have gone’

This form extends the range of the conjectural to parallel the pluperfect in its counterfactual function, further marked by the prefix *me-* as a non-indicative marker (see also Imperfect section 3.2.2.2d).

agar čašm-i Yodgor-ro andeša-i oyanda-i sioh torik
 if eye-EZ Y.-DO thought-EZ future-EZ black dark
na-[kard-agī] me-bud
 not make.CNJECT IPFV.be.PT.3s
 ‘if the thought of a black future had not darkened the vision of Yadgar . . .’ ~
 ‘if Yodgor’s vision had not been clouded by the prospect of a black future . . .’

3.2.6b.5 Synopsis of Tajik verb forms

The following is an overview of the innovated forms. Note that several forms are tense neutral, that is, they may refer to both a present and past context.

TABLE 8.11: TAJIK INNOVATED VERB FORMS, 3s

<i>Progressive forms</i>		
Present progressive	<i>raft-a istod-a-ast</i>	
Past progressive	<i>raft-a istod-a bud</i>	
(see also evidential, subjunctive and conjectural forms)		
<i>Evidential mode</i>		
Evidential durative	<i>me-raft-a-ast</i>	(tense neutral)
Evidential progressive pluperfect	<i>raft-a istod-a bud-a-ast</i>	(tense neutral)
<i>Subjunctive forms</i>		
Durative perfect subjunctive	<i>me-raft-a boš-ad</i>	(tense neutral)
Present progressive subjunctive	<i>raft-a istod-a boš-ad</i> (rare)	(tense neutral)
<i>The conjectural mood</i>		
Present-future conjectural	<i>me-raft-agī-st</i>	
Past conjectural	<i>raft-agī-st</i>	
Pluperfect counterfactual conjectural	<i>raft-agī me-bud</i>	
Present progressive conjectural	<i>raft-a istod-agī-st</i>	

3.2.7 Synopsis of Persian and Tajik verb systems

Table 8.12 shows the inherited basic verbs forms (in italics) and the innovative forms (non-italic) and their categories; several forms are listed more than once according to their functions. The arrangement of the table highlights the fact that in Tajik the progressive and the evidential mode have become major parameters of the system, intersecting with the tense parameter, present and past, here represented by pairs of tense forms in each subset. Note also that the conjectural forms constitute a subset together with the imperfective conditional (imperfect form in its counterfactual function).

4 SYNTAX AND USE OF FORMS I: NOUN PHRASE STRUCTURE

The NP may consist minimally of a single nominal, which may be a noun, independent pronoun, infinitive: *xānd-an āsān ast/xond-an oson ast* ‘reading is easy’, or any item that is nominalized or used in nominal function.

4.1 Coordination

4.1.1 Single and paired conjunctions

The two basic conjunctions are: (1) coordinating *valva*, enclitic form *-o, -vol-u, -yu,-v-u*, in open and closed compound, respectively, e.g. *man va tolman va tu* ‘I and you’ vs. *man-o tolman-u tu* ‘I-and-you’, so in fixed compounds such as number, e.g. *bist-o dolbist-u du* ‘22’; (2) disjunctive *yāl/yo* ‘or’.

Paired coordinators of homogeneous NPs (and VPs) include: *valva . . . valva* ‘both . . . and’, *hamlham . . . hamlham* ‘both . . . and’, *čelčī* ‘whether . . . whether, both . . . and’; *yāl/yo . . . (va) yāl(va) yo* ‘either . . . or’; *xāhl/xoh . . . xāhl/xoh* ‘be it . . . or, whether . . . or’; *nalna . . . (va) nal(va) na* ‘neither . . . nor’:

TABLE 8.12: SYNOPSIS OF VERB SYSTEMS, 3s

Stem forms: *rav-*, *raft-*, *raft-a* 'go, leave'; = same in form; * rarely used

PROGRESSIVE RAFTA ISTODA	BASIC	EVIDENTIAL RAFTA
IMPERATIVE, 2p		
PR	<i>rav-ed</i>	
PT	<i>rafta boš-ed</i>	
SUBJUNCTIVE		
PR * <i>[rafta istoda]</i> boš-ad	<i>rav-ad</i>	me- <i>[rafta]</i> boš-ad =
PT	<i>rafta boš-ad</i>	me- <i>[rafta]</i> boš-ad =
IMPERFECTIVE		
<i>Indicative</i>		
PR <i>[rafta istoda]</i> -ast	<i>me-rav-ad</i>	me- <i>[rafta]</i> -ast =
PT <i>[rafta istoda]</i> bud	<i>me-raft</i>	me- <i>[rafta]</i> -ast =
<i>[rafta istoda]</i> buda-ast EV		<i>[rafta]</i> istoda buda-ast PROG
<i>Non-Indicative</i>		
Conjectural - <i>agī-st-</i>	Counterfactual	Conjectural - <i>agī-st-</i>
PR * <i>[rafta istoda]</i> gī-st	<i>me-raft</i> =	me- <i>[rafta]</i> gī-st
PT	<i>me-raft</i> =	<i>[rafta]</i> gī-st
PERFECTIVE		
Preterite/Aorist	<i>raft</i>	<i>[rafta]</i> -ast =
RESULTATIVE-STATIVE		
<i>Indicative</i>		
PR PF	<i>rafta ast</i> =	<i>[rafta]</i> -ast =
PT PF	<i>rafta bud</i>	<i>[rafta]</i> buda-ast
<i>Non-Indicative</i>		
	Subj., Counterf.	Conjectural - <i>agī-</i>
PR PF	<i>rafta boš-ad</i> SBJ	
PT PF	<i>rafta (me-)bud</i> C-F	<i>[rafta]</i> gī me-bud

PERSIAN

magar [na man na to] lāyeq-e in kār hast-im?
 but not I not you/s fit-EZ this work EX-I
 'but are neither I nor you are fit for this task?';

TAJIK

ba in koyaz-(h)am [na dast-i to] ras-ad, [na dast-i man]
 with this paper also not hand-EZ you.s (SBJ.)reach-3s not hand-EZ me
 'and let [neither my hand] [nor yours] touch this paper'.

The contrastive construction *na tanhāl/na tanho . . . balkelbalki ~ hamlham* 'not only . . . but also' may link NPs, VPs or clauses:

[na tanhā] man, [balke] doxtar-am [ham] u-rā didl
[na tanho] man, [balki] duxtar-am [ham] ū-ro did
 not alone I but daughter-my also he-DO see.PT.3s
 'not only I, but my daughter too saw him'.

Adjectives are similarly concatenated, or are connected by *-e/i* if they qualify the preceding N + ADJ phrase:

<i>ketāb-e bozorg o mofid</i>	<i>zan-e zibā-ye mahru</i>
<i>kitob-i buzurg u mufid</i>	<i>zan-i zebo-i mahrū(y)</i>
book-EZ big and useful	woman-EZ beautiful-EZ moon-faced
'a big and useful book';	'a beautiful moon-faced woman'.

4.1.2 Concatenation and group inflection

Nouns and noun phrases (and any other constituents) may be concatenated, typically by adding *-o/-u*, to each constituent except the last, or to the penultimate one alone, or without addition. Grammatical markers such as the direct object *-rā/-ro* are normally added only to the last of several closely coordinated group of nominals:

[u(-v)-o mā-o šomā]-rā mo'ayyan kard-and
[vay-u mo-v-u šumo]-ro ta'in kard-and
 he-and we-and you.p-DO identified do.PT-3p
 'they designated [him and us and you];

[ketāb, daftar va qalam]-rā ru-ye miz gozāšt-am
[kitob, daftar va qalam]-ro rū-i stol mond-am
 book notesbook and pen-DO face-EZ table place.PT-1s
 'I laid [the book, notebook and pencil] on the table'.

4.2 Constituents following the head: *Ezāfe*

Dependent nominals, including adjectives, nouns and nominal phrases, follow the head noun and are linked in the EZ construction by the connective clitic *-e/i*, as described under Morphology, allowing multiple embedding.

4.2.1 Basic structure

4.2.1.1 NP levels

The EZ noun phrase has two levels and two basic types, depending on the syntactic status of the complement:

(1) adjectival, noun + adjective:

ketāb-e bozorg
kitob-i buzurg
 book-EZ big
 'big book'.

(2) compound, noun (+ adjective) + noun (+ adjective):

ketāb-e pedar
kitob-i pidar
 book-EZ father
 'book of father/father's book'.

(3) phrasal: noun + adverbial phrase:

qarn-hā-ye piš az eslāmī
sada-ho-i peš az islom
 centuries-EZ before from Islam
 ‘the centuries before Islam’.

(4) clausal: noun + nominalized phrase or clause, prominently in Tajik:

pesar-e [az madrase gorixt-e]
pisar-i [az madrasa gurexta-e]
 boy-EZ from school flee.PART
 ‘boy fled from school’, i.e. ‘truant boy’.

4.2.1.2 NP clitics

There are three sets of NP clitics which attach according to level:

(1) the indefinite *-il-e*, (2) the personal enclitics, (3) the postposition *-rāl-ro*. The former two function on the level of the adjectival noun phrase, the latter functions on the level of the NP as a whole.

(1) Thus, as a rule, *-il-e* is attached to the singular and plural N, but to the adjective in an adjectival EZ:

<i>so’āl-il</i>	<i>so-āl-hā-il</i>	<i>[so’āl-e xub]-i</i>	<i>[so’āl-hā-ye xub]-i</i>
<i>suwol-e</i>	<i>suwol-ho-e</i>	<i>[suwol-i xub]-e</i>	<i>[suwol-ho-i xub]-e</i>
‘a question’,	‘some/certain questions’;	‘a good question’,	‘some, certain good questions’.

In case of several adjectives it may be attached to the noun directly, without *-il-e*:

Esfahāni-hā [mardom-ān-e besyār mehmāndust]
 Isfahani-p people-p-EZ very hospitable
va mehrabān, samimi va širin-zabān]-i and
 and kind sincere and sweet-tongued-INDEF are

vs.

Esfahāni-hā [mardom-ān]-i hast-and [besyār mehmāndust]
 Isfahani-p people-p-INDEF are very hospitable
va mehrabān, samimi va širin-zabān]
 and kind sincere and sweet-tongued
 ‘the Isfahanis are (a) very hospitable, kind, sincere and sweet-tongued people’.

(2) The pronominal clitics are similarly attached:

<i>so’āl-aml</i>	<i>so-āl-hā-yaml</i>	<i>[so’āl-e xub]-at</i>	<i>[so’āl-hā-ye xub]-at</i>
<i>suwol-am</i>	<i>suwol-ho-yam</i>	<i>[suwol-i xub]-at</i>	<i>[suwol-ho[-ixub]-at</i>
‘my question’,	‘my questions’;	‘your good question’,	‘your good questions’.

These two sets are mutually exclusive.

(3) The postposition *-rāl-ro* is clitic to the NP as a whole in rightmost final position irrespective of its internal structure, be it a single N, simple NP or an extended NPP as a whole.

The basic EZ structures are thus as follows (N = noun, A = adjective, NPP = compound noun phrase; 3s -aš exemplifies personal clitics; all with nouns in the singular):

TABLE 8.13: CLITIC ATTACHMENT

N	=	<i>ketāb</i> [-aš or -i] [-rā]/ <i>kitob</i> [-aš or -e] [-ro]					
		book his, her a	DIROBJ				
		'his, her book' or 'a book'.					
NA	=	<i>[ketāb-e bozorg]</i> [-aš or -i] [-rā]/ <i>[kitob-i buzurg]</i> [-aš or -e] [-ro]					
		book-EZ large his, her a	DIROBJ				
		'his, her large book' or 'a large book'.					
NPP	=	<i>ketāb-e bozorg-e</i> [mo'Allem-e javān] [-aš or -i] [-rā]/ <i>kitob-i buzurg-i</i> [mu'allim-i javon] [-aš or -e] [-ro]					
		book EZ large EZ teacher EZ young his, her or a	DIROBJ				
		'the large book of his young teacher'.					

4.2.1.3 Alternate attachment of indefinite marker

The noun alone (or with plural marker) without any determiner, modifier, or enclitic, is either non-specific or definite, and can be used either in a generic or non-generic sense. Without a determiner they are usually indefinite.

While *-i/-e* is usually added to the end of the EZ phrase, in the adjectival EZ it may optionally be attached to the head noun or NP, in which case the EZ particle is not used:

<i>kas-e</i>	<i>digar-il</i>	or	<i>kas-i</i>	<i>digarl</i>
<i>kas-i</i>	<i>digar-e</i>		<i>kas-e</i>	<i>digar</i>
person-EZ other INDEF			person-INDEF other	
'another person'			'someone else'.	

(1) *-i/-e*: As discussed under Morphology (section 3.1.1.2), both singular or plural nouns may be followed by the indefinite (specific or non-specific) clitic:

<i>ketāb-ilkitob-e</i>	<i>ketāb-hā-ilkitob-ho-e</i>
'some ~ any book' or	'some ~ any books' or
'a book';	'some ~ certain books'.

(2) *yek/yak*. Alternatively, the noun may be preceded by the quasi-indefinite article *yek/yak* 'one, a' and may further be followed by the indefinite enclitic *-i/-e*. The admission of plural and the indefinite enclitic distinguishes the quasi-article *yek/yak* from its role as numeral:

Singular:	Plural:
<i>[yek] fenjān[~i-]</i>	<i>[yek] fenjān-hā [~i-]</i>
<i>[yak] payola [~e]</i>	<i>[yak] payola-ho [~e]</i>
'a (certain) cup';	'some (certain) cups'.

Tajik also allows plural without enclitic *-i*: *[yak] piyola-ho*.

4.2.2 Types of noun phrases and complements

4.2.2.1 Overview

Fundamental for identifying the function of the *ezāfe* construction is the fact that it encompasses not only nouns, but also adjectives. Heads may be nouns, including verbal nouns, as well as adjectives and participles, prominently those belonging to the Arabic loan component and, rarely, a pronoun. Complements may similarly be nominal and adjectival, or participial, or may be sentential nominalizations. (A good number of the latter is further reduced to quasi-compound words; see section 7.1 Word formation.)

<i>kār-e</i>	<i>pesar, u, sāxtan</i>	<i>kār-e</i>	<i>moškel, anjām-šodel</i>
<i>kor-i</i>	<i>pisar, vay, soxtan</i>	<i>kor-i</i>	<i>muškil, anjom-šuda</i>
work-EZ boy,	he, building	work-EZ difficult	having been completed
‘the work of the boy, him, building’;		‘difficult, completed work’.	

In Tajik, moreover, the nominal EZ in combination with verbal nouns, especially the infinitive and other non-finite verb forms, has greatly extended its range of application (arguably under the influence of Uzbek), and may condense into a single sentence what would appear more naturally in Persian or English as a subordinate VP (see section 6.5.4 Tajik Turkic-type nominal clauses).

4.2.2.2 Nominal heads and valence relationships

The nominal EZ can be seen as essentially the nominalization of underlying case relationships, or neutralized valences:

Predicate	<i>āsmān-e ābilosmon-i obi</i> ‘blue sky’;
Event	<i>ruz-e enqelābrūz-i inqilob</i> ‘the day of revolution’ ~ ‘revolution day’;
Possessor	<i>ketāb-e Hasanlkitob-i Hasan</i> ‘the book of Hasan’ ~ Hasan’s book’;
Agent	<i>kār-e mardomlkor-i mardum</i> ‘the work of people’;
Patient	<i>qatl-e Hoseynlqatl-i Husayn</i> ‘the murder of Hoseyn’;
Purpose	<i>dāru-ye geripldāru-yi gripp</i> ‘influenza medicine’;
Goal	<i>rāh-e Tehrānlroh-i Dušanbe</i> ‘the road of/to Tehran, Dushanbe’ ~ Tehran, Dushanbe road’;
Location Time	<i>mardom-e injā, emruzlmardum-i in jo, imrūz</i> ‘people (of) here, of today’;
Origin	<i>ahl-e Tehrānlahl-i Dušanbe</i> ‘inhabitant of Tehran, Dushanbe’;
Source, Cause	<i>āb-e češmelob-i čašma</i> ‘water of well’ ~ well-water’, <i>bim-c mowj</i> ‘fear of the waves’;
Substance	<i>gombad-e talālgunbad-i- talo</i> ‘dome of gold’;
Element	<i>anbuh-e sā’el-ānlanbūh-i so’il-on</i> ‘crowd of pilgrims’;
Part	<i>do nafar-e ān-hāldu nafar-i on-ho</i> ‘two (persons) of them’.

Predication includes the following relationships, also marked by *-el-i*:

Genus-species	<i>gol-e sorx/gul-isadbarg</i> ‘rose (flower)’;
Class-gender	<i>dust-e doxtar, pesar/dūst-i duxtār, pisar</i> ‘girl-, boyfriend’;
Class-name	<i>asb-e nar, mādelasb-i nar, moda</i> ‘male, female horse’; <i>šahr-e Tehrān/šahr-i Dushanbe</i> ‘Tehran/Dushanbe city ~ city of Tehran/Dushanbe’;
Address-last name	<i>āqā-ye, xānom-e Javādiljanob-i, xonum-i Javodi</i> ‘Mr., Mrs. Javadi’
Address-first name	<i>āqā-ye Ahmadljanob-i Ahmad</i> ‘Mr. Ahmad’, <i>xānom-e Maryam/xonum-i Maryam</i> ‘Mrs. Maryam’;
First-family name	<i>Ahmad-e Javādil/Ahmad-i Javodi</i> ‘Ahmad Javadi’, <i>Maryam-e Javādil/Maryam-i Javodi</i> ‘Maryam Javadi’;
Series-number	<i>dars-e panj, panj-om/dars-i panj, panj-um</i> ‘lesson five, fifth lesson’; <i>sā’at-e selso’at-i se</i> ‘three o’clock’.

A typical result of neutralization may be overt identity: *dust-e doxtar/dūst-i duxtār* may be both predicative ‘girlfriend’ and ‘a friend who is a girl’, as well as possessive ‘a friend of the girl’. Similarly, *ešq-e dust-ān/išq-i dūst-on* may be both subjective ‘love by the friends’ and objective ‘loving (the) friends’.

4.2.2.3 Possessor

4.2.2.3a Preposed possessor

The possessor can be topicalized by inversion and anaphoric pronominal clitic:

<i>esm-e to/ism-i to</i> name-EZ you.s	<i>ketāb-e Hasan/kitob-i Hasan</i> book-EZ H.
>>	
<i>[to] esm [-at]/tu ism-at</i> you.s name-your-s ‘your name’;	<i>[Hasan] ketāb [-aš]/Hasan kitob-aš</i> H. book-his ‘Hasan’s book’.

Syntactically such NPs are clause initial:

<i>be-bin-am,</i>	<i>[to] esm[-at]</i>	<i>či-st?!</i>
<i>bin-am,</i>	<i>[tu] ism[-at]</i>	<i>či-st?</i>
SBJ.see.PR-2s you.s name-your-s what-COP.3s		
‘let me see, what’s [your] name?’		

4.2.2.3b Tajik Turkic-type possessor inversion

In the Northern dialects of Tajik, a topical possessor inversion may serve to bind a complex NP such as a reduced relative clause. This construction uses the postposition *-a* (<*-ro*) which at earlier stages of Persian had also a dative function, but has become the unmarked construction, essentially a Turco-form version of the EZ (see Postposition *-rā, -ro* section 3.1.4.2):

<i>kitob-i in mard</i>	‘book of this man’ (Tajik) >
book-EZ this man	
<i>in mard[-a] kitob-aš</i>	‘this man-to book-his’.
this man[-RA] book-his	

4.2.2.4 Adjectives and ordinals

4.2.2.4a Superlatives and ordinals

These two canonically precede the head:

<i>behtar-in ketābl</i>	<i>avval-in ruz-e tābestānl</i>
<i>bihtar-in kitob</i>	<i>avval-in rūz-i tobiston</i>
‘the best book’;	‘the first day of summer’.

In Persian, superlative adjectives always precede the head, as do ordinals when in focused position.

<i>bozorg-tar-in māšin</i>	vs. <i>māšin-e bozorg-tar</i>	‘the largest car’ vs. ‘the larger car’;
<i>dovv-om-in sālgard</i>	~ <i>sālgard-e dov-om</i>	‘the second anniversary’.

In Tajik, (1) superlatives, (2) ordinal adjectives in *-in*, (3) simple ordinal adjectives in *-um*, (4) adjectives in *-ngī*, and (5) phrases with the postposition *barin* may either precede or follow the head:

<i>kalon-tar-in šahr</i>	~ <i>šahr-i kalon-tar-in</i>	‘the biggest town’;
<i>oxar-in bor</i>	~ <i>bor-i oxar-in</i>	‘the last time’;
<i>du-yīm poyezd</i>	~ <i>poyezd-i du-yūm</i>	‘the second train’;
<i>dina-ngī rūz-noma</i>	~ <i>rūz-noma-yi dina-ngī</i>	‘yesterday’s paper’;
<i>man barin šaxs</i>	~ <i>šaxs-i man barin</i>	‘a person like me’.

4.2.2.4b Preposed affective adjectives

When used affectively or evaluatively, adjectives may precede the noun, notably when exclamatory. When evaluative, Persian often has indefinite *-i*:

<i>dust-e</i>	<i>bičāre-ye manl</i>	<i>pesar-e xub(-i)</i>	<i>astl</i>
<i>dūst-i</i>	<i>bečora-i man</i>	<i>pisar-i xub</i>	<i>buda ast</i>
friend-EZ	poor-EZ I	boy-EZ good(INDEF)	is/seems to be
>		>	
[<i>bičāre</i>]	<i>dust-e manl</i>	[<i>xub</i>] <i>pesar-i astl</i>	
[<i>bečora</i>]	<i>dūst-i man</i>	[<i>xub</i>] <i>pisar bud-a ast</i>	
poor	friend-EZ I	good boy	is/appears to be
‘my poor	friend!’;	‘he’s/seems to be a good boy’,	non-evidential-mirative.

4.2.2.4c Pronouns as heads

Pronominal heads of EZ constructions are marginally productive and mostly confined to deplorable qualifications: *man-e bičāre*, *falak-zademan-i be-čora*, *falak-zada* ‘poor, fate-stricken me’, etc.

4.2.2.5 Nominalized complements

4.2.2.5a Verbal nouns as heads

Verbal nouns, in both Persian and Arabic, behave essentially like nouns, those derived from transitive verbs exhibiting ambivalence between agent and patient, i.e. subject and direct object (here using roughly equivalent Persian infinitives and Arabic verbal nouns):

Transitive	Intransitive
<i>bar gozid-an-e ~ entexāb-e Hasan</i>	<i>rasidan-e ~ vorud-e vazir</i>
<i>bar guzidan-i ~ intixob-i Hasan</i>	<i>rasidan-i ~ vurud-i vazir</i>
‘choosing ~ choice-EZ H; both ‘choice by Hasan’ and ‘choosing Hasan’.	‘the arriving ~ arrival-EZ minister’.

Archaizing diction allows the direct object to be attached (e.g. in sub-headings):

<i>košt-an-e Rostam [Sohrāb-rā]</i>	
<i>kuštān-i [Suhrob-ro]</i>	
kill.INF-EZ R. S.-DO	
‘Rostam’s killing of Sohrab’, lit. ‘the Sohrab killing of Rostam’.	

4.2.2.5b Compound participles and adjectives as heads

A few adjectives and a relatively limited set of participles forming compound verbs behave similarly to nouns and verbal nouns in neutralizing valence relationships. They consist mainly of Arabic active and passive participles. Common examples include the following (see Ahadi 2001: 105–109):

Persian adjectives	
<i>dočār beldučor ba</i> ‘affected by’:	<i>dočār-e eškāl ast</i> ‘is encountering difficulties’;
<i>šifte az/šeftra az</i> ‘enamored by’:	<i>šifte-ye češm-hā-yaš ast</i> ‘is enamored by her eyes’;
Arabic participles	
<i>motavallid dar/mutavallid dar</i> ‘born in’:	<i>āšeq bālošiq bo</i> ‘being in love with’:
<i>motavallid-e Irān ast</i>	<i>āšeq-e doxtar ast</i>
<i>mutavallid-i Tojikiston ast</i>	<i>ošiq-i duxtar ast</i>
‘was born in Iran/Tajikistan’;	‘is in love with the girl’;
<i>-rā monker/ro munkir</i> ‘denying s.th.’:	<i>montazer-elmuntazir-i</i> ‘waiting for’:
<i>monker-e xodā ast</i>	<i>montazer-e dust ast</i>
<i>munkir-i xudo ast</i>	<i>muntazir-i dūst ast</i>
‘is a denier of God’;	‘is expecting the friend’.
<i>tahvil beltahvil ba</i> ‘handing over to’:	
<i>tahvil-e u kard</i>	
<i>tahvil-i vay kard</i>	
‘he handed it over to him’.	

A distinct subtype are reduplicated adjectives:

xub-e xub/xub-i xub ‘very good’, lit. ‘good of good’, etc.

4.2.2.5c Participial phrases as complements

(See also section 6.3.7 Tajik nominalized relative clauses)

Participial relative clauses as complements tend to be idiomized compounds and adjectival, particularly in Persian:

<i>mard-e</i> [dast-šekast-e]l	<i>ādam-e</i> [kešti-šekast-e]l
<i>mard-i</i> [dast-šikast-a]	<i>odam-i</i> [kašti-šikast-a]
man-EZ hand-break.PART	person EZ ship-break.PART
‘the man with a broken hand’;	‘ship-wrecked person’;
but not * <i>mard-e māšīn-suxte</i> ‘the man with a burned motor’.	

Prepositions can be retained:

forsat-hā-ye [az dast raft-e]l
fursat-ho-i [az dast raft-a]
 opportunities EZ from hand give.PART
 ‘missed opportunities’, lit. ‘given out of the hand’;

ham-ān pesar-e [az madrase gorixt-e]-i *bud-am ke...*
 very same boy-EZ from school flee.PART-INDEF be.PT-1s who ...
 ‘I was again that same truant boy who ...’

4.2.2.5d Tajik nominalized relative clauses

In Tajik copular relative clauses may be nominalized as a possessive EZ construction:

mard(-e) ki mū-yaš safed ast ‘a man whose hair is white’ >
mard-i [mū-yaš safed]
 the man-EZ hair his white
 ‘the white-haired man’.

4.2.2.5e Tajik constructions with participle in -agī

While in Persian participial constructions are limited, Tajik has constructions with the participle in *-agī* that are fully productive (see also section 6.3.7.1):

korgar-on-i [az Evropa boz-gašt-agī]
 worker-p-EZ from Europe return.PARTII
 ‘workers having returned from Europe’,
 cf. Persian: *dānešju-yān-e* [az farang bar-gašt-e]
 ‘students having returned from Europe’.

4.3 Constituents preceding the head noun

4.3.1 Overview

Determiners including demonstratives, indefinite and interrogative pronouns, quantifiers and numerals with or without classifiers and their interrogatives, and others precede the head noun:

DEM ~ INDEF-INTERROG – CLASS – NUM – NOUN

Examples are: *ān mardlon mard* ‘that man’; *har haftelhar hafta* ‘every week’; *če now’lčī xe?* ‘what sort?’; *do kilo guštldu kilogramm gūšt* ‘two kg of meat’, etc. (see also examples under Morphology).

4.3.2 *Classifiers*

4.3.2.1 Numerative classifiers

Numerative classifiers constitute a small set of terms that follow the cardinal number or the indefinite and numeral interrogative *čandlčand* ‘several; how many?’ and precede the head nominal without any connective marker: Numeral – Classifier – Noun.

In Persian the head noun is in the singular following the singular rule after numbers, while Tajik allows the plural as an option when the numerand denotes a conventionally or contextually defined group. A classifier is obligatory in the absence of a head nominal, but optional otherwise. The choice of these classifiers is conditioned by the semantic class of the counted nominal, with the basic distinction of human and non-human.

The general semantically unmarked classifier is unstressed *tā/to* ‘unit, item’. In Persian, there is the distinction of plurality vs. singularity: *tā* follows numbers larger than one, *do tā pesar, medād* ‘two boys, pencils’, *čand tā pesar, medād* ‘several boys, pencils’ and ‘how many boys, pencils’, as opposed to *dāne* ‘grain’ for single non-human items, *yek dāne medād* ‘a single pencil’.

In Tajik, *to* (colloquial *ta*) has no such restrictions and is widely used for all classes of numerands and all numbers: *yak to zan* ‘one [item] woman’, *sad to kurta* ‘a hundred shirts’, while *dona* ‘grain’ is the classifier for smallish inanimate objects.

Most semantically distinguished classifiers are the same in Persian and Tajik, which once constituted an extensive series; some of those still in use are the following:

(1) Human: *nafar/nafar* ‘individual’ (Ar.) and *tan* ‘body, person’, *haft nafar pirezanihaft nafar kampir* ‘seven old women’; *panj tan/panj tan* ‘five (bodies)’; Tajik plural numerand of contextually definite group:

az hamin šaš nafar-on faqat ba du nafar
 from this-same six persons only to two person
dast na-me-ras-on-em
 hand not-IPFV.reach.CAUS.PR-1p
 ‘of these six men we will leave unharmed only two’.

(2) Prominent part of whole, larger animals: (a) *sar/sar* ‘head’ or (b) *ra’s/ra’s* (Ar.), *se sar boz/se sar buz* ‘three head of goats’.

(3) Prominent part for whole, inanimate: (a) houses, *bāblbob* ‘door’ (Ar.), *dah bāb xānel dah bob xona* ‘group of ten houses’; (b) books, *jeld/jild* ‘skin, book cover’ (Ar.), *se jeld ketāblse jild kitob* ‘three volumes’.

(4) Small-sized objects, including small animals, plants, and other smallish things: (a) *dānel dona* ‘grain’ *dah dona bodring dih-ed* ‘give (me) ten cucumbers’, *čand dona anor* ‘several pomegranates’ and ‘how many pomegranates?’, or (b) *adad* ‘number’ (Ar.), *se adad mix* ‘three nails’.

(5) Shape: *taxte* ‘flat board’, etc., *yek taxte qāli* ‘one (item) ‘carpet’.

(6) Small sets: (a) Pers. *dast* ‘hand’, and *yek dast lebās* ‘one suit’, *se dast sandali* ‘set of three chairs’; (b) small groups (animal, people, objects): Pers. *daste* ‘handful’, mostly with EZ, *yek daste(-ye) gol* ‘a bunch of flowers’; Tajik has *dasta* for both (a) and (b) without and with EZ, *yak dasta gul, dasta-i jinoyatkor* ‘a criminal group’.

(7) Buildings, apartments: *dastgāh/dastgoh* ‘compound unit, set’, *se dastgāh emārāt/se dastgoh imorat* ‘three buildings, flats’.

The classifier is obligatory only if the number is mentioned without the numerand: *čand (tā) ketāblčand (to) kitob – se tā/se to* ‘how many books – Three’, *šomā čand nafar bud-id?!šumo čand nafar bud-ed? – se nafar/se nafar* ‘how many were you? – Three’.

Numbers and classifiers may also be joined with a dependent nominal or pronouns, frequently enclitic: *har do-(tā)-yaš qašang ast* *har du-(to)-yaš naʔz ast* ‘both of them’, lit. ‘of it’, ‘are nice’, *har se tā-yešān* *har se to-yašon* ‘all three of them’, *yek-dāne-yaš-rā bi-yandāz/lyak-to-yaš-ro parto* ‘toss me one of them’.

4.3.2.2 Mensurative classifiers

Mensurative classifiers identifying weight, length, and other measures designate a portion of a mass and combine freely with mass nouns. Syntactically they behave like numerative classifiers: *do fenjān čāy/du payola čoy* ‘two cups (of) tea’, *čahār-o nim kilo sib-zaminilčoru nim kilo kartoška* ‘two and a half kilo (of) potatoes’, *panjāh kilometr rāhl panjoh kilometr roh* ‘fifty kilometers (of) way, road’.

Mensurative noun phrases contrast with partitive EZ constructions: *se qatre xun/se qatra xun* ‘three drops of blood’ vs. *se qatre-ye xun/se qatra-i xun* ‘three drops of the, that blood’.

4.3.2.3 Kind and manner classifiers

Kind and manner classifiers likewise behave like numerative classifiers. These include *now/nav* ‘kind, sort; breed, species’ way, method, manner, mode’: Pers. colloq. *tovr/taʔwr* ‘sort, kind, type; manner, style’; Pers. *jur*, ‘sort, kind, variety; manner’; *raqam/raqam* ‘number’; ‘type, kind, sort’: *če now parandelči naw paranda* ‘what kind of bird?’; Pers. *če-toʔwr ādam-i ast?* ‘what kind of person is he?’; Pers. *in jur miz* ‘this kind of table?’; *in če raqam ketāb astlin čī raqam kitob ast* ‘what type of book is this?’.

5 SYNTAX AND USE OF FORMS II: CLAUSE STRUCTURE

5.1 Order of constituents

The canonical sequence in main, coordinate and dependent clauses is:

(SUBJECT) – (OBJECT) – VERB.

The parentheses indicate that both subject and object may be lexically omitted since the person of the subject is incorporated in the verb, and the object of transitive verbs may be contextually implied.

Thus the simplest clause may comprise a single finite verb:

<i>āmad-andl</i>	<i>dozdid-andl</i>
<i>omad-and</i>	<i>duzdid-and</i>
‘they’ve come, are here’;	‘they stole it’.

With three-valence verbs such as “give”, the definite or specific direct object generally precedes the beneficiary (indirect object), but follows it in preverbal position if indefinite. Note that in the preverbal position of the unmarked object *ketāb/kitob* exemplifies the generic function of unspecified nominals in preverbal position as part of compound verbs:

<i>(Ahmad) [ketāb-rā] be man dād</i>	<i>(Ahmad) be man [ketāb] dād</i>
<i>Ahmad [kitob-ro] ba man dod</i>	<i>Ahmad ba man [kitob] dod</i>
‘Ahmad gave me the book’;	‘Ahmad gave me a book, books.’

Word order is flexible and allows for “scrambling” (particularly in spoken registers). Typically, such re-ordering requires non-canonical marking patterns involving shifts of stress (and pauses). The more frequent and less marked of such patterns are discussed in the following.

5.1.1 *OSV and OVS*

5.1.1.1 Object raising

For emphasis on agent OSV may occur:

sag-rā man košt-aml
sag-ro man kušt-am
 dog I killed (it)
 ‘I was the one who killed the dog’.

5.1.1.2 Concomitant subject lowering

For emphasis on the patient, or the activity, OVS may occur with concomitant subject lowering:

sib-rā ne-mi-xar-id šomā?!
seb-ro na-me-xar-ed šumo?
 the apple-DO not-IPFV.buy.PR-2p you.p
 ‘aren’t you going to buy any apples?’

Note the contrastive use of the direct object marker *-rā/-ro* even with indefinite non-specific objects, approximately ‘as to the apples there (in addition to yoghurt, etc.)’.

5.1.2 *Adverbials*

Adverbs of time and sentence adverbials are often placed first in the sentence. The unmarked order of other is:

TIME – MANNER – LOCATION/DIRECTION,

and is determined by the valency of the verb, such as the means and direction of “going” in the following example:

ma’mul-an, mā har ruz bā otobus be madrase mi-rav-im
odat-an, mo har ruz bo avtobus ba madrasa me-rav-em
 usually we every day with bus to school we go
 ‘we usually go to school every day by bus’.

Otherwise the position of adverbials is relatively free, such as the focused locative preceding the preverbal beneficiary of “giving” instead of following the initial time adverbial:

diruz Ahmad ketāb-rā [dar ketāb-xāne] be man dād
dina-rūz Ahmad kitob-ro [dar kitob-xona] ba man dod
 yesterday Ahmad the book in library to me he gave
 ‘Ahmad gave me the book in the library yesterday’.

Particularly in colloquial speech direction, but not location, follows the verb, usually without the preposition *belba* ‘to’:

raft madrasel
raft madrasa
 ‘she went to the school’.

Adverbials pragmatically connecting and/or qualifying the sentence as a whole (often Arabic loan adverbials and phrases) are typically in first position:

ettefāq-an man ham did-am *zāher-an fardā mi-ā-yad*
ittifoq-an man ham did-am *az aft-āš pagoh me-o-yad*
 ‘incidentally, I saw it, too’; ‘from the looks of things, he’ll be coming tomorrow’.

5.1.3 Questions

Word order in questions is normally the same as for statements.

(1) Yes-no questions, *āyā/oyo*

In yes-no questions interrogativity is established suprasegmentally by means of intonation and pitch contour, which ends at mid-height rather than falling as in statements.

In addition to the primary contour, Common Persian is one of the languages that has developed an overt question marker, the particle: *āyā/oyo* (initial stress) ‘whether, isn’t it that’, which may be followed by the clipped tag question *yā na/yo na* ‘or not’:

āyā to hamrāh mi-ā-i yā na?!/
oyo tu hamroh me-o-yi yo na?
 Q you.s along IPFV.come.PR-2s or not
 ‘are you coming along, or not?’.

In colloquial Tajik, the Uzbek interrogative clitic *-mī* is added to the final word:

seb na-me-xar-ed-mī?
 apple not-IPFV-buy.PR-2p-Q
 ‘won’t you buy some apples?’

(2) Surprise and indignation, *magar/magar*

In questions connoting surprise or indignation (including rhetorical questions) the particle *magar* in the sense of ‘but is it not true that, but’ may precede or follow the subject:

magar u ān-hā-rā na-did-e ast!
ū magar on-ho-ro na-did-a-ast?
 but he those-DO not-see.PART COP.3s
 ‘(but) hasn’t he seen them?’

magar may also imply possibility in questions or statements (for its use as a conjunction ‘unless’ see section 6.5.3.7):

magar šomā ham mi-dān-id!
magar šumo ham me-don-ed?
 whether you.p also IPFV.know.PR-2p
 ‘maybe you know, too?’

(3) Interrogative pronouns and adverbs

Persian has *in situ* constructions. Interrogatives for parts of speech appear in the same slot as in statements when unmarked, but may appear at the beginning when focused (before or after an adverbial phrase):

Subject	Direct object
<i>[ki] emruz mi-ā-yad/</i>	<i>šomā ān-jā [če kas-i]-rā did-id?/</i>
<i>[kī] imrūz me-o-yad</i>	<i>šumo on-jo [čī kas]-ro did-ed?</i>
who today IPFV.come.PR-3s	you.p there what person-DO see.PT-2p
‘who will come today?’;	‘what person, whom did you see there?’
>> <i>emruz [ki] mi-ā-yad?</i>	>> <i>[če kas-i]-rā ān-jā did-id?</i>

Adverb

<i>ketāb-rā [kojā] gozāšt-e i?/</i>	<i>bahār emsāl [key] mi-ā-yad/</i>
<i>kitob-ro [kujo] mond-a i?</i>	<i>lbahor insol [kay] me-o-yad?</i>
book-DO where put.PART COP.2s	spring this year when IP ^F V.come.PR-3s
>> <i>kojā ketāb-rā gozāšte i?</i>	>> <i>key emsāl bahār mi-ā-yad?</i>
‘where did you leave, put the book?’;	‘when will spring come?’

5.1.4 Exclamations

(1) Vocative

The vocative has primary stress on the first syllable, and may be preceded by the particle *eyley* or *āy*:

<i>bačče, injā bi-āl</i>
<i>bača, in jo bi-o</i>
child here come-IMP.2s
‘boy, come here!’.

(2) Selective reference

Surprise and admiration are marked by *če N-i* ‘what (a) N’ and *če ADJ* ‘how ADJ’ (similar to questions):

<i>če havā-ye xub-i!</i>	<i>če xub!</i>
what weather-EZ good-INDEF	what good
‘what a fine weather!’;	‘how good!’.

(3) Tags

These include (a) *-(h)ā(n)* ‘take note, careful, beware’ and (b) *dige*, literary *digar* ‘already, simply’, lit. ‘other’:

<i>goft-am-ā!</i>	<i>goft-am-dige!</i>
‘I said so, didn’t I!’;	‘I’ve already told you!’.

5.2 Definiteness and subject and object marking

5.2.1 Marking of definiteness and specificity

The marking of a NP as definite, indefinite or specific is interdependent with case relations, which in turn are expressed chiefly by syntactic means.

5.2.1.1 Definiteness

The stem form of a noun has two uses, apart from paratactic mention in lists, titles, etc. These are:

(1) Generic, when focus is on the concept it conveys or the generic referent, rather than a specific member of the class or number of items:

<i>dar maqāze [ruznāme] hast?!</i>	<i>ketāb be-xān!!</i>
<i>dar magazīn [gazeta] hast?</i>	<i>kitob xon!</i>
in store newspaper EX.3s	book read.IMP.2s
‘are there newspapers in the store?’; ‘read a book!’, i.e. any book, or books in general;	

<i>pedar-hā-yemān [dehqān] and!</i>
<i>padar-ho-yamon [dehgon]-and</i>
father-p-our peasant COP.3p
‘our fathers are peasants’, lit. ‘peasant’.

(2) Definite, when a referent has been mentioned, or is contextually defined, for instance as a proper noun, independent or clitic pronoun, demonstrative, or equivalent interrogatives. A definite count noun denoting a plural requires a plural marker and, usually, plural verb agreement:

Singular

<i>deh az injā xeyli rāh nist!</i>	<i>pedar-bozorg āmad!</i>
<i>qišloq az injo xele roh nest</i>	<i>bobo omad</i>
village from here much way not.COP.3s	grandfather come.PT.3s
‘the village isn’t far from here’;	‘grandfather has come, arrived’.

Plural

<i>deh[-hā] xeyli rāh nist[-and]!</i>
<i>qišloq[-ho] xele roh nest[-and]</i>
village-p much way not.COP.3s
‘the villages are not far’.

5.2.1.2 Specificity

The specific NP occupies a status between the indefinite NP, the referent of which is known to neither speaker nor listener, and the definite NP, the referent of which is known to both. It constitutes information known to the speaker, but introduced for the first time to the listener; with this form, the speaker signals his ability or willingness to give further information.

A specific NP is marked with *-il-e* and/or *yek/yak*, and as a direct object it requires the marker *-rāl-ro*, while an indefinite noun normally does not.

5.2.2 Marking of subjects

5.2.2.1 Verbal agreement

(1) 3rd person. The Classical Persian norm for 3rd person verbal concord (inherited from Indo-European) is that inanimate subjects have singular agreement and anaphora. In

turn, humans and sentient beings (including higher animals, notably horses) have plural agreement: this is still a valid pattern, but the use of singular and plural is determined by the speaker's perception of the degree of individuation of the subject:

<i>dust-hā āmad-andl</i>	<i>ketāb-hā injā-st ~ (hast-)andl</i>
<i>dūst-ho omad-and</i>	<i>kitob-ho injo-st ~ (hast-)and</i>
friend-p come.PT-3p	book-p here is (EX)-3p
'the friends have arrived-3p';	'the books are here'.

Similarly, collective nouns and pronouns, as well as plural inanimates, may have a singular or plural verb, more commonly in Tajik:

human collective-plural	inanimate plural-singular
<i>mardom dar āmad-andl</i>	<i>šāx-hā-ye deraxt-hā mi-jombidl</i>
<i>mardum daromad-and</i>	<i>šox-ho-i daraxt-on me-jumbid</i>
people out.come.PT.3p	bough-p-EZ tree-p IPFV.shake.PT.3s
'the people came out',	'the boughs of the trees shook',
individually, in small groups;	all together.

(2) Combined persons. Subjects combining 1st person with either or both the other two have 1st plural agreement, subjects combining 2nd and 3rd person have 2p plural agreement:

<i>man o šomā mi-dān-impl</i>	<i>man o Ahmad bā ham raft-im</i>
<i>man u šumo me-don-em</i>	<i>man u Ahmad bo ham raft-em</i>
I and you.p IPFV.know-1s	I and A. together go.PT-1s
'I and you we know';	'I and Ahmad went together'.

(3) Honorific marking

Social convention requires respectful reference to 3rd persons with plural agreement, particularly in their presence. Similarly, the addressee is called by the plural pronoun *šomā/šumo* 'you' with the corresponding verb ending, while the speaker may use the self-deprecating plural *mālmo* 'we' (see section 3.1.3.1 Personal pronouns):

<i>ostād šaxs-e bozorg-i bud-andl</i>	<i>šomā bud-idl mā če mi-dān-impl</i>
<i>ustod šaxs-i buzurg-e bud-and</i>	<i>šumo bud-ed mo ċī me-don-em</i>
professor person-EZ great be.PT-3p	you.p be.PT-1p we what IPFV.know.PR-1p
'the professor was a great man';	'it was you'; 'what do I know?'

5.2.2.2 Copular predicates and ellipsis

As noted under Morphology, the copula form of 'be' in the present tense may be either enclitic or derived from the stems *hast-lhast-*, *bāš-lboš-*, or *nist-lnest-*, and in past tenses from the stem *bud-lbud-*.

In Persian colloquial the present tense forms are obligatory, except with certain question words and phrases: *ku?*, *kojā?* 'where (is): Hasan *ku*, *kojā?* 'where is Hasan?'. Exclamatory rhetorical phrases include deictic forms like *inak*: *inak—parčam!* behold/here (is) the banner!.

In colloquial Tajik, present tense forms may be more readily omitted, especially in questions: *Nodir kujo?* 'where's Nadir?', including basic identification statements: *man tojik* 'I'm a Tajik'.

5.2.2.3 Specified and unspecified subjects

Since the person is expressed in the verb ending, the minimal clause consists of an inflected verb, and there may be no overt subject noun phrase:

sib mi-xar-amlseb me-xar-am
apple IPV.buy.PR-1s
'I'll buy (~ some) apples'.

Unspecified human subjects are expressed by the 3p verb ending (see section 5.5.7.2 Agency and causation), while impersonal inanimate subjects are expressed by the 3s verb ending only: e.g. *bāš-adlboš-ad* 'let it be, OK', *mi-šav-adlme-šav-ad* 'it's possible, one can (do it)', and prominently in indirect verb constructions (section 5.3.2).

'Weather-verbs' such as *bār-*, *bārid-lbor-*, *borid-* 'pour from above' and similar verbs require a subject noun: *bārān*, *barf mi-bār-adlboron*, *barf me-bor-ad* 'it rains, snows' ('it pours down rain, snow').

5.2.3 Marking of direct objects

5.2.3.1 Definite direct objects

Definite objects are obligatorily marked with *-rāl-ro*:

ketāb-rā xarid-aml u-rā did-aml
kitob-ro xarid-am vay-ro did-am
book-DO buy.PT-1s he-DO see.PT-1s
'I bought the book'; 'I saw him'.

5.2.3.2 Indefinite specific and non-specific direct objects

An indefinite noun as direct object is not normally marked with *-rāl-ro*. However, while a *specific* NP is also marked with *-il-e* and/or *yeklyak*, as a direct object it requires the object marker:

(yek) zan[-i]-rā did-aml
(yak) zan[-e]-ro did-am
one woman-INDEF-DO see.PT-1s
'I saw a certain woman';

(yek) ketāb-e xub[-i]-rā xānd-aml
(yak) kitob-i xub[-e]-ro xond-am
one book-EZ good-INDEF-DO read.PT-1s
'I've read a good book'.

Such sentences as a rule require, or imply, following explanatory clauses, typically relative clauses, such as, for the latter example, 'a good book – which I might now identify or describe for you, if you show interest'.

5.2.3.3 Scale of specificity

Specificity is given overtly when a noun or noun phrase is, or contains, a name, personal or demonstrative or interrogative pronoun, and implicitly when a specific or certain item

or set is referred to by the speaker. Accordingly, *-rāl-ro* also attaches to items or sets that are marked by the indefinite *-il-e*, when specific, such as *kas-i-rā did/kas-e-ro did* ‘he saw someone, a certain person’; in comparison:

<i>gorg-rā košt-and</i>	‘they killed the wolf’;
(<i>yek</i>) <i>gorg-i-rā košt-and</i>	‘they killed a certain wolf’;
(<i>yek</i>) <i>gorg-i košt-and</i>	‘they killed some wolf’;
<i>gorg košt-and</i>	‘they killed wolves, were wolf-killing’;
<i>gorg-hā košt-and</i>	‘they killed many wolves’, <i>-hā</i> with emphatic stress.

5.3 Use of pronominal clitics

The enclitic series of the pronouns is used primarily to express attributive possession and its metaphors (see section 3.1.3.2):

ketāb-am/kitob-am ‘my book’, *varaḡ-hā-yaš/varaḡ-ho-yaš* ‘its pages’.

This use includes the Turco-form possessive construction in the northern Tajik dialects discussed above:

in mard[-a] kitob[-aš]
 this man-RA book-his’
 ‘this man’s book’.

In addition, the clitics function both as direct objects and indirect objects, including the experiencer. In Tajik they also function in a possessive construction, including nominalized clauses (see section 6.3.7).

5.3.1 Pronominal direct and indirect objects

The pronominal clitics may also express the direct or indirect object of a verb to which they attach.

(1) direct object, *did-am-etān/did-am-aton* ‘I saw you (*-etān/-aton*)’.

The pronominal clitic may replace the direct object phrase, and is attached to the verb, and in compound verbs usually to the preverbal nominal or to the verb itself (mostly colloquial, see section 5.5.6.2):

[u-rā] did-am/[vay-ro] did-am
did-am[-aš]/ did-am[-aš]
 ‘I saw him’;

Persian	Tajik
<i>[u-rā] bidār kard-am/</i>	<i>[vay-ro] bidor kard-am</i>
<i>bidār[-aš] kard-am/</i>	<i>bidor[-aš] kard-am</i>
<i>bidār kard-am[-aš]/</i>	<i>bidor kard-am [-aš]</i>
‘I woke him up’, <i>bidār kardan/bedor kardan</i> , lit. ‘make awake’.	

In spoken Tajik these forms often elide the initial vowel and add the object marker *-a* (< *ro*): *me-gir-an[-š]* ‘they’ll catch him (*-š*)’; *did-am-t[-a]* ‘I saw you (*-t*)’.

(2) indirect object, *goft-am-aš/guft-am-aš* ‘I told him (*-aš*)’.

These replacements are found with a small number of simple and compound verbs with beneficiary valence:

Persian	Tajik	Persian	Tajik
[be u] goft-aml	[ba vay] guft-am	[be u] komak kard-aml	[ba vay] yorī dod-am
goft-am[-aš]l	guft-am[-aš]	komak[-aš] kard-aml	yorī[-aš] dod-am
‘I said to, told him’;		‘I gave him help, helped him’.	

5.3.2 Indirect experience

Indirect verb constructions express bodily or mental experience as happening to, or overcoming, a person without his or her doing. They are most frequent in casual styles, and appear to be an expanding class.

The verb is a compound of a noun or adjective plus ‘be’ or ‘become’ and a small set of other verbs in the impersonal 3s, while the experiencer is referred to by the personal clitics, at earlier stages of Persian also by the independent personal pronoun plus *-rā* in its dative function.

The source is usually indicated by *az* ‘from’. Some sensations and events can only be expressed that way, while others can be expressed either by an indirect construction or by verbs which imply the active involvement of the speaker, allowing for options. Among the commonly used indirect verb constructions, and their active correspondents, there are prominently those expressing bodily sensations, emotions, and mental activity. The following examples illustrate the pairing of active expressions with indirect construction, and the typical pairs of verbs which are similar to those used in the quasi-active/passive pairs of other types of compound verbs:

Active	Experiential
<i>man gorosne [am]</i>	<i>gorosne [-am] ast</i>
I hungry COP.1s	hungry-to me COP.3s
‘I am hungry’;	‘I feel hungry’;
<i>man bāvar ne-mi-kon[-am]</i>	<i>bāvar[-am] ne-mi-šav-ad</i>
I belief not-IPFV.make.PR-1s	belief-to me not-IPFV.become.PR-3s
‘I don’t believe it’;	‘I cannot believe it’;
<i>qahr kard[-im]</i>	<i>qahr[-emān] šod</i>
quarrel make.PT-1p	quarrel-to usbecome.PT.3s
‘we quarrelled’, lit. ‘made quarrel’;	‘we got to quarrelling’;
<i>[ān-rā] dust dār-im</i>	<i>[az ān] xoš[-am] mi-ā-yad</i>
that-DO friend have.PR-1p	from that well to me IPFV.come.PR-3s
‘we like him, it’, lit. ‘have as friend’;	‘I like it’, lit. ‘from it it comes well’;
	cf. earlier <i>[ma-rā] xoš āmad</i> .
Compound with <i>yād</i> ‘memory’, all ‘I remember’	
<i>[ān-rā] yād dār[-am]</i>	<i>[ān] yād[-am] hast</i>
that-DO memory have.PR-1s	that memory-to me EX.3s
‘I have it in memory’;	‘that is to me in memory’;
<i>[ān-rā] yād mi-ā-var[-am]</i>	<i>[ān] yād[-am] mi-ā-y-ad</i>
that-DO memory IPFV.bring.PR-1s	that memory-to me IPFV.come.PR.3s
‘I bring to memory’;	‘it comes to my memory’;

Typically colloquial is the expression *čē-t-e?* ‘what is the matter with you?’ – *čiz-i-m nist!* ‘nothing!’ (< *čē-at ast* ‘what to you is?’ and *čiz-i-am nist* ‘a thing to me is not’).

5.4 Subject and object complements

Nominal complements implying a property are non-specific (generic) and are unmarked, whether intransitive or transitive. The verb is a verb of state and persistence or change, most typically ‘be’, including copula, the existential *hast-lhast-*, neg. *nist-lnest-*, the subjunctive *bāš-lboš*, past all *būd-lbud-*. Other verbs include: intransitive *mān-*, *māndlmon-*, *mond-* ‘remain’, *šav-*, *šod-lšav-*, *šud-* ‘become’; transitive *kon-lkun-*, *kard-*, *dān-*, *dānest-l don-*, *donist-* ‘know, recognize’, *entexāb k.lintixob k.* ‘elect’, etc.

<i>qabl-an čupān bud-andl</i>	<i>u-rā dust mi-dān-aml</i>
<i>qabl-an podabon bud-and</i>	<i>ū-ro dūst me-don-am</i>
before shepherd be.PT-1p	he-DO friend IPFV.know.PR-1s
‘before they were shepherds’;	‘I consider him a friend’.

The same rules applies to adjectival complements.

However, when modified by a specifying nominal EZ plural marking is obligatory:

<i>qabl-an čupān-hā-ye mā bud-andl</i>	<i>ān-hā-rā dust-hā-ye mā mi-dān-ml</i>
<i>qabl-an podabon-ho-i mo bud-and</i>	<i>on-ho-ro dūst-ho-i mo me-don-im</i>
before shepherd-p-EZ we be.PT-3p	they-DO friend-p-EZ we IPFV.know.PR-1p
‘before they were our shepherds’;	‘we consider them our friends’.

5.5 Complex predicates

The following general characterizations can be made:

The combined morphology and categorical functions of this system suggest the following parameters:

Aspect. The primary parameter of the system is aspect, intersecting with the parameters of tense and mood. It is defined by the polarity between imperfective and perfective, framing the central perfective-aorist.

The *perfective preterite (aorist)* is the unmarked form and refers to an event simply as completed, and is central to the system. It refers most frequently to past events, and is traditionally equated with a Western preterite, or simple past. However, it may also refer to present or future contexts, and, in addition, may have subjunctive function, notably in subordinate clauses.

The imperfective is undifferentiated and refers to an event simply as not completed. Thus, the indicative present and past may refer to punctual, habitual-iterative or durative-progressive events, in present and past.

The resultative-stative refers to prior action, result, and temporary state in present and past contexts.

Mood. Basically, real action is marked by the indicative, potential action or state by the subjunctive, and unlikely and unreal action by the counterfactual. Both subjunctive and the counterfactual have binary distinctions, imperfective and perfective/resultative-stative. In form, the imperfective counterfactual is identical with the imperfect, while the subjunctive is morphologically aligned with the present. Both occur in independent clauses, as well as in subordinate clauses.

5.5.1 Possession

Possession is expressed by the verb *dār-*, *dāšt-/dor-*, *došt-* ‘have, keep’.

In Tajik, the pronominal clitics in their dative function together with ‘be’ may also express (alienable) possession of a predicate (see Pronominal direct and indirect objects section 5.3.1):

safar pul[-aš] boš-ad agar, me-ti-yam-t[-a]
 S. money-his may be if, IPFV.give.PR-1s-you.s-RA
 ‘if Safar has any money, I’ll give it you.’

5.5.2 Modal constructions, verb + verb

5.5.2.1 Definite future

xāh-am raft/xoh-am raft ‘I shall, will (definitely) go’.

This modal construction with *xāh-* ‘will’ is isolated and expresses definite future, contrasting with the unmarked imperfective present. It is formed from the unmarked present stem of the verb *xāh-/xoh-* ‘will, want, wish’ marked for person plus the short infinitive: *(na-)xāh-am raft/(na-)xoh-am raft* ‘I will (not) go’; compound verbs: *kār (na-)xāh-and kardlkor (na-)xoh-and kard* ‘they will (not) work’.

In Tajik, more so than in Persian, this construction is regarded as “literary” and generally avoided in favor of the simple present-future. See also section 8.3.2 Uncertainty constructions.

5.5.2.2 Volition

Volition is expressed by fully inflected *xāh-*, *xāst-/xoh-*, *xost-* ‘want, wish’ followed by the subjunctive.

<i>bačče če</i>	<i>mi-xāh-adl</i>	<i>mi-xāst-am</i>	<i>xāne be-rav-aml</i>
<i>bačča či</i>	<i>me-xoh-ad</i>	<i>me-xost-am</i>	<i>xona rav-am</i>
child what	IPFV-want.PR-3s	IPFV-want.PT-1s	house SBJ-go.PR-1s
‘what does this/that child want?’		‘I wanted to go home’, lit. ‘that I go’.	

5.5.2.3 Intention and expectation

(1) [*raft-an-ī*] (*hast-)*am ‘I am about to, supposed to’

Verbs of motion, and particularly *raft-an/raft-an* ‘go, leave’, are formulated in a construction formed with the future participle, or gerundive (see section 3.2.2.2e Nominal forms) and forms of ‘be’ to express intention.

In Persian this construction is marginal. Used predicatively, it is found with a semantically highly restricted set of verbs and contexts: intransitive, rarely used in tenses other than the present, and rarely negated: *man raft-an-i hast-am* ‘I have to, am about to go’.

In Tajik there are no such restrictions. It is used with both intransitive and transitive verbs, occurs with *boš-/bud-* ‘be’ and *šav-/šud-* ‘become’ (and colloquially, without a verb), may be negated, allows for any inflected form of the verb; examples of negated, evidential, and progressive forms are:

Intransitive:

man raft-an-ī-yam ~ *hast-am raft-an-ī* *šud-em*
 I go.GERV-COP-1s ~ EX-1s go.GERV become.PT-1p
 'I'm going, I'm off'; 'we were all set to go'.

Transitive:

heč hujum [kard-an-ī] nest-and [*guft-an-ī*] *na-[bud-a]-ast*
 any attack make.GERV not-COP-1p say.GERV not.be.PART.EV-COP.3s
 'they are not going to attack'; 'it appears she's not about to tell';

ma-ro ba dušman [dod-an-ī šud-a] istod-a and
 I-DO to enemy give.GERV become.PART PROG.PART COP-3p
 'they're getting ready to give me to the enemy'.

This productive construction is thus distinct from the use of the predicative use of the gerundive in both Persian and Tajik with forms derived from the transitive verb which may be understood as a passive: Persian *xeyli did-an-i ast* 'it is very much to be seen, worth seeing', Tajik *on jo heč did-an-ī nest* 'that place isn't worth seeing'.

(2) *rav-ona* 'about to, supposed to go'

In Tajik, the most commonly found form in the adjectival derivative is *-ona*:

man estalinobod ravona bud-am ke na-šud
 I St. go.ADJ be.PT-1s that not-become.PT.3s
 'I was about to/supposed to go to Stalinabad, but it didn't work out'.

In Persian, there is *ravāne šod-an; kard-an* 'set out, start; send off', but no corresponding construction.

5.5.2.4. Obligation and presumption

5.2.2.4a Obligation and necessity

Obligation and necessity are expressed by means of the impersonal defective verb *bāyad/boyad*, neg. *na-bāyad/na-boyad* (less frequent variants with milder force are (*mi-*)*bāyest-(i)/me-boist* with *mi-/me-* in its counterfactual function and the frozen archaic counterfactual enclitic *-i*, respectively).

(1) Impersonal generalized statements

bāyad raft/boyad raft 'one (we, etc.) must go'

This construction is formed with the short infinitive (= Stem II).

(2) Personal expressions

bāyad be-rav-ad/boyad rav-ad 'he must go'

This construction is formed with the present subjunctive:

bāyad fardā parvāz be-kon-aml
boyad fardo parvoz kun-am
 must tomorrow flight SBJ.make.PR-1s
 'I have to fly tomorrow';

mi-bāyest in deraxt-rā az bix be-kan-impl
me-boist in daraxt-ro az bex kand-a bar-or-em
 would have to this tree-DO from root pull.PART up-SBJ.bring.PR-1p
 ‘we ought to uproot this tree’.

Frequent other expressions are *lāzem ast*, *budllozim ast*, *bud* ‘it is necessary’ and *majbur hast*, *budlmajbur ast*, *bud* ‘he is obliged, forced to . . .’, etc., similarly followed by the subjunctive.

(3) Counterfactual forms

These are marked by the imperfect in its counterfactual function in the dependent verb, given that *bāyad* does not have a functional past form (similar to English “must”).

(a) imperfective form

bāyad mi-raft/boyad me-raft ‘he had to, should have gone’

bāyad zud-tar mi-raft-il
boyad zud-tar me-raft-ī
 must early-COMP IPFV.go.PT-2s
 must earlier would leave
 ‘you had to leave, should have left earlier’.

This construction also refers to an action that had to be done instead of another:

tāze qabl-aš ham [bāyad mi-raft-am], barādar-am rā madrase be-gozār-am
toza qabl-aš ham [boyad me-raft-am], uka-am-ro maktab mon-am
 ‘before that I had to go drop my brother off at school (and did not come here right way)’.

(b) resultative form (rare)

bāyad raft-e bud ‘he had to, should have gone (earlier, before then)’.

(4) Tajik infinitival construction

In Tajik, obligation in any tense is expressible by a circumlocution such as *darkor* (‘in the act, appropriate’) preceded by a nominalized clause construction:

xobond-an-i vay-ro darkor bud
 laying-EZ he-DO necessary be.PT.3s
 ‘we had to lay him down’.

5.5.2.4b Presumption

bāyad raft-e bāš-ad/boyad raft-a boš-ad

Presumption about a state or the completion of an action is also expressed by *boyad* and is followed by the perfect subjunctive or the present subjunctive of ‘be’. With change of state verbs the completion usually corresponds to a present in English:

bāyad dar xāne bāš-ad/ bāyad tā hālā rasid-e bāš-and/
boyad dar xona boš-ad boyad to holo rasid-a boš-and
 must in house (SBJ) be-3s must till now arrive.PART (SBJ.)be-3p
 ‘she must be at home’; ‘they must have arrived by now’-;

bāyad ānjā istād-e bāš-and/
boyād on jo istod-a boš-and
 must there stand.PART (SBJ.)be-3p
 ‘they must have stood up = be standing there now’.

5.5.2.5 Ability

Ability is expressed by the fully inflected verb *tavān-*, *tavānest-/tavon-*, *tavonist-* ‘be able’. Similar to *bāyad/bojad* there are two constructions:

(1) Generalized impersonal statements

mi-tavān raft/me-tavon guft

This construction uses the bare present tense form plus the short infinitive, highly marked as literary:

mi-tavān goft
me-tavon guft
 IPFV-can say.INF2
 ‘one can (not) say’.

(2) Personal expressions

mi-tavān-ad be-rav-ad/me-tavonist rav-ad ‘he can go’ (present form)

This construction has fully inflected modal and dependent verb forms in the subjunctive with person agreement. The cover term “ability” includes permission as well:

<i>ne-mi-tavānest-am</i>	<i>be xāb</i>	<i>be-rav-am!</i>	<i>mi-tavān-am</i>	<i>be-rav-am?!</i>
<i>na-me-tavonist-am</i>	<i>xob</i>	<i>rav-am</i>	<i>me-tavon-am</i>	<i>rav-am?</i>
not-IPFV-can.PT-1s	to sleep	SBJ.go.PR-1s	IPFV-can.PR-1s	SBJ-go.PR-1s
‘I could not go to sleep’;			‘can I go, leave?’.	

(3) Tajik participial constructions

[raft-a] me-tavon-ad (present form)

Though the Persian subjunctival construction is available in Tajik, the preferred idiom is a participial construction. Specifically, instead of the dependent verb following the modal in the subjunctive, this is a conjunct construction where the past, or perfect participle precedes the modal verb: *[raft-a] me-tavon-am* ‘I can go’ (here with the modal in the evidential mode in the first example, and in the progressive in the second):

vay [tojikī gap zad-a] na-tavonist-a-ast
 he Tajik talk hit.PART not can.PT-EV-COP.3s
 ‘he couldn’t speak Tajik, as far as I know’.

5.5.2.6 Possibility

The fully inflected verb *šav-*, *šod-* ‘become’ may express possibility as well as permission. It also provides the pendant to impersonal (*ne-*)*mi-tavān* ‘one can’. Whereas the latter implies the ability proper to living beings, reflecting the original meaning ‘be strong’ of *tavān-/tavon-*, *šodan/šudan* implies not so much physical as logistical or moral feasibility reflecting its original meaning ‘go, proceed’ (cf. German *es geht* ‘is possible, can be done’, lit. ‘it goes’).

(1) Generalized impersonal statements

mi-šav-ad raft | *me-šav-ad raft* ‘one can go’

In this construction the modal is followed by the short infinitive:

[*ne-mi-šav-ad*] *maze-ye ān-hā-rā* [*farāmuš kard*]
not-IPFV-become. PR-3s taste-EZ those-DO forgotten make.INF2
‘one cannot forget the taste of those’;

čerā na-raft-i?– *na-šod*
why not-go.PT-2s not-become.PT.3s

‘why didn’t you go? – it didn’t work out, was not possible’.

Tajik

man ba maskaw raft-an-ī bud-am, lekin na-šud

I to Moscow go.GERV be.PT-1s but not-become.PT.3s

‘I intended to, was supposed to, was ready to go to Moscow, but it didn’t come off’.

(2) Personal construction

mi-šav-ad be-rav-ad | *me-šav-ad rav-am* (present form)

Persian also has a personal construction, similar to *tavānestan*, where the 3s impersonal modal is followed by the inflected dependent verb in the subjunctive:

ne-mi-šav-ad *be-rav-am*
not IPFV-become.PR-3s SBJ-go.PR-1s
‘it is not possible for me to go ~ can’t I go?’, colloquial *ne-mi-š-e be-r-am?*

5.5.3 Aktionsart constructions, *V + V*

Aktionsart constructions are mostly confined to colloquial speech in Persian (see Windfuhr 1979: 102–105), but are less restricted in Tajik. They are expressed by a small number of verbs with dependent finite verb, infinitive, or participle. Of these, the infinitive construction is inherited from earlier stages of Persian, which is retained in Tajik, but in Persian tends to be assimilated to post-auxiliary position. Typically, most Persian constructions are double finite, while Tajik constructions are participial (see section 5.5.3), as is the case with the progressive construction discussed in section 3.2.6a/b.2:

Persian		Tajik		
(<i>dār-am</i>)	<i>mi-rav-am</i>	<i>rafta</i>	<i>istoda</i>	<i>am</i>
PROG.PR-1s	IPFV-go.PR-1s	go.PART	PROG.PART	COP.1s

both: ‘I am about to leave, leaving’.

5.5.3.1 Inception, non-intentional

In Persian, *raftan* ‘go, leave’ followed by the subjunctive expresses the notion of ‘be going to’ with the connotation of possibility or likelihood, usually in combination of *dāštan*:

mi-rav-am yavās yavāš be in natije be-res-am ke . . .
IPFV-go.PR-1s slow slow at this result SBJ-arrive.PR-1s that
‘I am slowly beginning to understand that . . .’;

mi-rav-ad farāmuš be-šav-ad
 IPFV-go-3s forgotten SBJ-become.PR-3s
 '(this old tradition) is now in the process of being forgotten'.

5.5.3.2 *Inception*

Ali šoru' kard [be neveštan]/
Ali [ba nivištan] šurū' kard
 A. begin made [to writing-INF]
 'Ali began to write'.

5.5.3.3 *Inception, completive*

Persian uses the infinitive in the literary, but TMA marking in the colloquial register; Tajik uses only the former:

Persian

(1) infinitive

asb david-an gereft
 horse run-INF take.PT.3s
 'the horse began to gallop'.

(2a) Past

gereft-and xābid-and
 take.PT-3p sleep.PT-3p

(2b) Imperative

hama-š na-gir be-xāb
 all-its not take.IMP-2s SBJ-sleep.IMP-2s
 'they took to sleeping, fell asleep'; 'don't fall asleep all the time'.

Tajik

korkar-i xud-aton-ro kard-an gir-ed!
 work-EZ own-your-p-DO do-INF take.IMP-2p
 'you get on with/begin your work!'.

5.5.3.4 *Inception, ingressive*

The incipient function of the verb *xāstan* 'want' derives from its modal connotations. This innovative construction is to be distinguished from the long-established, but isolate definite future construction with *xāh-* + past stem: *xāh-am raft/xoh-am raft* 'I will go' (cf. section 8.3.2 Uncertainty constructions). While in colloquial Persian this construction is restricted to ingressive function, in Tajik it is not.

Persian

bičāre mi-xād be-mir-e
 poor IPFV-want.PR-3s SBJ-die.PR-3s
 'the poor fellow is about to die'.

Tajik

Sobir čahār-čub girifta, xost ki bi-rav-ad
 S. frame take.PART want.PT.3s CL SBJ-go.PR-es
 'Sobir picked up the frame and made to go'.

Some of these verbs may be combined in a double Persian finite construction:

māšin *dār-e* *mi-r-e* *be-suz-e*
 car PROG.PR-3s IPFV.go.PR-3s SBJ-burn.PR-3s
 ‘the car is about to begin to burn’.

5.5.4 Tajik conjunct verb constructions

Also known as serial verbs, these are a salient feature of Tajik, especially in the literary language and the Northern dialects, where they seem to have originated as calques on Turkic usage. The category may still be evolving and expanding. In fact, the progressive *kard-a istod-a-*, and the ability construction *kard-a tavon-Itavonist-* discussed above, are grammaticalized instances of the type, and other uses will be noted.

In its most systematic manifestation, the construction joins an inflected form of one of a class of auxiliaries, which provides an Aktionsart or adverbial nuance, to the non-finite “past participle” of the semantically (and grammatically) main verb. There are about 16 such modals, the most commonly used being:

(1) *dih-/dod-* ‘give’ (other-benefactive):

nom-aš-ro *navišt-a* *me-dih-am* *ašula* *xond-a* *dod-and*
 name-his-DO write. PART IPFV-give.PR-1s song sing.PART give.PT-3p
 ‘I’ll write its name down for you’; ‘they performed songs’,
xon-/xond- ‘sing’.

(2) *gir-/girift-* ‘take’ (self-benefactive):

afsona-ro *navišt-a* *girift-am* *mavj-i* *kūtoh-ro* *došt-a* *me-girift-em*
 tale-DO write.PART take.PT-1a wave-EZ short-DO hold.PARTIPFV-take.PT-1p
 ‘I copied down the tale’; ‘we used to get the short wave’, *dor-/došt-* hold’.

(3) *bin-/did-* ‘see’ (tentative):

non-ro *xūrd-a* *did*
 bread-DO eat. PART see.PT.3s
 ‘he tasted the bread’, *xūr-/xūrd-* ‘eat’.

(4) *šav-/šud-* ‘become’ (completion of an action):

kitob-ro *xond-a* *šud-am*
 book-DO read. PART become.PT-1s
 ‘I finished the book’.

(5) *firist-/firistod-* ‘send’ (sudden, violent action):

xand-a *kard-a* *firistod*
 laugh.PART do.PART send.PT-1s
 ‘he burst out laughing’.

(6) *mon-/mond-* ‘stay, remain’ (sudden action followed by continuing state; non-completion or forestalling of an action, depending on the sense of the main verb):

zan-aš kasal šud-a mond man tanho pušt-i on adam-ro did-a mond-am
 'his wife fell ill'; 'I only saw that man from the back'.

(7) *rav-raft-* 'go' (motion away from the speaker; finality; progressivity, depending on the sense of the main verb):

bača david-a raft daftar-am gum šud-a raft
 child run.PART go.PT.3s notebook-my lost become.PART go.PT-1s
 'the boy ran away'; 'my notebook got lost';

ohista-ohista sihat šud-a me-rav-ī
 little little health become.PART IPFV-go.PR-2s
 'little by little you'll get well'.

(8) *o-lomad-* 'come' is similarly used for motion toward the speaker.

5.5.5 Light verb constructions

Light verb constructions (LVC), also referred to by the ill-defined term "compound" verbs, consist of a nominal and a limited set of simple ("light") verbs. They represent a class of complex predicates that is distinct from modal, aktionsart, and conjunct verb constructions, even though lexically their sets of function verbs partially overlap. Most prominent are transitivity-voice pairs of LVCs with 'do' and 'become': *xošk kardan* :: *šodan* 'make :: become dry'. Essentially, then, LVCs are analytical types of de-nominal and de-adjectival verbs as opposed to derivation, such as *fahm-*, *fahm-id-* 'understand' (Ar. *fahm* 'understanding'), *xošk-*, *xošk-id-* 'dry (out), drain' (Pers. *xošk* 'dry'; see 3.2.1.2 above). For many centuries the emergence of LVCs, now estimated at over a thousand, has been the only source of verbal innovation, including basic verbal concepts, e.g. *dust dāštan/dūst doštan* 'love, like,' lit. 'have as friend', *kār kardan/kor kardan* 'work,' lit. 'do work', and in turn has led to the decreasing use of simple verbs, of which only some 150 have retained higher frequency.

These constructions consist of nominals that provide the main lexical meaning and about 18 light (simple) verbs that carry the full inflectional load (including auxiliaries). The nominals include adverbials, nouns, adjectives, or nominalized forms of verbs, such as verbal nouns and active and passive participles—many of which are integrated Arabic loans—both with or without adpositions. The LVCs therefore range from relatively little to considerable morphological, syntactic, and semantic complexity. Thus, Ahadi (2001: 53) found eleven syntactic subclasses of varying complexity, consisting of one- to four-place NPs, determined by the specifics of both the verb and the obligatory complements (even though some of the latter may be contextually omitted under definable conditions). Others distinguish more or fewer subclasses.

Syntactic criteria for testing the syntactic properties and subclassification include prominent place of verbal rection, permissibility of insertions, deletions, and permutations, and semantic shifts (for the insertion of the amplifier *-hā* see Windfuhr 1979: 33). While to a certain extent these hundreds of LVC can be classified, ultimately each has its own syntactic and semantic range. Moreover, they may vary from speaker to speaker. The following example shows the set of the insertion options related to *fekr* 'thought, thinking' plus *kardan* 'do, make'. This LVC has partially replaced the simple literary verb *andišidan* 'think, reflect; plan, devise; fear, be apprehensive' in terms of semantic range (cf. Meyer-Ingwersen 1974):

Nominalization, EZ

fekr kardan-e u

'his thoughts, thinking'.

Insertions permitted

*(yek) fekr-i**mi-kon-am*

'I'll think of s.th.', lit. 'make a thought';

*fekr-hā-(i)**mi-kon-am*same, action amplified by plural *-hā*;*in fekr-rā**mi-kon-am ke*

'I think this that . . .';

*fekr-e xub(-i)**mi-kon-ad*

'he has good ideas', lit. 'makes a good thought';

*fekr-e pul-eš-rā**mi-kon-am*

'I'm thinking of the costs of it',

lit. 'its money-DO';

*be pul-aš fekr**mi-kon-am*

'I think of the costs (that might be involved)',

lit. 'at its money'.

Preverbal

*xub fekr**mi-kon-ad*

'he has very good ideas', lit. 'thinks well';

*hamin-rā fekr**mi-kon-am*

'I am thinking the same-DO';

*barā-ye to kār-i fekr**mi-kon-am*

'I am thinking of some things for you to do

('a job', etc.).

The example of *fekr kardan* contrasts with the LVC *otu zadan* 'iron (clothes)', lit. 'hit with iron', where the insertion of the direct object marker *-rā* dissolves the LVC: *otu-rā zad* 'he hit the iron (in anger)', etc.

There are predictable correlative transitive vs. intransitive pairs using these and other auxiliaries (as had been noted from the earliest Western grammars of Persian onward), the semantically least specified pair being *kardan* vs. *šodan*:

*dorost kařdanl**durust kardan*

'make correct, prepare s.th.';

*dorost šodanl**durust šudan-*

'be corrected, prepared'.

*gerd āvordanl**gird ovurdan*

'amass, convene' ('bring around');

*gerd āmadanl**gird omadan*

'gather, congregate' ('come around').

*rāh andāxtanl**roh andoxtan*

'set in motion, start' ('throw on road')

*rāh oftdanl**roh uftodan*

'set out, start' ('fall on road').

From a semantic perspective, there are two kinds of light verbs:

(1) Those involving metaphorical use of a nominal, adverbial, etc., where the semantic load is spread fairly evenly between both elements and the auxiliary may vary idiomatically:

sar kardanl/sar kardan

'begin', lit. 'make start'

az barl/bar kardan

'learn by heart', lit. 'make by memory';

sar dādanl/sar dodan

'let go', lit. 'give start',

< *sar* 'head, start';*dar bar gereftanl/giriftan*

'embrace', lit. 'take into bosom',

< *bar* 'bosom; memory', not used independently.

(2) Transparent conversions of a nominal using the same set of "dummy" auxiliaries, the semantic load remaining with the nominal:

<i>rāst konl</i>	<i>qofl kon- ~ zan-l</i>	<i>tahiye kon- ~ bin-</i>
<i>rost kun-</i>	<i>qulf kun- ~ zan-</i>	<i>tahiya kun- ~ bin-</i>
‘straighten’,	‘lock, bolt’,	‘prepare’,
lit. ‘make straight’;	lit. ‘make ~ hit lock’;	lit. ‘make ~ see to preparation’.

In literary style, the dummy *kon-*, *kard-lkun-lkard-* may be replaced by *nemā-*, *nemudl namo-*, *namud-* (originally ‘appear; show’), and *šav-lšav-*, *šod-lšud-* by *gard-lgard-*, *gašt-lgašt-* ~ *gardid-lgardid-* (originally ‘turn’):

<i>vazife-yešān-rā ejrā</i>	<i>‘nemud-e andl</i>	<i>vazife-yešān ejrā gašt ~ gardidl</i>
<i>vazifa-ašon-ro ijro</i>	<i>namud-a-and</i>	<i>vazifa-ašon ijro gašt ~ gardid</i>
duty-their-DO	deed do.PART COP-3p	duty-their deed become.PT.3s
‘they have carried out their duty’;		‘their duty was carried out’.

Semantically, LVCs frequently mark general categories such as “duration”. Thus, the two verbs *zadan* ‘hit’ and *kešidan* ‘draw, pull, stretch’ are paired as *dast zadan* vs. *dast kešidan* both ‘touch’. However, *kešidan* connotes continuous (moving) touch, and *zadan* once or quickly repeated action, including making sounds: *vāks zadan* ‘shoeshine’ (*vāks* ‘wax’); *harf zadan* ‘speak’ (*harf* ‘word’); *violon zadan* ‘play violin’; *gul zadan* ‘deceive’ (*gul* ‘deception’).

5.5.5.1 Reflexivity

While the light verbs tend to become semantically bleached, they also tend to develop a more generalized meaning. Thus, *kešidan* belongs to the subset of verbs that are both transitive and reflexive, as in *derāz kešid* ‘(s)he stretched out, lay down’. So does *xordanl xurdan* ‘eat, drink’, which as a light verb connotes “take to self”, and is one of the partners of *zadan*: *gul xordan* ‘be deceived’.

5.5.5.2 Experiencer-subject LVCs

In Persian, a distinct and open subset of LVCs consists of reflexive, or indirect, verb construction where the affected experiencer is marked by a personal enclitic in its dative function: *az ān xoš-aš āmad* ‘he liked it’, lit. ‘from that it became him well’; *qahr-emān šod* ‘we began to quarrel’, lit. ‘to us happened quarrel’; *sard-am ast* ‘I am cold’, lit. ‘to me it is cold’; coll. *če-t-e?* ‘what’s the matter with you?’ (< *če-at ast* ‘what is to you’). Today, experiencer LVCs are particularly frequent in non-formal speech, but they continue an impersonal construction well-documented throughout the diachronic stages of Persian, where it alternated with *-rā* in dative function: *ō-rā xwaš āmad*, lit. ‘it became him well’.

5.5.6 Voice and agency

5.5.6.1 Passive voice

(1) The passive voice is formed for all tenses with the past (perfect) participle of the main verb and the inflected auxiliary *šav-*, *šodlšav-*, *šud-* ‘become, get (intrans.)’:

<i>košt-e mi-šav-adlkušt-a me-šav-ad</i>	<i>košt-e šodlkušt-a šud</i>
‘(s)he is being, will be killed’	‘(s)he was killed’
or ‘is, will be getting killed’;	or ‘got killed’.

This includes the evidential forms and the Tajik conjectural forms, e.g. Tajik *ovard-a me-[-šud-a]g-i-st* ‘it is probably being taken’.

(2) In compound verbs, *šav-*, *šod-[-šav-*, *šud-* replaces the transitive auxiliary *kon-*, *kard-[-kun-*, *kard-* ‘do, make’.

Tajik prefers the “long passive”, i.e. with passivization of *kard-an*: *kard-a šav-[-šud-*, which is at best rare in Persian:

<i>ān</i>	<i>ketāb tarjome</i>		<i>šodl</i>
<i>on</i>	<i>kitob tarjuma</i>	<i>(kard-a)</i>	<i>šod</i>
	that book translation (do.PART) become.PT.3s		
	‘the book was translated’.		

(3) The passive construction has no overt agent, though the agent may be referred to indirectly, where felt necessary, by adverbial EZ phrases such as *az taraf-elaz taraf-i* ‘from the side of’, *be vāsete-yelba vosita-i* ‘by means of’, *az, be dast-e* ‘from, by the hand of’:

<i>in</i>	<i>nāme az taraf-e</i>	<i>edāre</i>	<i>nevešt-e</i>	<i>šodl</i>
<i>in</i>	<i>maktub az taraf-i</i>	<i>rafiq-am</i>	<i>nivišt-a</i>	<i>šud</i>
	this letter	from direction-EZ	office/my friend write.PART	become.PT.3s
	‘this letter was written by the office/by my friend’.			

The passive is comparatively little used in both Persian and Tajik, and is relatively more restricted semantically in Persian than in Tajik, and typically a feature of literary and bureaucratized diction.

The most prominent means of indicating an unspecified agent is the use of third person plural:

<i>ān-rā</i>	<i>tarjome kard-andl</i>	<i>ān</i>	<i>tarjome</i>	<i>šodl</i>
<i>on-ro</i>	<i>tarjuma kard-and</i>	<i>on</i>	<i>tarjuma kard-a</i>	<i>šud</i>
	that-DO translation make.PT-3p		that translation make.PART	become.PT.3s
	‘they translated it’;		‘it was (made) translated’.	

5.5.6.2 Scale of agency and causation

The following illustrates the position of the passive construction on the scale of agency, using past tense forms (Windfuhr 1979: 105–113):

(a) Agent specified, causation:

<i>pesar āb-rā</i>	<i>rixtl</i>	‘the boy poured the water out’.
<i>pisar ob-ro</i>	<i>rext</i>	
boy	water-DO	pour.PT.3s

(b) Agent unspecified, causation:

<i>āb-rā</i>	<i>rixl-andl</i>	‘they, someone poured the water out’
<i>ob-ro</i>	<i>rext-and</i>	~ ‘the water was poured out’.
water-DO	pour.PT-3p	

(c) Agent unspecified, causation:

<i>āb</i>	<i>rixl-e šodl</i>	‘the water was poured out’.
<i>ob</i>	<i>rext-a šud</i>	
water	pour.PART	become.PT.3s

(d) Agent unspecified, no causation:

āb rixtl ‘the water poured out’.
ob rext
water pour.PT.3s

6 SYNTAX AND USE OF FORMS III: COMPLEX SENTENCES

The basic word order in subordinate clauses is essentially the same as in main clauses.

6.1 Clause linking

6.1.1 Overview

There are two distinct types of clause linking: one in which the dependent verb is finite, the other where the dependent verb is a nominal form of the verb. In Tajik, both types exist side by side, due to its partial convergence with Uzbek, whereas in Persian the nominal type is relatively marginal (see section 4.2 *Ezāfe*).

In the finite type, clauses may be paratactic or connected by conjunctions and particles. Devices for clause linking inside the second clause include the use of pronominal anaphora and the verb ending when same subject, as well as gapping. These rules apply irrespective of the type and syntactic status of the clauses involved. Noteworthy for all varieties of Persian is the gapping of the direct object due to the lack of an overt independent non-deictic 3s pronoun (such as English *it*):

did-am, xarid-aml
did-am, xarid-am
I saw, I bought
‘I saw it, and bought it’.

A further linking feature is the strict distinction between indicative vs. non-indicative mood in the subordinate clause, depending on the semantics of the verb phrase in the matrix clause (a distinction referred to as “realis” vs. “irrealis” in current typological discussions).

For the use of mood, the logic of the three main modal degrees of reality in the verb system applies: (1) events or states that are expressed as factual use the indicative forms; (2) those that are expressed as possible or contingent, or describe a general characteristic, use the two subjunctive forms (imperfective and perfective, e.g. *be-rav-am/rav-am* and *rafte bāš-amlrafta boš-am*, and more complex forms in Tajik); (3) those that are expressed as counterfactual, ranging from unlikely to impossible, use the two irreal forms (imperfective and perfective, *mi-raft-aml/me-raft-am* and *rafte bud-amlrafta bud-am* in their counterfactual function, and more complex forms in Tajik).

The use and various combinations of these three devices, morphological connectives, anaphora, and mood, by the speaker results in a continuum from more tightly to less tightly integrated sentences, partially conditioned by the semantic-pragmatic context.

6.1.2 Coordinating conjunctions and clitics

The coordinating single and paired conjunctions are the same as those used to join NPs, but may have additional functions.

(1) Independent *valva* ~ enclitic *-ol-u*, etc. These denote a second event contemporaneous with or closely following the first, or a consequence of it, the latter emphasized by *valva*:

Same subject:

<i>in soxan-e talx-aš-rā goft</i>	<i>va az otāq birun</i>	<i>raftl</i>
<i>in suxan-i talx-aš-ro goft</i>	<i>va az xona bar-omad-a</i>	<i>raft</i>

this speech-EZ bitter-his-DO say.PT.3s **and** from room out come.PART go.PT.3s
‘he said these bitter words and left the room’.

Different subject:

<i>“keš” goft-am-o</i>	<i>gonješk be havā parvāz kard</i>	<i>raftl</i>
<i>“kiš” guft-am-u</i>	<i>čumčuk ba havo parvoz karda</i>	<i>raft</i>

“pshsh” say.PT-1s **and** sparrow to air flight make.PT.3s/ go.PT.3s
PART
‘I shouted “kish”, **and** the sparrow flew up into the air’, *karda raft*, cf. conjunct verbs.

(2) The conjunctions *ammālammo*, *valilvāle*, *lākenllekin* ‘but, however’ (all borrowed from Arabic) introduce the second of two contrasted clauses:

<i>havā abri šod,</i>	<i>vali bārān na-bāridl</i>
<i>havo abr šud,</i>	<i>lekin boron na-borid</i>

weather cloudy/cloud become.PT.3s but rain not rain.PT.3s
‘it became cloudy, but it didn’t rain’.

(3) *hamlham* and *kelki*

Ham. The clitic *ham* ‘also, too’ may have focusing-contrastive function, as does the generalized complementizer *kelki*. Note that the use of the latter is typically colloquial (see section 6.4.5 Idiomatic *ke*):

<i>to ān kār-rā</i>	<i>bo-kon,</i>	<i>man ham ~ ke</i>	<i>mi-rav-aml</i>
<i>to on kor-ro</i>	<i>kun,</i>	<i>man ham ~ ki</i>	<i>me-rav-am</i>

you.s that work-DO SBJ.do.IMP.2s I also though IPFV.go.PR-1s
‘you do that work, I though ~ but I am leaving’.

In colloquial speech, the members of the preceding sets of linkers may occasionally be used in sequence, which can be rendered only approximately in English. Therefore these linkers mark three successive syntactic levels.

The *kelki*-clause here can be attached after the clitic *hamlham* (but not the coordinating conjunctions), which suggests that syntactically it is a subordinate clause.

(4) Anaphoric clitic *-aš*

The 3s clitic *-aš*, or the independent pronoun *ān/on*, refers to a preceding situation, mainly in colloquial speech:

<i>ba'd-eš ~ ba'd az un mā raft-iml</i>	<i>āxer-eš raftl</i>	<i>ku-š-eš? (Persian)</i>
<i>ba'd-aš ~ ba'd az on mo raft-em</i>	<i>oxar-aš raft</i>	
after-it after from that we go.PT-1p	final-it go.PT.3s	where-there-his?
'there(-aš)-after, then we left';	'in the end he left';	'where is he?'

(5) Anaphoric clitic *-rā*

The phrase clitic *-rā* in its general function of topicalizing reference (see section 3.1.4.2a) introduces clipped clauses:

Hasan-rā – ne-mi-dān-am če šode ast
 'as to Hasan, I don't know what happened'.

6.2 Subordinate clauses and linking

Subordinate clauses fall into three major types: (1) relative (property) clauses; (2) complement (propositional) clauses; and (3) adverbial clauses. Unmarked relative clauses are embedded in the matrix clause. Unmarked complement clauses, both subject and object clause as well as purpose clauses, are post-verbal. Unmarked adverbial clauses fall into two sets: those that precede the main clause, and those that follow it. This unmarked distinction follows the logical or temporal sequence of the events or situations in the respective main and subordinate clause.

Typologically significant is the existence of a generalized complementizer, *kelki* (partially comparable to English *that*).

This complementizer is an enclitic, and may introduce most types of subordinate clauses. As such, it precedes them when they follow the matrix clause, but is inserted after their first constituent component when they precede the matrix clause. Moreover, it combines with demonstrative pronouns and/or nouns in conjunctive constructions. Significant for the function of this complementizer are two facts: it may also attach to most semantically specified conjunctions (e.g. *čun kelčun (ki)* 'because'), and second there are pragmatic-semantic conditions where it is said to be optional. This suggests that *kelki* is not semantically empty.

Subordinate clauses may be linked as follows: (1) without morphological linker (asyndetic), (2) with the general complementizer *kelki*, and (3) with complex conjunctive constructs consisting of a nominal head + *kelki*. The head may be:

- (a) minimally a demonstrative prop, *inlin* or *ānlon*: *in ke lin ki*;
- (b) an EZ construction with the pronominal prop: *N-e in ke lN-i in ki*;
- (c) a noun + linker *-i*, with optional prop: *(ān) N-i kel(on) N-e ki*.

Depending on the role of the head in the matrix clause, these phrases may be modified by primary and secondary adpositions (see section 3.1.4.1 Prepositions), including the direct object marking postposition *-rāl-ro*: *in-rā kelin-ro ki*, *N-e in-rā kel N-i in-ro ki*, *N-i-rā kelN-e-ro ki*.

In 'this' is the default demonstrative prop in most contexts, whereas *ān* 'that' tends to be preferred in the head of relative clauses. The choice may be semantically significant: *in* pointing to the present or immediate situation, or anticipated a future event, and *ān* referring to a more remote situation or past state. – Both demonstratives may have focused forms marked by stressed *ham-* 'same', *haminlhamin*, *hamānlhamon* 'this, that very'.

Focusing, topical reordering and extraposition obligatorily requires the use of such a prop, minimally *kelki* (cf. also Aghaei 2006).

Note: In the following the interlinear annotation will be simplified.

6.3 Relative clauses

6.3.1 Restrictivity and specificity

Relative clauses (RCs) distinguish between restrictive and non-restrictive types. The former requires the linker *-il-e* before the generalized complementizer *kelki*, the latter does not. This distinction corresponds to the EZ construction marked by the linker *-el-i* vs. zero linkage in appositional NPs (Windfuhr 1979: 65):

Relative linkage	EZ linkage	
<i>N-i kelN-e ki</i>	<i>N-el-i</i>	complement
<i>N-keN-ki</i>	<i>N-</i>	complement

Like other nouns, the head noun of relative clauses may be marked by the indefinite *-il-e*, which is homophonous with the relative linker. When combined, these two markers coalesce into a single *-il-e*.

In addition, heads may be further marked by a demonstrative pronoun (*ān* being the default choice), or by indefinite *yek/yak* ‘one, a’. The following shows the basic patterns, here using Persian only (DEM = demonstrative; I = indefinite marker; L = linker).

TABLE 8.14: ORDER OF RELATIVE CLAUSES

	DEM	HEAD I	L	PROPERTY	PREDICATE
<i>Restrictive</i>					
(1a)	<i>ān</i>	<i>doxtar</i>	<i>-i</i>	<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>
(1b)	<i>yek</i>	<i>doxtar -i</i>	<i>-i > -i</i>	<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>
<i>Non-restrictive</i>					
(2a)	<i>ān</i>	<i>doxtar</i>		<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>
(2b)	<i>yek</i>	<i>doxtar-i</i>		<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>

Restrictive:

(1a) ‘(that) the girl who loves Ali is here’ (not the other one).

(1b) ‘a (certain) girl who loves Ali is here’ (not another one).

Non-Restrictive:

(2a) ‘(that) the girl, who I think, etc. loves Ali, is here’.

(2b) ‘a girl, who I think, etc. loves Ali, is here’.

6.3.2 Gapping and anaphora

There are no dedicated relative pronouns. Agreement and case roles of the head inside the relative clause are indicated by the combination of the general complementizer *kelki* and anaphoric reference. The latter is marked in the verbal ending if same subject, and by an anaphoric pronoun otherwise; non-anaphoric repetition is an emphatic option. The occurrence of the anaphoric phrases ranges from obligatory to optional. When

optional, occurrence marks emphasis (examples here from Persian only; translation is literal; parentheses = optional; double parentheses = strongly emphatic):

Basic cases:

Subj., <i>N who</i> , opt.:	<i>doxtar-i ke ((u)) Ali-rā dust dār-ad</i> 'the girl that she Ali loves';
Dir. obj., <i>N whom</i> , opt.:	<i>doxtar-i ke Ali ((u-rā)) dust dār-ad</i> 'the girl that Ali her loves';
Ind. obj., <i>N to whom</i> , obl.:	<i>doxtar-i ke Ali be u gol dād</i> 'the girl that Ali to her flowers gave';
Abl., Comp., <i>N from/than whom</i> , obl.:	<i>doxtar-i ke Ali az u kuček-tar ast</i> 'the girl that Ali from her younger is';
Poss., <i>N whose</i> , obl. (mostly clitic):	<i>doxtar-i ke Ali dust-aš ast</i> 'the girl that Ali friend-her is'.

Adverbial:

Time, <i>N when</i> , opt.:	<i>ruz-i ke ((dar ān)) in ettefāq oftād</i> 'the day that (on) this happened';
Place, <i>N where</i> , opt.:	<i>deh-i ke (dar ān) zendegi mi-kard</i> 'the village that in it she lived';
Reason, <i>N why</i> , none:	<i>ellat-i ke harf na-zad</i> 'the reason that she didn't talk';
Manner, <i>N how</i> , none:	<i>towr-i ke ((be ān)) in kār-rā anjām dād</i> 'the manner that in it this work she accomplished'.

Clausal:

CL, <i>which</i> , none:	<i>aslan harf na-zad, ke xeyli ajib ast</i> 'at all she didn't talk, which is very strange'.
--------------------------	---

6.3.3 Ordering of restrictive RCs

Restrictive relative clauses distinguish an unmarked and a marked order: (1) When unmarked, they immediately follow the head phrase, and are thus embedded within the matrix clause as in the example above: (2) When marked, they are post-verbal, and have focusing-contrastive function. The post-verbal position is also typical for relative clauses that identify a characteristic of the head, irrespective of whether the head is indefinite or definite: (*ān*) N(-*il-e*):

(1) Unmarked:

<i>ān doxtar-i [ke Ali-rā dust dār-ad] raft</i> that girl-L that A.-DO friend have.PR-3s go.PT.3s '(that) the girl whom Ali loves left'.
--

(2) Contrastive:

<i>ān doxtar-i raft [ke Ali-rā dust dār-ad]</i> that girl-L go.PT.3s that A.-DO friend have.PR-3s go.PT.3s '(that) the girl [whom Ali loves] left', not the other one.
--

(3) Topical:

yek ketāb-i-rā *towsiye* *me-kon-am* [*ke xod-am xānd-am*]
yak kitob-e-ro *tawsiya* *me-kun-am* [*ki xud-am xond-am*]
 one book-INDEF-DO commendation IPFV.make.PR-1s that self-my read.PT-1s
 'I'll recommend a (certain) book that I have read myself'.

Copular relative clauses are necessarily post-verbal because the copula is enclitic. However other forms of 'be', such as subj. *bāš-/boš-*, past *bud-/bud-* and the existential verb *hast-/hast-* 'be there' do allow for positional options:

ān doxtar-i-st [*ke Ali-rā dust dār-ad*]
 that girl-L COP.3s who Ali-DO friend has
 'she is the girl who loves Ali'.

6.3.4 Use of moods

Inside relative clauses, the moods express the three degrees of reality in the verb system, from factual to potentially factual to counterfactual, all necessarily pragmatic responses to situations, whether with explicit reference or as general statements (then often with present subjunctive). Typically, relative clauses may be relativized adverbial clauses (see section 6.5 Adverbial clauses).

(1) Indicative

(a) Facts:

duxtarča [*ki dar gird u peš-i xud modar-i xud-ro na-yoft*]
 little girl [that in around and front-EZ self mother-EZ self-DO not.find.PT.3s
ba girya daromad
 to cry.INF in come.PT.3s
 'the little girl, who did not find her mother anywhere near, began to cry', implicitly also 'when' and 'because she did not find'.

(b) Referential facts:

in-jā kas-i-st *ke kār mi-kon-adl*
injo kas-e-st *ki kor me-kunad*
 here person-L COP.3s that work IPFV.do.PR-3s
 'here is someone who works, is working'.

(c) General statements:

Indicative

kas-i *ke kār mi-kon-ad* *xošhāl astl*
kas-e *ki kor me-kun-ad* *xursand ast*
 person-L that work IPFV.do.PR-3s happy COP.3s
 'someone who works is happy'.

(2) Subjunctive

(a) General characteristic:

in-jā kas-i nist ke kār be-kon-ad
injo kas-e nest ki kor bi-kun-ad
 here person not.COP.3s that work SBJ.do.PR-3s
 'there is no one here to work';

(b) General comparison:

mesl-e mard-i ke aslan kār-i na-karde bāš-adl
misl-i mard-e ki aslan kor-e-ro na-karda boš-ad
 like-EZ man-L that at all work-INDEF-DO not.do.PART (SBJ)be.PR-3s
 'he looks like a man who has done absolutely nothing';

(c) Purpose:

be zabān-i ke mā ham be-fahm-im harf be-zan-id
ba zabon-e ki mo ham bi-fahm-em gap zan-ed
 in language-L that we too SBJ.understand.PR-1p word speak.IMP-2p
 'speak the/a language that we can understand';

(d) Wish:

ketāb-i mi-xāh-am ke gerān na-bāš-adl
kitob-e me-xoh-am ki qimat na-boš-ad
 book-INDEF:L IPFV.want.PR-1s that expensive not (SBJ.)-be.PR-3s
 'I want a book that hopefully is not expensive';

(e) Condition and consequence:

on ki ba mo dastdarozi kun-ad, albata, bozi kun-ad
 that that to we handraising SBJ.make.PR-3s surely, play SBJ.make.PR-3s
 'he who raises his hand against us surely risks his head', Tajik.

(4) counterfactual condition

čiz-i be man foruxt ke xod-aš ne-mi-xaridl
čiz-e ba mo furūxt ki xud-aš na-me-xarid
 thing-INDEF-L to I sell.PT.3s that self-his not.IPFV.buy.PT.3s
 'he sold me something that he himself wouldn't buy'.

6.3.5 General-indefinite RCs

(1) Indefinite relatives '-ever'

Clauses introduced with the distributive pronouns *har* 'each, every' designate a restricted class of entity, defined only by the clause itself. The heads are relative constructs with indefinite pronouns and adverbial nouns, and have the structure:

(a) *har kelharki* 'whoever' vs. *har-čelhar-čī* 'whatever', *har kas kelhar kas ki* 'whoever (person)' vs. *har ān čelhar on čī* 'whatever'; *har čandlhar čand* 'however much, although, notwithstanding'; (b) *har jā kelhar jo ki* 'wherever (place)'; *har kojā kelkujo ki* 'wherever'; *har vaqt kelhar vaqt ki, vaqt-i kelvaqt-e ki* 'whenever'; *har tovr-i kelhar tavr-e ki* 'however (way)'. Note that these are the only type of relative heads that distinguish

human and non-human gender (see section 3.1.3.7 Indefinite pronouns). The use of tense and mood is the same as in other relative clauses.

(2) Specific indefinite heads (optional demonstrative requiring *kelki*):

har ān ke dānā bāš-ad tavānā bāš-ad
har on ki dono boš-ad, tavono boš-ad
 each that that knowledgeable (SBJ.)be.PR-3s powerful (SBJ.)be.-3s
 ‘whoever is knowledgeable is wise; free variation of saying:
har ke dānā bov-ad tavānā bov-ad, which retains the archaic subjunctive stem *bov-*for
bāš-.

(3) Non-specific pronominal heads

Non-specific heads do not require *kelki*:

har kojā ke be-rav-i, be-gu ke mā ki hast-im
har kujo ki rav-i, falimon ki mo ki-st-em
 wherever that SBJ.go.PR-2s IMP.say/understand that we who EX.1s
 ‘wherever you go, tell people who we are’;

har kas-i-rā mi-did, az u pul qarz mi-kardl
har kes-e-ro me-did, az ū pul qarz me-kard
 each person-L-DO IPFV.see.PT.3s from he money loan IPFV.make.PT.3s
 ‘he borrowed money from whomever he met’ (Lazard 1992: 235).

(4) Non-pronominal indefinite heads:

Similar to other post-verbal object clauses, which do not require *kelki*, lexical heads that introduce object clauses do not require the complementizer. (In such cases, there tends to occur a pause when uttered):

mi-xāst-am jā-i bud-am- na ādam-i bāš-ad, na dard-esar
me-xost-am jo-e bud-am- na odam-e boš-ad, na dard-i sar
 IIPFV.want.PT-1s place-L be.PT.1s -not person-INDEF (SBJ.)be.-3s not headache
 ‘I wished I were somewhere where there were no people and no worry’;

barā-ye ān kār šab-i mi-xāh-ad
baro-i on kor šab-e darkor
 for-EZ that work night-L IPFV.want.PR-3s/needed

– *sobh na-dāšte bāš-ad*
 – *subh na-došta boš-ad*
 – morning not.have.PART (SBJ.)be.3s
yā ruz-i mi-xāh-ad – *āftāb-āš*
yo ruz-e darkor ast oftob-aš
 or a day IPFV. want.PR-3s/needed COP.3s – sunshine-its

foru na-rav-ad

furū na-rav-ad

does not(SBJ.)-go.PR-3s

‘in order to do that, one needs a night that has no morning, or a day when the sun never sets’.

6.3.6 *Head incorporation*

Head incorporation is an alternative to the anaphoric strategy: the case rection of the correlating NP in the subordinate clause is moved to the head noun phrase; this process tends to trigger focused lexical repetition in the main clause:

- (1) No change of sequence
 (a) Direct object *-rāl-ro*:

- (1) indefinite-specific object:

<i>magar [zan-i</i>	<i>ke]</i>	<i>ensān u-rā dust mi-dār-ad]</i>	<i>almās-o javāher-ast?</i>
<i>magar [zan-i-rā</i>	<i>ke]</i>	<i>ensān dust mi-dārad]</i>	<i>almās-o javāher-ast)?</i>
but a woman-L DO	that	man her loves	diamond and jewels is

‘but is the woman that a person loves diamonds and jewels?’

(Ali Dashti, cited in Lazard 1992: 232)

- (2a) Prepositional phrase, Persian

<i>bāyad dānest [zabān-i</i>	<i>ke]</i>	<i>mā [bā ān]</i>	<i>emruz soxan mi-gu-yim . . .</i>
<i>bāyad dānest [bā zabān-i</i>	<i>ke]</i>	<i>mā</i>	<i>emruz soxan mi-gu-yim] . . .</i>
must know with language-L	that we	with it	today speech say

hamān zabān ast ke . . .

same language is that

‘one must realize that the language with which we communicate today . . . is the same language which . . . (Kasravi Tabrizi, cited in Lazard 1992: 232)

- (2b) Tajik

<i>[</i>	<i>ba hamon jo-e ki az ān]</i>	<i>omad-i</i>	<i>bar gard!</i>
<i>[az jo-e</i>	<i>ki]</i>	<i>omad-i], [ba hamon jo]</i>	<i>bar gard</i>
from the place	that	you came	to same place return

‘go back to where you came from!’.

6.3.7 *Tajik nominalized relative clauses*

Two types of nominalized clauses were discussed in sections 4.2.2.5d, e. Both are EZ constructions where the dependent verb is nominalized as a perfect participle. The latter is the perfect participle in *PT-e/PT-a* (*raftelrafta*) which alternates with the participle in *PT-agi* (*raftagi*) in Tajik. While in Persian the participial EZ construction is of limited length, Tajik has developed an elaborate system that goes far beyond the inherited pattern and allows for full-length participial clauses.

Typologically significant for this pattern is the fact that participialization does not involve the complete loss of tense, mood, and aspect: while tense distinction is lost (as in a number of finite forms, see Table 8.8), modal and aspect marking are retained. That is, the participial forms may be marked by the prefix *me-* for imperfectivity and non-indicative mood as well as by the participle *istoda* for progressivity. Given the two Tajik participles, the system of the participial forms are shown in Table 8.15 (*rav-/raft-* ‘go, leave’).

TABLE 8.15: TAJIK FINITE AND PARTICIPIAL FORMS

Finite forms			
<i>rafta istoda ast</i>	<i>me-rav-ad</i>	<i>rafta ast</i>	PRS
<i>rafta istoda bud</i>	<i>me-kard</i> , and <i>(bi-)kun-ad</i>	<i>rafta bud</i>	PST SBJ
Participial forms in <i>-a</i>			
<i>rafta istod-a</i>	<i>me-rafta</i>	<i>rafta</i>	
begin of progress or in progress	general process or continuing state, and modality	process concluded or temporary state	
Participial forms in <i>-agī</i>			
<i>rafta istod-agī</i>	<i>me-raft-agī</i>	<i>raft-agī</i>	
begin of progress or in progress	general process or continuing state, and modality	process concluded or temporary state	

The distinction between these two sets is partially a matter of register, the *-agī* forms being notably more frequent in colloquial speech.

As is the case with the reduced participial construction shared by both Persian and Tajik, the passive interpretation of these constructions with transitive verbs is partially suspended, and voice interpretation depends on the context, including the animacy of the subjects. Most distinct from Persian is the fact that in these constructions the agent-subject may be marked by the possessive clitics.

6.3.7.1 Level of nominalizations

Compared to finite subordination, the process from finite clause to full nominalization and Turkification involves not only the deletion of conjunction and nominalization, but, as mentioned, the replacement of the finite subject verb ending (and pronominal subject) over several stages with a possessive pronoun, and, in a final step, the inversion of head and clause. Schematically, the sequence of these stages may be shown using *-agī* forms (SUBJ = subject; VB = verb form; END = finite ending; POSS = possessive clitic):

Unmarked: *raft-agī*

(0) *ob-e [ki mo az hawz ovard-em]*
water-L that we from pool bring.PT-1p
'the water [that we brought from the pool]'

>>

(1) *ob-i [mo az hawz ovard-agī]*
(2) *ob-i [az hawz ovard-agi-amon]*
(3) *[az hawz ovard-agi-amon] ob*

Indicative or subjunctive: *me-raft-agī*,

corresponding to present or past according to context

(0) *ob-e [ki mo az hawz me-ovard-em]*
'the water that we from pool IPFV.bring.PT-1s
'the water that we are ~ were to bring from the pool'

>>

(1) *ob-i [mo az hawz me-ovard-agī] EZ*
(2) *ob-i [az hawz me-ovard-agi-amon]*
(3) *[az hawz me-ovard-agi-amon] ob*

Progressive: *rafta istod-agī*

corresponding to present or past according to context

(0) *ob-e [ki mo az hawz ovarda istoda em]*
 water-L that we from pool bring. PART PROG.PART COP-1p
 >>

(1) *ob-i [mo az hawz ovarda istod-agī]*

(2) *ob-i [az hawz ovarda istod-agi-amon]*

(3) [az hawz ovarda istod-agi-amon] **ob**

As is true for any other relative clause, the case rection of a nominalized relative clause in the matrix clause is irrelevant for its internal structure, and is indicated by a preposition preceding its head, and by *-rā* following the entire clause.

Further examples

(1) *raft-agī*

Finite *ke navisonid-id > navisonid-agi-aton* ‘which you dictated’, lit. ‘caused to write’:

mazmun-i maktub-ho-i [navisonid-agi-aton]-ro ba ū naql kard-em
 content-EZ letters-EZ [having dictated-your.p]-DO to he repeat make.PT-1p
 ‘we repeated to him the content of the letters [you had dictated]’.

(2a) *me-raft-agī*, indicative

mon ki on urf-u odat-ho-i
 allow that that traditions and custom-p-EZ
[zan-ro kaniz, mard-ro yulom me-kard-agī]
 [woman-DO slavegirl, man-DO slave IPFV.making.]
az miyon bar-došta šav-ad
 from among abolish.PART (SBJ)-become-2s
 ‘let those traditions that make women into servant girls and men into slaves be abolished’; gerundial ‘the traditions of making . . .’

(2b) *me-raft-agī*, modal function

man zan-i [sar me-dod-agī] na-dor-am
 I wife-EZ [release IPFV-giving] not.have.PR-1s
 ‘I don’t have a wife to divorce’, i.e. ‘I’m not about to divorce my wife’.

(3a) *rafta istod-agī*, active interpretation

(indicated by presence of personal clitic)

qozi-ho poy-ho-i [ba taht-i šikam kašida-istod-agi-ašon]-ro ham
 the qadis foot-p-EZ [to under-EZ belly being pulling-their]-DO even
ba kor dar-ovarda, du-poya gurext-and
 to work bring.PART] two-footed flee.PT-3p
 ‘even the qadis brought their legs [that they had been pulling under their bellies] into action, and fled hotfooting’.

(3b) *karda istod-agī*, passive interpretation

(indicated by lack of possessive clitic)

agar etz-i [vay yod doda istod-agi]-ro
 if thing-EZ [he memory give.PART PROG]-DO
durust gufta na-tavon-ad
 correct say.PART not(SBJ.)-can-3s
 'if he cannot correctly repeat something that he is being taught . . .'

In the ultimate reduction, these participializations are compacted to nouns:

[ba qur'on amal namud-agī] kas [dar xona bud-agi]-ho
 [to Koran act doing] person [in room being/-plural
 'a person acting acc. to the Koran'; 'those who are/were in the room';

[qalam be javob na-me-girift-agī]
 [pen without permission not-IPFV.take.ADJ]
 'who is never taking the pen without permission'.

6.4 Complement clauses

Complement clauses comprise subject and object clauses as well as purpose (or final) clauses. The former two are typically in the indicative, and will be discussed in the following, while purpose clauses are typically in the subjunctive and will be discussed under Adverbial clauses (see section 6.5).

6.4.1 Subject and object clauses

Both subject and object clauses have the same basic syntactic structures and follow the same rules of permutations (see also Aghaei 2006). The order of matrix clause and complement clause defines three levels of functional markedness:

- (1) In the unmarked order, the complement clause, with or without the complementizer *kelki*, follows the matrix clause.
- (2) In the focused order, a cataphoric demonstrative, mostly *in*, with or without a lexical head, is inserted in the matrix clause, and the complement clause is introduced by the complementizer.
- (3) In the topical order, there is inversion: the complement clause precedes the matrix clause, and is introduced by the cataphoric demonstrative plus complementizer, *in kelin ki* (identical to one of the heads introducing relative clauses), with object marker *-rāl-ro* inserted if object clause.

Example patterns

((1) = object clause; (2) = subject clause; (3) = relative clause; AG = agent; parentheses: optional). Note that agents are optional since subjects are obligatorily marked in the verb ending.

TABLE 8.16: FOCUSED AND TOPICAL ORDER

(1) Basic unmarked pattern, post-verbal:

(AGENT)	PREDICATE	+	COMPLEMENT
(a) (<i>Hasan</i>) Hasan	<i>mi-dān-ad</i> knows	+	[<i>ke Ali in xāne-rā sāxt</i>] that Ali this house-DO built
(b) (empty)	<i>doruq-ast</i> lie is	+	[<i>ke Ali in xāne-rā sāxt</i>] that Ali this house-DO built

(a) '(Hasan) knows that Ali built this house';
(b) 'it is a lie that Ali built this house'.

(2) Focused pattern, post-verbal:

(AG) IDEM (HEAD)	PREDICATE	+	COMPLEMENT
(a) (<i>Hasan</i>) <i>jin-rā</i> Hasan this-DO	<i>mi-dān-ad</i> knows	+	[<i>ke Ali in xāne rā sāxt</i>]
(b) <i>jin</i> (<i>xabar</i>) this information	<i>doruq-ast</i> is lie	+	[<i>ke Ali xāne-rā sāxt</i>]

(a) 'Hasan does know (this), that Ali built this house';
(b) 'this (matter) is a lie, that Ali built this house'.

(3) Topical pattern, preposed:

PROP IDEM (HEAD)	Topical inversion COMPLEMENT (AG)	PREDICATE
(a) <i>jin</i> (<i>xabar</i>)(- <i>rā</i>)	[<i>ke Ali in xāne-rā sāxt</i>] >	(<i>Hasan</i>) <i>mi-dān-ad</i>
(b) <i>jin</i> (<i>xabar</i>)	[<i>ke Ali xāne-rā sāxt</i>] >	<i>doruq-ast</i>

(a) '(this matter) that Ali built this house, Hasan knows (about) it';
(b) 'this matter) that Ali built this house, that's a lie'.

6.4.1.1 Use of TMA

Epistemic matrix verbs of observation (cognition) of facts (see, know, understand) as well as verbs and expressions governing potential actions or states (modal verbs; want, must, can; decide, try, etc.) do not map their tense onto the dependent verb. Rather the former require the indicative (present or present perfect), and the latter the subjunctive (present subjunctive and perfect/past subjunctive):

fahmid-am ke ādam-e xub-i ast
fahmid-am ki odam-i xub-e ast
understand.PT-1s that person-EZ good-INDEF COP.3s
'I realized that he was a good person';

did ke hame rafte and
did ki hama rafta and
see.PT.3s that all go.PART COP.3p
'he saw that all had left';

tasmim geraft-am be ān-jā be-rav-am
qaror dod-am ba on-jo rav-am
decision take/give.PT-1s to there SBJ.go-PR-1s
'I decided to go there'.

6.4.2 Embedded speech

6.4.2.1 Levels of integration

Embedded speech may be direct, indirect, or integrated indirect.

(1) Direct speech

Persian and Tajik partially differ in the representation of direct speech. In Persian, the speech string follows the verb of saying, which may be introduced by the general complementizer *kelki*:

mo' allem goft ke al'ān daftar-hā-yetān-rā be-gir-id
 teacher say.PT.3s that now notebook-p-your.p-DO SBJ.take.PR-2p
 'the teacher said, now take your notebooks'.

(2) Indirect speech and questions

mo' allem goft ke hālā daftar-hā-yemān-rā be-gir-im
muallim guft ki aknun daftar-ho-yamon-ro gir-em
 teacher say.PT.3s that now notebook-p-your.p-DO SBJ.take.PR-1p
 'the teacher told us to take our notebooks now'.

The interrogative marker *āyālo* 'whether' is used in with indirect questions, with or without the complementizer *kelki*, paired *ke āyā . . . yā . . . , ki . . . yā . . .*:

ma'lum nist ke āyā u hamrāh mi-ā-yad, yā nal
ma'lum nist ki oyo vay hamroh me-o-yad, yo na
 clear not.COP.3s that whether he along IPFV.come.PR-3s or not
 'it isn't clear whether he is coming along or not'.

(3) Implicit indirect speech

Direct speech may be further integrated with verbs of cognition:

šomā xod-etān mi-dān-id ke u kojā raft
šumo xud-aton me-don-ed ki ū kujō raft
 you.p self-your.p IPFV.know.PR-2p that he where go.PT.3s
 'you know where he went';

mi-dānest-am ke mo' allem-emān dir bi-ā-yad
me-donist-am ki muallim-amon der bi-o-yad
 IPFV.know.PT-1s that teacher-our late SBJ.come.PR-3s
 'I knew that our teacher would be late'; here the subjunctive implies a response to the question about the teacher's absence.

6.4.3 Adpositional and EZ complement clauses

Distinct subsets of complement clauses are those that depend on complex verbs consisting of adpositional noun phrases or EZ constructions based on adjectives, participles, and nouns, many of which are Arabic loans:

(1a) preposition (a) *azlāz* 'from' (source, cause), *tarsidan azlāz* 'be afraid of', *xošhāl budan azlāz* 'be happy about'; (1b) preposition *bālbo* 'with', *movāfeq*

budan bālmuvofiḡ budan bo ‘agree with’ (Arabic participle); (2) EZ construction, *montazer-e . . . budan ~ entezār-e . . . dāštan/muntazir-i . . . budan ~ intizor-I . . . doštan* ‘be expecting, waiting for’ ~ ‘have the expectation of’ (Arabic participle and noun).

In these subsets, the retention of adpositions and EZ construction marks topicalization.

(a) source *az* ‘from’:

xošhāl-am ke āmad >>

az ān *xošhāl-am ke āmad*

‘I am happy that he came, is here now’ >>

‘what I am happy about is that he came’, lit. ‘from that I am happy that . . .’;

(b) objective EZ:

montazer bud-am ke bi-ā-yad >>

muntazer-e ān bud-am ke bi-ā-yad

‘I was waiting for him to come’ >>

‘what I was waiting for was for him to come’, lit. ‘I was waiting that, that he come.’

6.4.4 Tajik Turkic-type nominal clauses

In Tajik there are not only the Persian type, as above, but also at least three variants of a Turkic conjunct-type of construction, which are more frequent in literary Tajik. In addition there are finite two-clause and nominalized one-clause constructions.

(1) Two-clause constructions

(a) The matrix clause follows the speech string, and the verb of saying is incorporated parenthetically within the speech string, or more typically is enclitic to it (rather than clause-final):

[*man*] *guft* [*ba šahr na-me-rav-am*]
I say.PT.3s to town not.IPFV.go.PR-1s
‘I, she said, am not going to town’;

[*uka, bisyor kalon-girā na-kun-ed!*] *guft* *Šokir ba Ahmad*
little brother, very high-mighty not.IMP.do-2p say.PT.3s Shokir to Ahmad
‘little brother, don’t be so high-and-mighty, said Shokir to Ahmad’.

(b) with nominalized forms

(ba) Clitic conjunct participle. If the matrix verb of saying is anything other than *guft-an* ‘say’, a conjunct construction using the participle *guft-a* is followed by an inflected form of the specific speech verb:

[*xud-i ū ki-st?*] *gufta man az Rahim Qand pурсid-am*
[self-EZ he who COP.3s] say.PART I from Rahim Qand ask.PT-1s
‘[who is he, actually?] I asked Rahim Qand’.

(bb) A loose conjunct construction using either *guft-a* or *gū-yon* frequently marks an idealized speech string which may be in mixed indirect and direct speech, as a sentential complement (often a purposed phrase or clause).

on-ho grup-a kard-and, [hamin rais-a na-me-xoh-im]
 they group-DO make.PT-3p this very chairman-DO not-IPFV-want.PR-1p
gufia
 say.PART
 ‘they organized [in order to protest (the appointment of) this chairman]’;

[xūrjin-ro ham ba šumo [[mukofot-gūyon] dih-ad]
 saddlebag-DO also to you.p reward say.PART.PR (SBJ.)give.PR-2p
 ‘he’ll give you the saddlebag too, [as a reward]’, lit. ‘... “reward” saying ...’

This is similar to a typically Turkic construction, using *dep* ‘saying’ in Uzbek.

(c) Quotative-evidential phrase.

A speaker repeating someone else’s words to a third party may end with a quotative *buda-ast*:

Pulod ako,
 Polod sir,
[modar-am xūrok tayyor kard-and, roš-ed] bud-a-ast
 mother-my meal prepared made-3p hurry-2p be.PART.EV-COP3s
 ‘Pulod, I’m to tell you [my mother has cooked a meal, and to hurry up]’

(2) Single-clause constructions

In these constructions the verb of the subject and object clause is an infinitive and the agent is converted to a possessive enclitic, and/or *xod* ‘self’:

mo me-don-em ki [kujo me-rav-em] >>
[kujo raftan-i xud]-ro me-don-em
 ‘we know where we are going’ >>
 lit., ‘we know our where-going’;

vay guft ki ba qišloq me-rav-ad >>
vay [ba qišloq raftan-aš]-ro guft
 ‘he said that he was going to the village’ >>
 lit. ‘he said our to the village-going’.

Double nominalization:

darkor nest ki bi-gūy-em [ki in duxtar ki-st] >>
[kī budan-e in duxtar]-ro guftan darkor nest
 ‘it is not necessary that we tell who this girl is’ >>
 lit. ‘saying the who-being of this girl-DO is not necessary’.

6.4.5 “Idiomatic” clitic *ke*

In the Persian colloquial register, the complementizer *ke* may have so-called “idiomatic”, i.e. asseverative function. This is because it does not seem to introduce dependent clauses, but is inserted after the first syntactic constituent of main clauses, mostly brief responses or observations. Pragmatically, it appears to have strong or weak emphatic as well as contrastive function and accordingly tends to be rendered in English by ‘though, however, but, why’ and similar, depending on the context:

in ke kār-i na-dār-e
 this KE work-INDEF not.have.PR-3s
 ‘this, though, does not require any major effort’, or, smoother:
 ‘why, that’s no problem’ (Hayyim, *Farhang-e yekjeldi-e fārsi-englishi*, p. 691).

It has been difficult to identify the syntactic role of this particle. Similar to others, Öhl and Korn (2008: 142) regard this *ke* as homonymous with, but distinct from, the complementizer *ke*, and suggest *Enunziativmarker* for its function.

However, the very designation “idiomatic”, combined with use in the colloquial speech register, should be taken as a hint that there may be no enigmatic non-subordinating function of *ke* involved at all (discussed in Windfuhr 2004). Rather, one expects elliptic constructions, which are typical for that register. Following these observations, it becomes apparent that what is elliptic here are contextually understood stereotype subject and object matrix clauses. The general tenor of such introductory clauses can be easily imagined, given the context of discourse: *ma’lum ast* ‘it’s obvious’, *be nazar-am* ‘in my opinion’, *nist* ‘isn’t it (true)’, *mesl-e in ke* ‘it seems like’ for subject clauses, and *mi-bin-id* ‘you see’, etc., for object clauses. The second position of *ke* is explained by the ellipsis: with the deletion of the introductory phrase, this enclitic necessarily shifts to second position, which may be the clause itself. The syntactic process in the example cited above and some other examples may be shown as follows:

Non- “idiomatic”	“Idiomatic”
<i>be nazar-am [ke šām xord-i]</i>	<i>šām ke xord-i?</i>
in view-my that supper eat.PT-2s	
‘it appears that you did have supper’	‘you did have supper, didn’t you?’
<i>mi-bin-id [ke jā na-dār-e]</i>	<i>jā ke na-dār-e!</i>
IPFV.see.PR-2p that place not-have.PR-3s	
‘you see (yourself) that there is no space!’	‘but there is no space!’
(in response to the suggestion to sit down),	
<i>nist [ke javāb na-dāšt?]</i>	<i>javāb ke na-dāšt</i>
not.COP.3s that answer not.have.PT.3s	(no question contour)
‘isn’t it (true) that there was no answer?’	‘there was no answer; right?’
<i>motma’en am [ke in kār-i na-dār-e] in ke kār-i na-dār-e</i>	
sure COP.1s that this work-INDEF not-have.PR-3s	
‘I’m sure that this doesn’t require much effort’	‘why, that’s no problem’.

6.5 Adverbial clauses

The basic rules for the sequence of matrix and adverbial clauses depends on the logical or temporal relation of their events or states: when the latter precede, the adverbial clauses are in first position; when they follow, their clauses are in second position. Thus, causal, temporal, conditional, and concessive clauses, as well as circumstantial clauses of place, manner, and degree, precede, while consecutive, resultative, and purposive (final) clauses, as well as causal-explanatory, interruptive-temporal, and conditional-exceptional clauses, follow. However, as is the case with relative and complement clauses, both focusing and

topical inversions are frequent, depending on the pragmatic context. In Tajik, these finite adverbial clauses compete with nominalized adverbial clauses.

There are four types of subordinators: parataxis; the general complementizer *kelki* (roughly comparable to English conjunctive 'as'); a small set of three primary conjunctions, with or without the complementizer, as well as two interrogatives; conjunctive phrases with the complementizer.

Parataxis and kelki occur prominently in colloquial speech. Being lexically empty, they may substitute for a number of other subordinators, depending on the pragmatic context:

pul ke mi-deh-i, mi-ā-y-am
 money KE IPFV-give.PR-2s, IPFV.come.PR-1s
 '(as) you give me money, I come', more specifically,
 'because ~ when ~ if ~ even though you give me money for this, I will come'.

The primary conjunctions include:

- (a) *čunlčun* 'because, since', also 'when' (only literary);
- (b) *vaqt-i/vaqt-e* 'when' (< conjunctive phrase *vaqt-i kel/vaqt-e ki*).
- (c) *tālto*, (ca) temporal and conditional duration or end point, 'till, until, as long ~ soon as' (note *tā na-lto na-* 'un-less'); (cb) comparison 'than';
- (d) *agarlagar* 'if';
- (e) *magar/magar* 'unless';
- (f) *zirālzero* 'for this reason, because';

Conjunctive phrases are numerous, and syntactically identical with the heads of relative clauses (see section 6.3) such as:

(PREP) N- <i>il-e</i>	<i>kelki,</i>	<i>(dar) vaqt-i kel/vaqt-e ki</i> 'at time that' = 'when';
(PREP) N- <i>el-i</i>	<i>in kelki,</i>	<i>(be) ellat-e in kel (ba) illat-i in ki</i> 'for the reason that' = 'because';
PREP	<i>in kelki,</i>	<i>bā in kelbo in ki</i> 'with this that' = 'although';
N *PREP	<i>in kelki,</i>	<i>ba'd az in kel/ba'd az in ki</i> 'after from this that' = 'after doing . . '

* The nominal may be a noun, adjective, or adverb.

6.5.1 Position and use of TMA

(1) Position. *Kelki* is an unstressed enclitic. In sentence-initial subordinate clauses it is typically inserted after the first constituent component of the clause; in sentence-final position, it is attached to the preceding matrix clause (similar to the enclitic coordinator *-ol-u* 'and'). Other clausal heads may optionally be in second position as well (see section 5.1 Order of constituents).

(2) Mood. The use of mood is predictable: Events or states require the indicative when real (factual), the subjunctive when potential, and the counterfactual when unlikely:

<p>Real <i>mi-xāh-i, mi-rav-iml</i> <i>me-xoh-ī, me-rav-em</i> you want IPFV.go.PR-1p ‘if you really want to, we will go’;</p> <p>Counterfactual <i>man mi-raft-aml</i> <i>man me-raft-am</i> I IPFV.go.PT-1s ‘I would go (if I were you)’.</p>	<p>Potential <i>raft-am, tā ~ ke az u be-pors-aml</i> <i>raft-am, to ~ ki az ū (bi-)purs-am</i> go.PT-1s till ~ that from he SBJ.ask.PR-1s ‘I went in order to ask him’.</p>
---	--

(3) Aspect is predictable as well. Notably the preterite (i.e. the perfective aorist form in the verb system) is required when an event or state is completed prior to that of the matrix clause (while English uses the present). Conversely, incomplete events or states require the subjunctive, notably those headed by *piš az in kel pēš az in ki* ‘before’, *be jā-ye in kelba jo-yi in ki* ‘instead of’, *bedun-e in kelbidun-i in ki* ‘without’:

<p><i>agar u-rā did-i, be-gu</i> <i>agar ū-ro did-ī, bi-gū</i> if he-DO see.PT-2s IMP.say-2s ‘once you see him, tell me’;</p>	<p><i>bedun-e in ke čiz-i be-gu-yad, raft</i> <i>be on ki čiz-e (bi)-gū-yad, raft</i> without this that anything SBJ.say.PR-3s go.PT.3s ‘he left without saying anything’.</p>
--	---

The basic unmarked patterns of clauses, together with verbal modality (the Persian *consecutio temporum*), may be shown as follows (cf. Windfuhr 1987: 539/2009: 457). Note that in each column, the subordinators for pre-matrix position may also occur in post-matrix position.

TABLE 8.17: PATTERNS OF CONJUNCTIONS AND MOODS

<i>Preceding Main Clause</i>		<i>Following Main Clause</i>					
Cause	<i>čunlčun</i>	‘because’	IND	Explanation	<i>zirā, zero</i>	‘(that is) because’	IND
Time	<i>vaqtilvaqte</i>	‘when’	IND	Interruption	<i>kelki</i>	‘when (suddenly)’	IND
Point/Stretch	<i>tāto</i>	‘as soon/ long as’	IND	End point	<i>tāto</i>	‘until (finally)’	IND
Condition	<i>tāto,</i> <i>tā na-l</i> <i>to na-</i>	‘until’ ‘unless’	SBJ SBJ	Purpose	<i>tāto</i>	‘in order to, so that’	SBJ
Condition	<i>agarlagar</i>	‘if’	SBJ, CND	Exception	<i>magar/</i> <i>magar</i>	‘unless, if not’	SBJ

6.5.2 Pre-matrix clauses

For these, see also the paradigm of relative clauses in Table 8.14.

6.5.2.1 Locative clauses

These are plain relative clauses:

be jā-i ke raft-i (be ān-jā) bar gard!!
ba jo-'e ki raft-i (ba on jo) bar gard!
 to place that go.PT-2s to there IMP.return.2s
 'go back to where you come from!'

6.5.2.2 Temporal clauses

These are mostly introduced by *vaqt-i (ke)/vaqt-e (ki)*. The noun may be substituted by similar time-nouns, such as *mowqe'-i kelzamon-e ki* 'the time, moment when'. Time extension may be indicated by *moddat-i kelmuddat-e ki* 'period, duration'. These may be preceded by prepositions, *dar, az Tldar, az T* 'at T, from T, since'. Punctual events are introduced by *hamin kelhamin ki* 'the very same that, just as', and simultaneous events by *dar hāl-i keldar hol-e ki* 'in the situation that, while'. Generalized events use *har vaqtlhar vaqt* 'each, every time, whenever'.

Posterior and anterior events use the adverbial phrases *ba'd ~ pas az ānlba'd ~ pas az on* 'after from that, after (he did ~ doing . . .)' and *pas ~ piš az ān kelpas ~ peš az on ki* 'before from that, before (he did ~ doing . . .)', which require the subjunctive as they refer to potential actions:

piš az ān ke ān kār-rā be-kon-i, fekr kon!
peš az on ki on kor-ro kun-ī, fikr kun!
 before from that that that work SBJ.do.PR-2s thought IMP.make.PR.2s
 'before you do that, think!';

to vazife-ye xod-at-rā tamām na-kon-i,
to vazifa-i xud-at-ro tayyor na-kun-ī,
 until dutyEZ self-your-DO accomplished not(SBJ.)-make.PR-2s
be kuče na-row!!
ba kuča baromada na-rav!
 to alley out.come.PART not.IMP.go.2s
 'until you've done your homework, don't go outside!'

6.5.2.2a Since-clauses

For the expression of the time since an event is taking or has taken place, Common Persian uses a subject clause construction of the type 'it BE(-COME) X time that':

panj sāl ast (~ mi-šav-ad) ke u-rā na-dide am
panj sol ast (~ me-šav-ad) ki ū-ro na-dide am
 five years COP.3s IPFV.become.PR-3s that he-DO not.see.PART COP.1s
 'I haven't seen him in ~ since ~ for about five years'.

6.5.2.3 Conditional clauses

These are introduced by *agarlagar* 'if', *har gāhlhar goh* 'as soon as, when, if'. In colloquial Tajik speech *agarlagar* may be placed at the end of the protasis, after the verb, similar to Uzbek.

Real conditions require the indicative, followed by the indicative present or imperative:
Present indicative

agar u-rā dust mi-dār-i, komak kon!
agar ū-ro dūst me-dor-i, yori dih!
if he-DO friend IPFV.have.PR-2s help do/give.IMP.2s
'if you really like him, help him!'

Preterite (perfective aorist)

agar u-rā did-i, be-gul
agar ū-ro did-i, gūy
if he-DO see.PT-2s IMP-2s
'once you see him, tell me'

Potential conditions require the subjunctive, also followed by the indicative present or imperative:

agar u-rā be-bin-id, salām-e ma-rā be-ras-ān-id!
agar ū-ro bin-ed, salom-i ma-ro ras-on-ed
if her SBJ-see.PR-2p see salām-EZ we-DO SBJ-taketo.PR-2p
'if you see her, give her my regards'

Counterfactual conditions have the imperfective past (in its modal function) in both clauses:

agar hamrāh-e man mi-raft-id, u-rā mi-did-id!
agar hamroh-i man me-raft-ed, ū-ro me-did-ed
if along-EZ I IPFV.go.PT-2p he-DO IPFV.see.PT-2p
'if you were to come with me, you would see him' or
'if you had come . . . you would have . . .'

Past anterior time may be specified by the use of a pluperfect in its modal function in one clause:

agar ham-rāh-i man raft-e bud-id . . .)
agar ham-roh-i man raft-a bud-ed . . .)
if along-EZ I go.PART be.PT-2p
'if you had gone with me . . .'

In Tajik, logical, or quasi-conditionals (which connote 'since, as appears, it is the case that', etc., use the durative perfect subjunctive (not found in Persian) for present reference:

agar ham-digar-ašon-ro dūst me-došt-a boš-and, digar čī?
if together-their-DO friend IPFV.have.PART (SBJ).be.PR-3p further what
'if they love each other, what of it?'

They use the past subjunctive for past reference in the protasis:

in kitob-ro xond-a boš-ed agar, čaro na-ovard-ed?
this book-DO read.PART (SBJ).be.PR-2p if why not.bring.PT-2p
'if you've read this book, why didn't you bring it?'

6.5.2.4 Concessive clauses

These are introduced by the following phrases, all ‘though, although, despite the fact that’: *agar-čelagar-či*, lit. ‘if, though’; *har čand kelhar čand(-e ki)*, lit. ‘however much that’; *bā (vojud-e) in kelbo vujud-i) in ki*, lit. ‘with the existence of this that’. They usually take the present or past subjunctive, according to time reference. The matrix clause may be introduced by *valilvale*, *ammālammo*, *likenllekin* ‘but, still’:

agar če qablan šomā-rā na-did-e bāš-am,
agar či peš-tar šumo-ro na-did-a boš-am,
 although earlier you.p-DO not.see.PART (SBJ).be.PR-1s
 (vali) *esm-e šomā be man āšnā budl*
 (vale) *nom-i šumo ba man ma'lum bud*
 (yet) name-EZ you.p to me familiar/clear be.PT.3s
 ‘even though I hadn’t seen you before, still your name was familiar to me’.

Tajik post-clausal *-ham* ‘also, even’

Frequent and peculiar to Tajik is use of the postposed particle *ham* as a concessive:

havo xunuk na-boš-ad ham,
 weather cold not(SBJ).be.PR-3s also,
barf bo dona-ho-i kalon-kalon-i laklakī me-borid
 snow with piece.p-EZ big-big-EZ fluffy IPFV.rain.PT.3s
 ‘although the weather was not cold, the snow was falling in large, fluffy flakes’;

note the generalized present subjunctive instead of the perfect subjunctive.

6.5.2.5 Causal clauses

These are introduced by: *čun kelčun (ki)* ‘since, as’; *mādām kelmodom-e ki* ‘while, since, inasmuch’; *barā-ye in kelbaro-i in ki* ‘for the sake of, because’ (see also section 6.5.3.4 Purpose clauses) with tense and aspect according to sense:

mādām ke ū āmad, be u be-gu-id ke in-jā tu bi-ā-yadl
modom-e ki ū omad, ba ū gū-ed ki injo dar-o-yad
 while EZ KE he come.PT.3s to he say.IMP-2p that here in SBJ.come.PR-3s
 ‘since he’s arrived, tell him to come in’.

6.5.3 Post-matrix clauses

6.5.3.1 Causal-explanatory clauses

These are introduced by: *čun kelčun (ki)*, lit. ‘because (that)’; *zirā/zero (kelki)*, lit. ‘for this (that)’; *čerā kelčaro ki*, lit. ‘what for that’; *banā bar in kelbino bar in ki*, lit. ‘built on this that’, all ‘because, therefore’:

ketāb-etān-rā be u na-dād-am, zirā ke u-rā na-did-aml
kitob-aton-ro ba ū na-dod-am, zero ki ū-ro na-did-am
 your book-DO to he not.give.PT-1s because he-DO not.see.PT-1s
 ‘I didn’t give him your book, because I didn’t see him’.

6.5.3.2 Temporal clauses

Post-matrix temporal clauses are syntactically focused clauses:

mi-xāst-am xāb be-rav-am ke kas-i dar zadl
me-xost-am xob rav-am ki kas-e dar zad
 IPFV-want.PT sleep SBJ-go.PR-1s that someone door knock.PT.3s
 'I was about to go to bed when someone knocked at the door'

6.5.3.2a Subordinator switch

A distinct syntactic operation involves the switch of the subordinator between two clauses. This is found typically in subsets where an ongoing or existing situation, or an event that is not yet completed, is interrupted by another event. When the latter is in post-matrix position, it is introduced by a simple *kelki*:

dar ān-jā nešaste bud-am, u dar āmadl
dar on jo nišasta bud-am, ū daromad
 I there sit.PART be.PT-1s he come.in.PT.3s
 'I was sitting there, and he came in'

>>

vaqt-i dar ān-jā nešaste bud-am, u dar āmadl
vaqt-e dar on jo nišasta bud-am, ū daromad
 unmarked, 'while I was sitting there . . .'

>>

dar ān-jā nešaste bud-am, ke u dar āmadl
dar on jo nišasta bud-am, ki ū daromad
 marked: '. . ., when he suddenly came in'

6.5.3.3 Result clauses

These are introduced by *tālto*:

ān-jā nešaste bud-am, belāxare āmad
on jo nišasta bud-am, biloxara omad
 there sit.PART be.PT-1s finally come.PT.3s
 'I sat there, until he finally arrived'

6.5.3.4 Purpose clauses

These are introduced by the following, all 'so that, in order to, to (do)': *kelki*, *tā kelto (ki)*, *barā-ye in kelbaro-i in kilbaro-i on ki*. Of these, *tālto* is mainly literary. All require the present subjunctive, irrespective of the matrix tense:

āmad-am, ke ~ tā u-rā be-bin-aml
omad-am, ki ~ to ū-ro bin-am
 I came that ~ till he-DO SBJ-see.PR-1s
 'I came to see him';

čandān kon, ke jā bāš-ad
čandon kun, ki jo boš-ad
 thus do.IMP.2s that place (SBJ).be.PR-3s
 'make it so, that there is space'

Colloquially, the purpose clause may be juxtaposed without a conjunction:

piš-e mehmān-at bo-row, az to nārāhat na-šav-ad
peš-i melmon-at rav, az to xafa na-šav-ad
 front-EZ guest-your.s IMP.go.2s from you.s offended not.(SBJ.)-become.PR-2s
 ‘go over to your guest, so he won’t be offended’. See also section 6.4.2 Embedded
 speech).

6.5.3.5 Consecutive clauses

These are introduced by *tā (in ke)to (in ki)* ‘until, as far as’, *čonān . . . kelčunon . . . ki, be qadr-i . . . kel (ham-)in qadar . . . ki* ‘so (much) that’, *towr-i ke* ‘in a manner that’, so’:

raft-o raft, tā be jangal-i rasid
raft-u raft, to ba jangal-e rasid
 go.PT.3s-and go.PT.3s till to forest-INDEF reach.PT.3s
 ‘on he went until he reached a forest’;

u čonān tond pašme mi-čīn-ad, ke hame hayrān mi-mān-and
ū čunon tez paxta me-čīn-ad, ki hama hayron me-mon-and
 she so fast cotton IPFV.pick.PR-3s that all astounded IPFV.remain.PR-3p
 ‘she picks cotton so fast that everyone is astounded’.

6.5.3.6 Comparative clauses

In Persian, these are introduced by *tālto*. Tajik has instead a single-sentence infinitival construction, *az INF dida* ‘seen from (his doing)’:

u bištar mi-xāb-ad, tā kār mi-kon-ad
 he more IPFV.sleep.PR-3s than work IPFV.do.PR-3s
 ‘he sleeps more than he works’;

vs. Tajik

ū az kor kardan did-a beštar xob me-rav-ad
 he from work do.INF see.PART more sleep IMP.go.PR-3s.

6.5.3.7 Conditional exception clauses

While *agarlagar* may introduce post-matrix conditional clauses, exceptional conditions are introduced by its negation, *magar/magar* ‘unless’:

man ne-mi-rav-am magar to ham bi-ā-yi
man na-me-rav-am magar tu ham ā-yi
 I not.IPFV.go.PR-1s unless you.s also SBJ.come.PR-2s
 ‘I will not go, unless you come, too’.

6.5.4 Tajik Turkic-type nominal clauses

As shown in some of the discussions above, the final clausal constructions represent only part of Tajik’s syntactic repertory. Most of them can be recast in the Turkic mold as a single sentence containing one or more non-finite verb forms representing nominalized VPs.

The means of nominalization is the nominal EZ which has greatly extended its range of application (arguably under the influence of Uzbek), by the conversion of the verb phrase of finite subordinate clauses to non-finite verb forms, especially (1) participle II in *-agī* and (2) infinitives (gerunds), thereby condensing into a single sentence what would appear more naturally in Persian or English as a matrix-plus-subordinate clause.

6.5.4.1 Participial relative clauses in *-agī*

(1) In the basic pattern, the augmented past participle in *-agī* (which is either active or passive in voice, and may be tensed analogously with the complex tenses which it helps to form, may be used in EZ and other constructions to perform the functions of a relative clause. Note that these participles retain the modal-aspectual distinction of perfective progressive, and conjectural (see section 6.3.7.1 above):

Perfective form	Progressive form
<i>kitob-i [man ovard-agī]</i>	<i>in duxtar-i [kitob xond-a istod-ag-ī]</i>
book-EZ [I having brought]	this girl-EZ [book reading-PROG]
'the book that I brought';	'the girl who is reading the book';
Conjectural form	
<i>zan-i [sar me-dod-agī]</i>	
wife-EZ [to divorce-PR-CONJECT]	
'a wife [to divorce]' (<i>sar dod-an</i> 'let go').	

(2) Inversions and topicalized relative constructions

The nominalization of a relative clause ranges from retention of the agent and word order, to the replacement of the agent pronoun to a pronominal clitic, to the fully Turkicized inversion:

HEAD-EZ [Clause] > [Clause] HEAD	
<i>kitob-i [man ovard-agī]</i>	book-EZ [I having brought] 'the book that I brought'.
<i>kitob-i [ovard-agi-am]</i>	book-EZ [my having brought-]
<i>[ovard-agi-am] kitob</i>	[my having brought] book.

6.5.4.2 Infinitival clauses

The infinitive, often in combination with conjunct verb forms, can participate in quite complex, nested NPs as sentential complements.

6.5.4.2a Object clauses

Infinitival object clauses are marked by the direct object marker *-ro* and precede the main clause:

(1) 'how do you know that doing this will not be worthwhile?'
az kujo me-don-i [(ki) az in kor foida-yaš na-bar-o-yad] >>
[az in kor foida na-bar-omad-an-aš]-ro az kujo me-don-ī?

from where do you know [that from this work benefit-its may not come out]
 [from this work benefit not-issuing-its]-DO from where do you know?

(2) ‘of course one needn’t tell anyone who this girl is’.
albatta darkor nest [ki gū-yem] [in doxtar kī-st]] >>
albatta [kī bud-an-i in duxtar] -ro guftan] darkor nest

of course necessary not is [that we say [this girl who is]]
 of course [[who being-EZ this girl]-DO saying] necessary not is

(3) ‘this excessive wheat isn’t worth for us to take back’.
in gandum-i ziyodatī na-me-arz-ad [boz gašt-a girift-a bar-em] >>
in gandum-i ziyodatī [ba [boz gašt-a girift-a burdan]] na-me-arz-ad

this wheat excessive is not worth [having taken back that we carry]
 this wheat excessive to [returned CONJUNCT taking] is not worth
 (*ba . . . arzid-an* ‘be worthy of, to’).

6.5.4.2b Adverbial clauses

In adverbial infinitival clauses, the infinitival clause is headed by a nominal replacing the conjunction, here with the example of a temporal clause:

‘when the lamp was being taken out of the room, its glass broke’;
dar vaqt-i [ki lamp-ra az xona bar-ovard-and]] šiša-aš šikast >>
[dar vaqt-i [az xona bar-ovardan-i lamp-ra]] šiša-aš šikast

at the time [that the lamp-DO from the room they took out] glass-its broke
 [at the time-EZ [from the room taking out-EZ the lamp-DO]] glass-its broke

6.5.4.3 Embedded infinitival clauses

With the concomitant gapping of an anaphoric independent or enclitic pronoun, infinitival clauses may further be embedded by an EZ construction:

‘we heard the sound of that man’s footsteps (*po*) as he was coming downstairs’.
[dar vaqt-e ki on kas poin me-furomad] sado-i po-i on kas-ro šiunid-em >>
[[dar vaqt-e poin furomadan-i on kas] sado-i po-i; vaγ]-ro šiunid-em
[sado-i po-i [poin furomadan-i on kas]]-ro šiunid-em

[[at time **when** that one **was coming down**] sound-EZ foot-EZ **that one**]-DO we heard
 [at **time-EZ coming down-EZ that one**] sound-EZ foot-EZ **him**-DO we heard
 [sound-EZ foot -EZ [**coming down-EZ that one**]]-DO we heard
 ‘we heard the sound of the footsteps of that one’s coming downstairs’.

Note the EZ embedding of infinitival temporal clause and the gapping of the anaphoric pronoun.

6.5.4.4 Persian conjunct participle

Persian allows for the participialization of finite verbs in sequences of clauses by reducing finite verbs to participles in all clauses except the last (known as *hazf-e fe’l* ‘ellipsis of tense-mood-aspect and person marking’). This strategy rarely includes subordinate clauses. It is well known from classical texts, and in contemporary Persian is confined to literary use. The following is an (abbreviated) example from an Iranian journal about old

dialect poetry (*fahlaviyāt*) found in various sources (M. Rezāyati Kische Khāle, *Guy-eššenāsil/Dialectology* 2.1 (Feb. 2006): 129).

Doktor Ali Ašraf Sādeqi, ke sāl-hā-st tashih-e fahlaviyāt . . . rā
Dr. Ali Ashraf Sadeqi who years is (that) emendation-EZ fahlavi poetry-DO

vajh-e hemmat-e xod qarār dāde,
direction-EZ effort-EZ-self own resolve having given

axiran . . . in yāzdah fahlavi-rā . . . tashih karde
recently this eleven fahlavis-D correction having made

va dombāl-e har do-beyti tahqiq-i mostowfi afzude
and after-EZ each two-liner research-INDEF comprehensive having added

va be bahs-i kolli pardāxte ast
and to discussion-INDEF general has undertaken

‘Dr. Ali Ashraf Sadeqi, who for many years has made the emendation of the *fahlaviyāt* the focus of his scholarly work, has recently not only corrected these eleven *fahlavis*, but has also been engaged in extensive discussions of the problems involved in general.’

7 LEXIS AND SOCIOLINGUISTICS

7.1 Word formation

The basic procedures of derivation and compounding are the same in Tajik and Persian. However, as the result of a degree of cultural separation during the last five centuries, and a flurry of conscious language planning in Tajik over the past 75 years, there are considerable differences of detail.

7.1.1 Homonymy and conversion

There is a fair amount of tolerance for homonymy, which is disambiguated syntactically. A word such as *xāb/xob* may be a noun with two distinct meanings, ‘sleep’ and ‘dream’; it may additionally function as a predicative adjective: *xāb ast/xob ast* ‘he is asleep’. The noun *ārāmlorom* ‘quiet, calm, tranquility’ may also be the adjective (both predicative and attributive) ‘quiet, calm, tranquil’.

Conversion without morphological change occurs most readily from adjective to noun status: *mard-e javān/mard-i javon* ‘young man’, *yek javān/yak javon* ‘a youth’, *javān-ān/javon-on* ‘young people, the young’.

7.1.2 Suffixation

Suffixation is the principal means of lexical derivation; more than 40 nominal suffixes may be identified. The following list is limited to those regarded as most productive:

-ilī (-g-il-g-ī after the vowel -a-): forms abstract nouns of quality from adjectives and type nouns, of activity from agentives, etc.: *pir-ilpir-ī* ‘old age’ < *pir/pir* ‘old; old man’; *zende-g-ilzinda-g-ī* ‘life’ < *zendelzinda* ‘alive, living’. Such nouns readily evolve concrete senses, as *širin-ilširin-ī* ‘sweets, candy (sweetness)’ < *širin/širin* ‘sweet’.

-kārīkōr, -garī-gar, -čīl-čī: form agent nouns and agentive adjectives from nouns of entity or activity (occasionally, adjectives): *xeyānat-kārīxīyonat-kōr* ‘traitor, treacherous’ < *xeyānat/xīyonat* ‘betrayal’; *kār-garīkōr-gar* ‘(manual) worker’ < *kārīkōr* ‘work’; *hīle-garīhīla-gar* ‘deceitful, cunning’ < *hīle/hīla* ‘trick, deception’; *xedmat-čīlxīzmat-čī* ‘employee, clerical worker’ < *xedmat/xīzmat* ‘service, employment’; Tajik *a’lo-čī* ‘honor student’ < *a’lo* ‘superior, excellent’.

-akl-ak:

(a) added to entity nouns and NPs, forms diminutives, often endearing or derogatory in connotation: *pesar-akl/pisar-ak* ‘laddie, kid’ < *pesar/pisar* ‘boy’; Taj. *modar-i mehrubon-ak* ‘mommy dear’ < *modar-i mehrubon* (EZ phrase) ‘kind mother’; Tajik has an intensive form, *-akak*: *kam-akak* ‘a pinch, spot’ < *kam* ‘little (in quantity)’;

(b) added to verbal agentives, it forms concrete instrumentives: Taj. *ob-poš-ak* ‘watering can’ < *ob* ‘water’ and Stem I of *pošid-an* ‘spray’; *band-ak* ‘pin, clasp’ < Stem I of *bast-an* ‘fasten’.

-čel-ča: forms diminutives from nouns and adjectives (neutral or endearing): *ketāb-čel kitob-ča* ‘booklet’ < *ketāb/kītib* ‘book’; Taj. *duxtar-ča* ‘lassie’ < *duxtar* ‘girl’; Taj. *safed-ča* ‘whitish’ < *safed* ‘white’.

-el-a: a nominal suffix from several ultimate sources, still productive in some functions (metonymic or specifying): *dast-eldast-a* ‘handle’ < *dast* ‘hand’; *xandelxanda* ‘laugh(ter)’ < *xand-id-an/xandid-an* ‘laugh’; *geryelgiry-a* ‘weeping’ < *gerist-an/girist-an* ‘weep’. Numerical expressions are made adjectival or adverbial: *do-ru-yeldu-rū-ya* ‘two-faceted, two-faced, hypocritical’ < *doldu* ‘two’, *rulrū(y)* ‘face’; *panj-sāl-elpanj-sol-a* ‘of five years (adj.)’, a five-year-old, five-year plan’ < *panj/panj* ‘five’, *sol* ‘year’.

-stānlston (after a consonant, *estānl-iston*) forms locative nouns, esp. names of countries: Pers. *kudakestan* ‘kindergarden’ < *kudak* ‘child’, *Baluchestānl/Balochiston*, *Baluchistan*’.

The following suffixes form adjectives from nouns:

-il-ī (after a vowel, *-g-il-g-ī* or *-v-il-v-ī*): the most general and neutral relative adjective, originating in both Middle Persian *-īk* and Arabic *-īyy*, and readily substantivized: *Irān-il’eron-ī* ‘Iranian, an Iranian’; *xāne-g-ilxona-g-ī* ‘domestic, household’; *šowra-v-ilsovet-ī* ‘deliberative; Soviet’ < *šowrālšavro* ‘council’; Taj. *partiya-v-ī* ‘(of the) Party’ (*billet-i partiya-v-ī* ‘Party card’).

Taj *-angī* (after a vowel, *-ngī*): forms relative adjectives, especially in colloquial Tajik: *dina-ngī* ‘yesterday’s’ < *dina* ‘yesterday’.

-ānel-ona (after a vowel, *-g-ānel-g-ona*) ‘characteristic of ~’: forms adjectives applied to non-humans, frequently used as nouns or adverbs: *āqel-āneloqil-ona* ‘intelligent, sensible’ < *āqelloqil* ‘intelligent’ (of a person); e.g. *kār-e āqel-ānelkor-i oqil-ona* ‘sensible action’; *bačče-g-ānelbača-g-ona* ‘children’s; childhood ~; childlike, childish’ < *bačče/bača* ‘child’; *māh-ānelmoh-ona* ‘monthly; salary’ (see also section 3.1.5 Adverbs).

-nākl-nok ‘possessing the quality of ~’, *xatar-nākl/xatar-nok* ‘dangerous’ < *xatar/xatar* ‘danger’: a very productive suffix in Tajik, generating adjectives or nouns from intangibles: *foida-nok* ‘useful, advantageous’ < *foida* ‘benefit’; *xarakter-nok* ‘characteristic, specific’ < *xarakter* ‘character(istic)’; *sado-nok* ‘vowel’ < *sado* ‘sound, voice’.

7.1.3 Prefixation

Productive nominal prefixes (mostly identical with prepositions) which form adjectives and related nouns are:

bālbo- ‘with’: *bā-este’dādībo-iste’dod* ‘talented’ < *este’dādīste’dod* ‘talent’.

bi-lbe- 'without': *bi-adablbe-adab* 'discourteous, impolite' < *adabladab* 'manners, courtesy'.

nā-lno- 'not, un-': *nā-dān/no-don* 'ignorant, ignoramus' < Stem I of *dānest-an/donist-an* 'know'; *nā-omidlno-umed* 'hopeless, desperate' < *omidlumed* 'hope'.

ham-lham- 'together', a particle connoting sharing, reciprocity or 'cooperation': *ham-rāhlham-roh* 'companion, fellow-traveler' < *rāhlroh* 'road, journey'; *ham-āhanglham-ohang* 'harmonious, consonant' < *āhanglohang* 'melody'.

kam-lkam 'little' (quantity): *kam-zur/kam-zur* 'weak' < *zur* 'strength; Pers. *kam-harf* 'taciturn' < *harf* 'word'.

ser- 'full, sated', rare in Pers.: *sir-āblser-ob* 'saturated' < *āblob* 'water'; Taj. *ser-odam* 'crowded, populous' < *odam* 'person'; *ser-masraf* 'often used, in demand, popular' < *masraf* 'consumption, utilization'.

por-lpur- 'full': *por-mā'nālpur-mā'no* 'meaningful' < *mā'nā* 'meaning' (literary variant of *mā'ni*); *por-šokuhl pur-šukūh* 'splendid' < *šokuhlšukūh* 'splendour'.

Prepositional EZ constructions (the following all Arabic loanwords) include:

yejyr-e-qānun-i/γayr-i-qonun-i 'illegal' < *qānun/qonun* 'law'.

zedd-elzidd-i- 'against, anti-, counter-': *zedd-e-havā'i/zidd-i-havo-i* 'anti-aircraft' < *havālhavo* 'air'.

ba'd-i- 'after': *ba'd-i jang-i* 'postwar' < *jang* 'war'.

Tajik *to-* 'up to, until': *to-maktab-i* 'pre-school' < *maktab* 'school'; *to-inqilob-i* 'pre-revolutionary' < *inqilob* 'revolution'. Note that this prefix qualifies a ready-formed adjective, as do the preceding three. (This use of *tā* is unknown in Persian, and was probably influenced by identical Russian use of *do-* 'up to, until'.)

7.1.4 Compounding

Compounding is most frequently effected by simple juxtaposition of stem nominals, or a nominal and a verbal radical, modifier before head. Such compounds are of two kinds:

Determinative, where the compound is of the same lexical category as the head (a noun): *hafte-nāmelhafta-noma* 'weekly (periodical)' < *haftelhafta* 'week' + *nāmelnoma* 'document, letter'; *piš-raftlpeš-raft* 'progress' < *pišlpeš* 'forward' + *raftlraft*, Stem II of *raft-anlraft-an* 'go'; *dast-nevisldast-navis* 'manuscript' < *dastldast* 'hand' + *nevislnavis*, Stem I of *nevešt-anlnavišt-an* 'write'. Structures of this last type are more commonly agentives or instrumentives: *kafš-duzlmūza-dūz* 'cobbler' < *kafš* 'shoe', *mōza* 'boot' + Stem I of *duxt-anldoxl-an* 'sew'; *bārān-sanjlboron-sanj* 'rain gauge' < *bārānlboron* 'rain' + Stem I of *sanjid-anlsanjid-an* 'measure'.

Possessive, generally where the modifier is an adjective or adverb, and the resulting compound refers to a third entity possessing, or characterized by, the compound quality (fundamentally, an adjective): *čahār-pālčor-po* 'quadruped, animal' < *čahārlčor* 'four' + *pālpo* 'foot, leg'; *seyāh-čašmlsiyoh-čašm* 'black-eyed' < *seyāhlsiyoh* 'black' + *čašmlčašm* 'eye'; *xoš-zabānlxuš-gap* 'well-spoken, courteous' < *xoš/xuš* 'good' + *zabān* 'tongue, language', *gap* 'speech'.

Modifier and head are occasionally reversed: *del-saxtdil-saxt* 'cruel' < *delldil* 'heart' + *saxtlsaxt* 'hard'; this is usual when the modifier is a participle: Pers. *ru-bast-e* 'veiled' < *ru* 'face' and Stem I of *bast-an* 'tie, close', Taj. *rū-toft-a* 'disobedient' < *rū* 'face' + past participle of *toft-an* 'turn (away)'.

There are a few common *copulative* compounds, where two nouns or verb stems are joined by addition of enclitic *-ol-u* 'and' to the first element: *āb-o havālob-u havo*

'weather', lit. 'water-and-air'; *raft-o-āmad/raft-u-omad* 'visit(ing), traffic' < Stem II *raft-an/raft-an* 'go', and Stem I of *āmad-an/lomad-an* 'come'; Taj. also *raft-u(o)y*).

Stems II of verbs. Adjectives and adverbs are similarly formed by compounding or reduplication of nouns with a connective element: *pey dar pey/pay dar pay* 'continuous(ly)' (*pey/pay* 'track, (in) pursuit, after'); Taj. *xel-ma-xel* 'various, of all sorts' (*xel* 'sort, kind').

Adjectives may be intensified by similar means: *dur-o-derāz/dur-u daroz* 'lengthy' (*dur* 'far', *daroz* 'long'), Taj. *garm-o-garm* 'extremely hot' (*garm* 'warm, hot'), but note Persian EZ construction *garm-e garm*.

By adding prefixes and suffixes to these structures, neologisms can be formed: *havā-peymā-bar* 'aircraft carrier' (*havā-peymā* 'air-plane', *bar-* 'bear, carry').

Tajik can generate neologisms of almost Russian length and complexity, a development that has been encouraged from the Soviet period on. Examples are: *mablay-jūdo-kun-ī* 'appropriation, disbursement of funds' (sum-separate-mak-ing); *avtomobil-kor-kard-a-bar-or-ī* 'automobile production' (auto-work-done-out-bring-ing).

Echoic reduplication is used in several ways, mostly in the colloquial language. Collective or generalized nouns are generated by repetition with a change of initial, to add a rhyming or echoic nonsense-word: *ketāb-metāb/kitob-mitob* 'books and papers'; Pers. *eynak-meynak* 'spectacles, glasses'; Taj. *bača-kača* 'kids and the like' < *bača* 'child'; *oš-poš* 'food and stuff' < *oš* 'food'; *mayda jūyda* 'odds and ends' (*mayda* 'small'; *jūyda, mitob*, etc. have no independent meanings).

7.2 Loanwords

Arabic vocabulary is the oldest and still the largest foreign element in the lexicon. In literary Persian it constitutes about 50 percent, in spoken about 25 percent.

7.2.1 Arabic

In Iran, official organizations, including the Farhangestān-e zabān-e Fārsi, targeted especially this stratum for replacement by native vocabulary during the 1930s and 1940s, and in Tajikistan during the Soviet language reforms of the same period; however, the presence of many everyday Arabic words for which there is no ready Persian equivalent (e.g. *ketāb/kitob* 'book', *havā/havo* 'air, weather') has mitigated the effects of purification.

Some specialized uses of Arabic and Arabicate plural forms continue to be used in Iran and have survived the general Soviet condemnation of archaisms in Tajikistan, since they have been lexicalized and fill a useful niche. Words with the suffixes *-āt/ot*, *-j-āt/j-ot*, Taj. *-v-ot* denote collectivities: *heyvānāt/hayvon-ot* 'animals, fauna', *mive-j-āt/meva-j-ot* 'fruit(s)', Pers. *sabz-ij-āt*, Taj. *sabza-v-ot* 'vegetables'. In Tajik, some plurals have developed a singular meaning: *taškil-ot* 'organization', *hašar-ot* 'insect' (with regular plurals *llarbāb-hā, taškil-ot-ho* and *hašar-ot-ho*).

Arabic "broken plurals" are often lexicalized, with collective or singular meanings: *atrāflatrof* 'environs, neighborhood' (cf. *taraf* 'side, direction'), *a'zāla'zo* 'member' (of an institution), *a'zā-ye badan/a'zo-i badan* 'parts of the body'; Pers. *arbāb* 'landowner', boss (cf. *rabb* 'Lord [God]'); Taj. *talaba* 'student' (cf. *tolib* '(religious) student').

7.2.2 *Turkic*

While Persian has integrated a considerable number of Turkic nominals (including a few of Mongol origin) such as *otāqlutoq* Pers. ‘room’ vs. Taj. ‘house, place’ and titles like *āqāl ogo* ‘sir, Mr.’, the Northern Tajik dialects are replete with pre-Uzbek Turkic and Uzbek vocabulary, even at the level of function words; a number of common loans have entered the literary language and are also widespread in Southern speech, e.g. *boy* ‘rich’, *tūy* ‘wedding, circumcision celebration’, *yaroq* ‘weapon’, *yordam* ‘help’, *qišloq* ‘village’ (Pers. *qešlāq* ‘winter quarters’), and several kinship terms such as *uka* ‘younger brother’, *yunga* ‘sister-in-law’ (see section 7.4.3).

7.2.3 *Russian*

In Tajik, the large Russian component pervades the speech of the cities, reinforced by code-switching and official communication (forms, notices, interaction in banks and post offices, etc.). The written language may be highly Russianized in political, economic and bureaucratic documentation or journalism, while in other cultural contexts it will display much more Persian vocabulary. At the height of the language reform movement in 1989–90, conservatives complained that unbridled substitution of Persianisms for established Russian loans (*doniš-goh* for *universitet* ‘university’, *havo paymo* for *samolyot* ‘airplane’, etc.) was making the language incomprehensible to the man in the street.

Abbreviations and acronyms are a feature of Soviet Russian bureaucratic language that have readily been adopted (in Russian) and adapted (to Tajik), e.g. *VABK*, for *Viloyat-i avtonom-i Badaxšon-i Kūh-ī* ‘the Gorno-Badakhshan Autonomous Region’.

7.3 Lexical distribution, Persian–Tajik

In the everyday Persian and Perso-Arabic vocabulary, even frequently-used words have been differentially distributed between Iran and Central Asia. Some Tajik words and expressions are either completely foreign to Iranians, or perceived as archaic or literary; the meanings of some common words have shifted considerably. In the following selection of shibboleths, words in parenthesis are Standard Persian equivalents (which sometimes coexist as variants in Tajik): *dina* (*diruz*) ‘yesterday’, *pagoh* (*fardā*) ‘tomorrow’, *be-goh* (*‘asr*) ‘evening; yesterday’, *tira-moh* (*pā‘iz*) ‘autumn’, *daryo* (*rud, rud-xāne*) ‘river’, *bahr* (*daryā*) ‘sea’, *paxta* (*panbe*) ‘cotton’, *tireza* (*panjare*) ‘window’, *xel* (*jur*) ‘sort, kind’, *tayyor* (*hāzer*) ‘ready’, *hozir* (*halā*) ‘now’, *vazn-in* (*sang-in*) ‘heavy’, *sang-in* (*sang-i*) ‘(of) stone’, *kalon* (*bozorg*) ‘big, great, old’, *mayda* (*xōrd*) ‘tiny; small change’, *xurd* (*kuček*) ‘small, little, young’, *kampir* (*pir-e zan*) ‘old woman’, *nayz*, *nek* (*xub, qašang*) ‘good, nice’, *ganda* (*xarāb*) ‘bad’, *kasal* (*mariz, bimār*), ‘sick’, *mond-a* (*xaste*) ‘tired’, *pok* (*tamiz*) ‘clean’, *iflos* (*kasif*) ‘dirty’, *aftid-an* (*oftād-an*) ‘fall’, *šīšt-an* (*nešast-an*) ‘sit’, *mond-an* (*gozāšt-an*) ‘let, put’, *partoft-an* (*andāxt-an*) ‘throw’, *yundošt-an* (*jam kard-an*) ‘gather’.

This applies even to quite recent neologisms: Taj. *madaniyat*, Pers. *farhang* ‘culture’; Taj. *ittifoq*, Pers. *ettehād* ‘(labor) union’. Many of the Tajik shibboleths are also common to Persian of Afghanistan (Kaboli, Dari).

Nominal compounds and conjunct verbs, the productivity of causatives, and adjectival formatives such as *-nok* and *-g-ī* constitute another large area of lexical difference with

Standard Persian. Compound neologisms, even if formed from native lexical stock and not calqued on Russian (for Tajik) or French (for Persian) may not coincide, either lexically or structurally: compare Taj. *bayn-al-xalq-ī*, and Pers. *beyn-al-melal-i* ‘international’, where the difference lies merely in the selection of the central (Arabic) lexeme; Taj. *yaroq-partoi*, Pers. *xal’-e selāh* ‘disarmament’, where the Uzbek-Tajik compound is a modern calque on the Arabicate compound, lit. ‘casting off arms’, still current in Standard Persian; Taj. *zioī*, Pers. *rowšan fekr* ‘(liberal) intellectual’, in which the underlying idea of ‘enlightened thinker’ is expressed in an Arabic-Persian derivative and a Persian-Arabic compound respectively.

7.4 Sociolinguistic aspects

7.4.1 Registers and ta’ārof

A crucial strategy of spoken and written communication is *ta’ārof* ‘formal mutual recognition’. It aims, in addition to its illocutionary function such as making a request, at the perlocutionary effect of self-presentation and respect, inferior and superior, respectively. It is one of the poles of the “communicative contexts in Iranian culture which are marked in terms of two continua: one of personal and communicative intimacy, from ‘inside’ (*bāten*) to ‘outside’ (*zāher*), and another of social hierarchy, from contexts indicating hierarchical relationships to those indicating equality” (Beeman 1988: Abstract). Linguistically, it is marked by a highly conventionalized phraseology and address terminology depending on the context, accompanied by appropriate metalinguistic behavior and gestures, including proper pacing and packaging of objectives, and is applied by speakers from the lowest to the highest speech registers.

Personal pronouns and endings are in the plural (see section 3.1.3.1a), *šomā* VB-*id* ‘you.p are . . .’ for the addressee, *išān* VB-*and* ‘they are . . .’ for a third party, while the speaker may use *mā* Vb-*im*, ‘we are . . .’, reflecting the role of the speaker as member of a social group, rather than as an individual. Lexically most notable is the use of *farmudan* ‘to order, command’ for any superior’s verb of action such as *be-farmā-id* ‘please (DO)’. Typical is the use of light verb constructions for simple verbs: superior raising includes *tašrif* ‘(your) honor’ in *tašrif āvordan* ‘bring honor’ = *āmadan* ‘come’, *tašrif bordan* ‘take honor’ = *raftan* ‘go, leave’, *tašrif dāštan* ‘have, keep honor’ = *budan* ‘be there, present’. Inferior lowering includes (*be*) *xedmat-e* ‘to the service of’ = *be, piš-e* ‘to, to the presence of’, e.g. *xedmat residan* ‘arrive at service’ = *āmadan* ‘come to’; *arz kardan* ‘make presentation’ = *goftan* ‘say, remark’, often in subjunctive, *arz kon-am* ‘let me remark’, and may use an oath-like address such as *qorbān* ‘your honor’ < *qorbān-e šomā be-rav-am* ~ *be-šav-am* ‘(I am ready) to be sacrificed for you’. Superior is addressed as *janāb(-e) āliljanob-i oli* ‘Sir’, ‘gentleman, aristocrat’, lit. ‘high side’, vs. *in jāneblin jonib* ‘I’, lit. ‘this side’, and *sarkār/sarkor* ‘Sir’ or ‘Madam’, lit. ‘leader, head of work’, vs. *bandelbanda* ‘I’, lit. ‘slave’ little used today, and others. All these are contextually selected from a paradigm of options for each, ranging from higher to lower.

Inferior typically introduces requests with one or a series of phrases, or their substitutes, rarely up to seven, and usually in this relative sequence: *janāb āqā-ye X, mā’zerat mi-xāh-am, agar ejāze be-deh-id, xāheš mi-kon-am, agar momken bāš-ad, lotf-an, bi zahmat, be-farmā-id . . .* ‘your excellency Mr. X, I beg your pardon, if you permit me, I ask of you, if it be possible, kindly, (if it is) no trouble, please . . .’

7.4.2 Kinship terms

Persian kinship terms for some basic consanguineal and affinal relationships are single terms, but others are composite descriptive terms expressed by EZ constructions.

In Tajik, alternate or additional terms (some of Uzbek origin) include: *oča* ‘mother’; *ako*, *aka* ‘elder brother’, *dodar*, *uka* ‘(younger) brother’; *apa*, *aya* ‘elder sister’; *yanga* ‘aunt (uncle’s wife), sister-in-law’; *taγoī*~ *amulamak* ‘uncle’, *padar-šuy*, *xusur* ‘father-in-law (of wife)’, *modar-šuy*, *xušdoman* ‘mother-in-law (of wife)’.

TABLE 8.18: PERSIAN KINSHIP TERMS

	Male	Female		
-3	<i>jadd-e bozorg</i>	<i>jadde-ye bozorg</i>	G.-great-grandparents	
-2	<i>jadd</i> Ar.	<i>jadde</i> Ar.	Great-grandparents	
-1	<i>pedar(-e) bozorg</i>	<i>mādar(-e) bozorg</i>	Grandparents	
-1	<i>pedar</i>	<i>mādar</i>	Parents	
-1	<i>amu</i> Ar.	<i>ame</i> Ar.	Paternal uncles, aunts	
-0	<i>dā'i</i> Tk.	<i>xāle</i> Ar.	Maternal uncles, aunts	
+1	<i>pesar</i>	<i>doxtar</i>	Children	
+2	<i>nave (navāde, navāse)</i>		Grandchild	
+3	<i>natije</i>		Great-grandchild	
+4	<i>nabire</i>		G.-G.-grandchild	
+5	<i>na-dide</i>		G.-G.-G.-grandchild	
	Male	Female		
	<i>dāmād</i>	<i>arus</i> Ar.	Bridegroom	Bride
			Son-in-law	Daughter-in-law
	<i>šowhar</i>	<i>zan</i>	Husband	Wife
		<i>havu</i>		Co-wife
	<i>pedar, mādar-e šowhar</i>	<i>zan</i>	Parents-in-law	
	<i>barādar, xāhar-e šowhar</i>	<i>zan</i>	Siblings-in-law	
	<i>zan-e barādar</i>		Brother’s wife	
	<i>šowhar-e</i>	<i>xāhar</i>	Sister’s husband	
	<i>bājnāq</i> Tk.		Husbands of sisters	
		<i>jāri</i>	Wives of husbands	
	<i>zan-e amu, dāi</i>		Uncles, aunts-in-law	
	<i>šowhar-e</i>	<i>ame, xāle</i>	Uncles, aunts-in-law	

7.4.3 Modes of address

Between members of the same peer and solidary groups (age, sex, occupation) address is essentially informal, given names being widely used, often with the endearing suffix *-jānljon* (*Ahmad-jānlAhmad-jon*, *Zohre-jānlZuhra-jon*).

Family members traditionally address one another in kinship terms rather than by their given name, again often adding *-jānljon*: *bābālbobo* ‘grandfather’, Pers. also ‘daddy’, coll. ‘fellow’; *bibilbibī* ‘grandmother’ (also mistress of house, etc.); *dada* ‘father’, Pers. also ‘daddy, mamma’, Pers. *āqā-jān* ‘daddy’. Most of these may be used figuratively to address unrelated acquaintances and strangers of the appropriate age and sex.

During the Soviet period in Tajikistan, Russian and Soviet modes of address became fashionable, including *rafiq* ‘comrade’ preceding the surname and the Russian-style name and patronymic, as *Ahmad Ibrohim-ovič*. Russian-style surnames were also widely adopted (*Yusup-ov*, *Nazir-ova*). Since the 1980s many Tajiks have re-Persianized their

surnames by means of relative or patronymic suffixes such as *-ī* and *-zoda*. In the change from *Yusupov* to *Yusuf-zoda* or *Yusuf-ī* the substitution of *f* for *p* further represents re-Persianization of an Uzbek reflex of the Perso-Arabic *Yusuf*.

8 DIACHRONY AND DIALECTOLOGY

8.1 Phases

The grammatical changes and typological shifts in the evolution of New Persian over fourteen centuries has been substantial, though shifts were mostly gradual and individual rather than bundled together. They involved temporal, social, as well as geographical parameters. Overall, between Middle Persian and contemporary Persian, the decisive monument for the history of Persian was the *Shahname* by the eleventh-century poet Ferdowsi of Tus in Khorasan.

In particular, recent detailed investigations of textual occurrence and relative frequency of features and their typology have led to clearer insights into the historical and geographical dynamics of their loss and innovation. In overview, the following table, based on Paul (2000b), shows the successive phases of innovation and relative consolidation, in correlation with related dynasties:

700–900	Transition from Middle Persian
900–1100	Earliest NP, Ghaznavids; <i>Shahname</i> of Ferdowsi of Tus
1100–1300	ENP, Mongols: in twelfth century: relative dialectal homogeneity and standardization, Persian international language in Eastern Caliphate
1300–1600	“Classical Persian”, Turko-Mongols to Safavids: considerable standard reached; earlier features may occur in emulation of “Classical” style – thereafter slower shifts
1600–1900	Early Modern Persian, Qajars
1900–present	Modern Persian, Pahlavids, Islamic Republic: rapid change in written register, influenced by spoken Tehrani Persian

The complex socio-linguistic dynamics of the development during the early phases of Persian are detailed in Perry (2009).

8.2 Comparative dialectology

The common Early New Persian vowel system had six members: *i ī ē*, *u ū ō*, *a ā*. There occurred shifts in the three main varieties of Persian: In Iranian Persian, the vowel system was restructured by the turn of the seventeenth century: The tense long mid vowels were raised, while the lax short high vowels were lowered: *ē ō*, *i u* > *ī ū*, *e o*: *bē* > *bi* ‘without’, *ō* > *u* ‘he, she’; *ki* > *ke* ‘which, that’, *-u* > *-o* ‘and’. More recently, word-final *-a* was raised to *-e*, but *-a* is still retained before enclitics in the Spoken Modern Standard: *ba* > *be* ‘to’, *hame-y-aš-rā*, but coll. *hama-š-o* ‘all of it’. In Tajik Persian, the restructuring involved the loss of length distinction under Turkic pressure, whereby the inherited short and long high vowels merged, and long *ā* was rounded to *o*. The fewest changes are found in Afghan Persian, where the lax high short vowels were lowered: *i o* > *e o*, and *ā* > *ā̄*. The development of the vowels is shown in the diagram given overleaf (see also Table 8.2).

Tajik		i	e		u	û		a	o
		↑	↑		↑	↑		↑	↑
Early NP		i	ī	ē	u	ū	ō	a	ā
		↓	↓	↓	↓	↓	↓	↓	↓
Afghan		e	ī	ē	o	ū	ō	a	ā
		↓	↓	↓	↓	↓	↓	↓	↓
Iran		e	.ī		o	ū		a	ā

8.3 Innovative constructions, summary

The development in the morphosyntax of these three varieties is the history of increasing differentiation, particularly in their verb systems through the development of new constructions. These include the progressive Aktionsart and the evidential mode. For Afghan Persian see also Farhadi and Perry 2009).

8.3.1 Progressive

Persian uses *dāštan* in its basic meaning ‘keep, hold’, where both auxiliary and dependent verb are finite. Tajik and Afghan use conjunct verb constructions with *istoda* ‘standing’ and *raftan* ‘go’, respectively: Pers. *dār-ad ketāb mi-xān-ad*, Taj. [*kitob-ro xond-a*] *istoda ast*, Afgh. [*ketāb-ra xānd-a*] *mē-rav-ad* ‘he is reading the book’.

8.3.2 Uncertainty constructions

Unlike Persian, Tajik and Afghan have developed modal constructions expressing uncertainty. In Tajik, the conjectural is based on participle II in *-agī* in its gerundive sense. In Afghan, the dubitative is based on the lexical marker *xāt*, a generalized form of 3s *xāh-ad* ‘it will probably be’ < *xāh-* ‘want, will’: [*zad-a*] *xāt bud-om* ‘I might hit’. These contrast with the literary definite future construction with the short infinitive, *xāh-ad raft* ‘he will go’, inherited from Early New Persian (for more detail, see Perry 2002).

8.3.3 Evidentiality

Evidentiality, while found in both Iranian and Afghan Persian, is fully developed only in Tajik, again clearly by interference from Turkic where it is expressed by the single marker *emiš*. Just as the verb forms of Turkic are mostly based on participles, so in Tajik one finds the development of participial formations with so-called converbs, where the participial main verb is followed by a varied set of verbs whose meaning is generalized to express various Aktionsarten.

8.4 Isoglosses West vs. East

8.4.1 Conjunct constructions and complementizer

Morphosyntactically, the innovative progressive and evidential constructions, and the uncertainty constructions in Tajik and Afghan Persian, differ drastically: they are

double-finite construction in Persian, and other western Persian varieties, but are nominalized conjunct verb constructions in Afghan and Tajik Persian. The earliest example of such Eastern conjunct constructions, which was integrated into common Persian, is the passive conjunct construction with *šodan* ‘go’: *ān kušta šud* ‘that one was killed’. These features evidence the membership of Tajik in a vast isoglossic area, to which belong not only Turkic, but also Hindi/Urdu (Windfuhr 2006: 279–281).

8.4.2 Persian homoglossia

In Tajikistan, the linguistic situation is bilingual, and the relationship between the literary standard and local spoken Tajik registers is locally homoglossic rather than diglossic in the sense of two distinct grammatical inventories.

Similarly, in Iran all registers share the same morphological and syntactical features and rules, though high and low differ in complexity. Phonologically, the lower register is characterized by the loss of a number of phonological distinctions which however involve only some 15 merger rules leading predictably from standard to low. The main distinction is lexical and phraseological. There are only the following shibboleths of exclusive use: high *niz* ‘also’ vs. shared *ham* ‘also’; *xiš* ‘oneself, own’ vs. shared *xod* ‘self, own’; low *vāse* ~ *vāsiye* ‘for’ vs. shared *barā-ye* ‘for (the sake of)’. Only the low aspectual particle *hey* ‘continuously’ (reminiscent of ENP *hamē*) does not have a corresponding high pendant.

Perry (2003) has shown that the major coding distinctions between the two registers of Persian amount to a difference of degree rather than kind. Those involve the selection, addition, deletion, reordering, or the re-glossing of elements familiar to both registers, but not exclusive grammatical features. They include:

- (1) Expanded and versatile use of the pronominal enclitics:
 - (a) use of personal enclitics with prepositions: *be-h-eš goft* ‘said to him’;
 - (b) use of 3s enclitic for subject, *raft-eš* ‘he went’;
 - (c) post-verbal position of the enclitic direct object: *bord-am-āš* ‘I took it’.
- (2) The double-finite progressive construction with *dāštan* ‘keep, hold’.
- (3) Violations of SOV: *raft bāzār* ‘he went to the market’, *na-did-am bābā-t-o* ‘I didn’t see your dad’.
- (4) The “idiomatic” use of *ke*.
- (5) The omission of the locative prepositions *be* ‘to’ and *dar* ‘in(to)’.
- (6) The nominal referential suffix *-e*.

However, all of these are to be found in earlier styles of Persian, notably in Classical poetry.

8.4.3 Topical references to Chapter 2

See also the sections in Chapter 2 **Dialectology and Topics** where Persian serves as the example for **the linguistic evolution**: the Iranian verbal quincunx system (section 3.4); **marking of aspect** (section 3.5); noun phrase and syntactic arguments, definiteness (section 4.1), genericity (section 4.1.2), and differential marking of the direct object (section 4.4); and clause complementation (section 4.5).

9 SAMPLE TEXTS

9.1 Persian

The following are the first six sentences of a report on an archaic custom in the region of Khomeyn, Central Lorestan, performed during the ten coldest days of winter (ca. Jan. 25–Feb. 5). A strong person is chosen to go into the Alvand mountains to assure the return of warmth (Anjavi Shirāzi, Seyyed Abolqāsem (1352) *Jashn-hā va ādāb va mo'taqedāt-e zemestāni* [Winter customs and beliefs], Tehrān). He is referred to as *Korde* 'that Kurd', which clearly reflects not only vague memory, but also its ancient connotation of the term *kord*, "mountain people".

Linguistically noteworthy in this passage, particularly in the initial sentences, is the use of the Persian evidential forms, here reflecting both the researcher's second-hand knowledge and the villagers' fading memory (L = linker in relative heads).

Korde be kuh

'(Sending) the Kurd to the mountain(s)'

Be towr-i ke mardom-e mantaqe-ye Rebāt-e Morād-e Xomeyn
the way-L that people-EZ region-EZ Rebāt-EZ Morād-EZ Xomeyn
'As the people of the region of Rebāt-e Morād in the Khomeyn district'

revāyat mi-kon-and,
narrative IPFV-make.PR-3p
'tell it,'

tā yek-sad-o panjāh sāl piš marsum bud-e
till one-hundred & fifty year(s) before customary be.PART.EV
'they had a custom until some 150 years ago'

hame sāle dar ayyām-e "Korde be kuh" yek nafar-rā
all year.ADJ in days-EZ K. one person-DO
'that during the days of K., one person'

ke besyār qavi va nirumand bud-e, entexāb mi-kard-e and.
that very strong & powerful be.PT.EV.3s choice IPFV-make.PART.EV-COP3p
'they would choose who was very strong and powerful'.

Tamām-e ahl-e mahall be u qazā va pušāk mi-dād-e and
whole-EZ people-EZ place to him food & clothing IPFV-give.PART.EV-COP3p
'All people of the village would give him food and clothing'

va u-rā bā tufang va lebās-e garm va āzuqe-ye dah-ruze
& him-DO with gun & dress-EZ warm & provision-EZ ten-days.ADJ
'and him, with gun, warm clothes and a ten-day provision,'

mojahhaz mi-kard-e va be onvān-e "Amu Kord Ali"
supplied IPFV.PART.EV & by title-EZ A.
'they would outfit, and, as "A."'

be kuh-e Alvand mi-ferestād-e and.
to mountain-EZ Alvand IPFV-send.PART.EV-COP3p
'would send him to Mt. Alvand'.

U in dah ruz-rā be kuh mi-raft
 he these ten days-DO to mountain IPFV-go.PT.3s
 'He would go to the mountain for these ten days,'

va bar mi-gašt-e ast.
 & back IPFV-turn.PART.EV COP.3s
 and return.'

Va asr-e ruz-e dah-om marāsem-i-rā
 & afternoon-EZ day-EZ ten-th ceremonies-INDEF-DO
 'And in the late afternoon of the tenth day, certain ceremonies

be xāter-e bar gaštan-e u dar jelo-ve u anjām mi-dād-e and
 for sake-EZ return.INF-EZ him in front-EZ him performance IPFV-
 give.PART.EV-COP3p
 'they would perform in front of him in appreciation of his return,'

va zemn-e rixtan-e ājil va noql va nabāt be sar-e u,
 & meantime-EZ pour.INF-EZ dried fruit & sweets & candy on head-EZ him
 'and, showering him with dried fruit, sweets, and candy,'

esteqbāl-o pazirā'i-ye garm-i az u be amal mi-āvord-e and.
 welcoming & reception-EZ warm-INDEF from him to action IPFV-
 bring.PART.EV-COP3p
 'they would give him an extensive reception and warm welcome.'

agar ān sāl kam-bārān va bad-i mi-šod-e
 if that year little-rain.ADJ & bad-INDEF IPFV-become.PART.EV
 'But if that year turned out to have little rainfall and a bad crop,'

šaxs-e digar-i-rā barā-ye in kār entexāb mi-kard-e and.
 Person-EZ other-INDEF-DO sake-EZ this work choice IPFV-make.PART.EV-
 COP3p
 'they would choose another person for this task.'

Zemn-an mardom se āš ham barā-ye salāmati-ye Amu Kord Ali mi-paz-and.
 meantime people 3 soup also sake-of health-EZ A. IPFV-cook.PR-3s
 'Also, for the sake of the health of A., people would cook three soups (one before his departure, one during his absence, and one on his return).'

Ham-čonin mo'taqed and ke agar āš na-paz-and
 same-thus convinced are-3p that if soup not.SBJ-cook.PR-3p
 'Also, they believe that if they would, lit. 'do', not cook those soups,'

va sag-hā-ye mahalle ham dar ayyām-e, Korde be kuh zuze be-keš-and
 & dog-p-EZ place also in days-EZ K. howl SBJ-continue.PR-3s
 'and also if he dogs of the village howl during the days of K'

va ow'ow kon-and va mardom kaš-hā-ye xod-rā
 & bark SBJ.make-3p & people shoe-p-EZ own-DO
 'and bark, and if (regarding their shoes) people'

pošt-o-ru na-gozār-and, yek-i az bozorg-ān-e ābādi
 back-&-front not.SBJ-place.PR-3p one-INDEF from elder-p-EZ village

mi-mir-ad.

IPFV-die.PR-3s

‘would, lit. ‘do’, not turn them around, one of the elders of the village will die’.

Pas bāyad āš be-paz-and va agar sag-hā ow’ow kard-and,
thus must soup SBJ-cook.PR-3p & if dog-p bark make.PT-3p

‘Therefore they have to cook the soups; and, once the dogs bark. lit. ‘barked’,

kafš-e xod-rā vārune kon-and va pošt-o-ru be-gozār-and
shoe-EZ own-DO turned (SBJ.)make.PR-3p & back-&-front SBJ-place.PR-3p
‘have to turn their shoes upside down and backwards.’

9.2 Tajik

The following is a paragraph from Sadriddin Ayni’s autobiographical novella *Maktabi kūhna* [The old(-fashioned) school], as published in the Ayni anthology *Aknun navbati qalamast* [Now it is the turn of the pen], Dushanbe, 1978, p. 126. Abbreviations and markings in the interlinear gloss:

A verb in upper case translates a non-finite conjunct form; + = connective in adjectival E phrase; CONJ = non-finite conjunct form; EVID/QUOT = evidential use of perfect tense, here two examples with quotative connotation; INT = intentional construction with future participle; BEN = benefactive verb in conjunct construction; PR-PROG = present progressive; SBJ = subjunctive; IMP = imperative.

The *duo* ‘prayer’ referred to in this text is a pious formula written on scraps of paper, which would be immersed in water, where the ink would dissolve, and the water would be drunk as a treatment; the *nazr-i domullo* is a charitable gift, here the fee, given to the *mullo* for his cures.

man az roh-i dur, az Obkena omad-am, guft on zan.
I from way-EZ far, from Obkena come.PT-1s said that woman.
‘I’ve come all the way from Obkena (Ābgina), the woman said.’

ma-ro hamsoya-amon šarofboy,
I-DO neighbor-our Sharofboy
‘My neighbor, Sharofboi (Sharāfbāy),’

ki zan-aš zoyid-a na-tavonist-a
that wife-his DELIVER-CONJ not-been-able-EVID/QUOT
‘his wife can’t deliver her baby’

dard-i saxt kašid-a istod-a-ast,
pain-EZ severe DRAW-CONJ PR-PROG,
‘and is in great pain,’

yak tanga dod-a
one tanga having-given-PART
‘gave me a tanga (“penny”)

baro-yi duo-yi kušoyiš firistod-an-ī šud,
 for prayer-EZ opening about-to-send-INT became.3s,
 ‘and was going to send me to fetch a childbirth prayer.’

zan-on-i hamsoya-g-on-i digar ham
 women-EZ neighbors-EZ other also
 ‘My other neighbors’ wives, too,’

ki har kadom dard-mand bud-a-and,
 that each which suffering they-have-been-EVID/QUOT,
 ‘who all have various aches and pains,’

baro-yi dard-ho-yi xud duo farmud-and
 for-EZ pains-EZ self prayer they-ordered
 ‘asked me to bring back prayers for their aches, too’

va har vaqt pul yob-em
 and any time money we-find-SUBJ
 ‘and as soon as they could afford it,’

nażr-i domullo-ro me-dih-em,
 fee-EZ master-ro we-give,
 ‘they’d send you the money’

burd-a me-dih-ī guft-and.
 TAKE-CONJ you.s-give-BEN they-said.
 ‘through me, they said.’

man ham az baro-yi savob
 I also from sake-EZ (spiritual) reward
 ‘So just to do a good deed’

ba in qadar roh po-yi piyoda xest-a omad-am
 by this much way foot-EZ on-foot RISE-CONJ I-came.
 ‘I’ve come all this way on foot’

šumo duo-ho-ro dihed,
 you.2p prayers-DO give-IMP,
 ‘So, give me the prayers.’

boz har vaqt on-ho pul dihand,
 again any time they money give-3p-SUBJ,
 ‘Then when they give me the money’

ovard-a me-dih-am . . .
 BRING-CONJ I-give-BEN . . .
 ‘I will bring it and give it to you . . .’

kadom-i in duo-ho ba kadom dard ast?
 which-EZ this prayers to which pain is?
 ‘which prayer is for which pain?’

guft-a on zan porsid.
 SAY-PART that woman asked.
 ‘the woman asked.’

BIBLIOGRAPHY AND REFERENCES

- Aghaei, B. (2006) *The Syntax of ke-Clause and Clausal Extraposition in Modern Persian*, Ph.D. dissertation, University of Texas at Austin.
- Ahadi, S. (2001) *Verbergänzungen und zusammengesetzte Verben im Persischen: eine valenztheoretische Analyse*, Wiesbaden: Reichert.
- (2002) *New Persian Language and Linguistics. A Selected Bibliography up to 2001*, Wiesbaden: Reichert.
- Alamolhoda, M. (2000) *Phonostatistics and Phonotactics of the Syllable in Modern Persian*, Helsinki: The Finnish Oriental Society.
- Ansarin, A. A. (2004) 'An acoustic analysis of modern Persian vowels', *SPECOM-2004z*; 315–318—http://www.isca-speech.org/archive/specom_04/spc4_315.pdf
- Bahār, M. T. (1958/1337) *Sabkshenāsi yā tārikh-e tatavvor-e she'r-e fārsi*, Tehrān: Amir Kabir.
- Barakaeva, G. B. (1968) *Luyat-i mukhtasar-i tojikī-anglisī* (Concise Tajik-English dictionary), Dushanbe: Universiteti Davlatii Tojikiston.
- Bashiri, I. (1994) 'Russian loanwords in Persian and Tajiki languages', in Marashi, M. (ed.) *Persian Studies in North America: Studies in Honor of Mohammad Ali Jazayeri*, Bethesda, MD: Iranbooks, pp. 109–139.
- Beeman, W. O. (1986) *Language, Status, and Power in Iran*, Bloomington: Indian University Press.
- (1988) 'Affectivity in Persian language use', *Culture, Medicine and Psychiatry* 12.1: 9–30.
- Birnbaum, S. A. (1950) 'The verb in the Bukharic language of Samarkand', *Archivum Linguisticum* 2.1: 60–73; 2.2: 158–176.
- Bosson, G. (1985) *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*, Tübingen: Narr, 1985.
- Cejpek, J. (1956) 'Die verbale Periphrase als ein wichtiges Unterscheidungsmerkmal zwischen Neupersisch and Tağikisch', *AO* 42.2: 171–182.
- Dehkhodā, A. A. (1994–1995/1372–1373) *Loqatnāme* (Dictionary), Tehrān: Mo'assese-ye Enteshārāt va Chāp-e Dāneshgāh-e Tehrān (some 26,000 pp., also digitalized version).
- Doerfer, G. (1963–1975) *Türkische und mongolische Elemente im Neupersischen. Unter besonderer Berücksichtigung älterer neupersischer Geschichtsquellen, vor allem der Mongolen- und Timuridenzeit*, 4 vols., Wiesbaden: F. Steiner.
- (1967) *Türkische Lehnwörter im Tadschikischen (AKM 37.3)*, Mainz: Harrassowitz.
- (1992) 'Central Asia xiv. Turkish-Iranian language contacts', in *EnIr* 5, pp. 231–233.
- Efimov, V. A. E. (1965) *Iazyk afganskikh khazara. Iakaulangskii dialekt* (The language of the Afghan Hazara. Yakaulang dialect), Moskva: Nauka.
- Eilers, W. (1959) *Deutsch-persisches Wörterbuch*, Wiesbaden: Harrassowitz.
- Estaji, A. and Bubenik, V. (2007) 'On the development of the tense/aspect system in Early New and New Persian', *Diachronica* 24.1: 31–55.
- Farhadi, R. A. G. (1975) *Le persan parlé en Afghanistan. Grammaire du kâboli accompagnée d'un recueil de quatrains populaires de la région de Kâbol*, Klincksieck: Paris (rev. of 1955).
- (1981/1360) *Zabān-e tājiki-ye Māvarā ol-nahr*, Kābul: Akadmi-i Ulum-i J. D. A., Markaz-i Zabānhā va Adabiyāt.
- Farhadi, R. and Perry, J. (2009) 'Kāboli', *EnIr* 15.3, pp. 276–280.
- Folli, R., Harley, H., and Karimi, S. (2005) 'Determinants of event type in Persian complex predicates', *Lingua* 115.10: 1365–1401.
- Gippert, J. (2003) 'Early New Persian as a medium of spreading Islam', in Paul, L. (ed.) *Persian Origins*, pp. 31–47.

- Glassman, E. H. (1972) *Conversational Dari*, Kabul: The Language Orientation Committee, International Afghan Mission (rev. ed.).
- Gomeshi, J. (2003) 'Plural marking, indefiniteness, and the noun phrase', *Studia Linguistica* 57.2: 47–74.
- Griunberg, A. (1963) *Iazyk severnoarzarbaidzhanskikh Tatov* (The language of the North-Azerbaijani Tats), Leningrad: Izd. Akademii Nauk SSSR.
- Griunberg, A. and Davydova, A. (1982) 'Tatskij jazyk', in *Osnovy* 3.1, Moskva, pp. 231–288.
- Haider, H. (1985) 'Chance and necessity in diachronic syntax – Word order typologies and the positions of Modern Persian relative clauses', in Fisiak, J. (ed.) *Papers from the 6th International Conference on Historical Linguistics, Poznan, 22–26 August 1983*, Amsterdam: John Benjamins, pp. 199–216.
- Hayyim, S. (1985) *The Larger English-Persian Dictionary*, Tehran: Farhang-e Mo'āser (numerous impressions).
- Henny, J. L. (1984) 'Enclitics in Pahlavi and Early Classical Persian: A theoretical analysis', in Skalmowski, W. and van Tangerloo, A. (eds) *Middle Iranian Studies. Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*, Leuven: Peeters, pp. 84–94.
- Ioannesian, Iu. A. 1999. *Geratskii dialekt iazyka dari sovremenogo Afganistana* (The Herat dialect of modern Afghan Dari), Moscow: Vostochnaia Literatura RAN.
- Jahangiri, N. (2000) *A sociolinguistic study of Persian in Tehran*, Tokyo: Inst. for the Study of Languages and Cultures of Asia and Africa.
- Jahani, C. (2000a) 'Expressions of indirectivity in spoken Modern Persian', in Johanson, L. and Utas, B. (eds) *Evidentials*, pp. 185–208.
- (2000b) 'Restrictive relative clauses in Classical and Modern New Persian and the marking of the antecedent', *Orientalia Suecana* 49: 33–56.
- (2008) 'Expressions of future in Classical and Modern New Persian', in Karimi, S. et al. (eds) *Aspects of Iranian Linguistics*, pp. 152–175.
- Johanson, L. and Utas, B. (eds) (2000) *Evidentials: Turkic, Iranian and Neighbouring Languages*. Berlin: Mouton de Gruyter (Empirical Approaches to Language Typology 24).
- Kalbāsi, Irān (2001/1380) *Sākhṭ-e eshteqāqi-ye vāzhe dar Fārsi-ye emruz* (The derivational structure of the word in Modern Persian), Tehrān: Pazhuheshgāh-e 'Olum-e Ensāni va Motale' āt-e Farhangi.
- Karimi, S., Samiiān, V. and Stilo, D. (eds) *Aspects of Iranian Linguistics*, Newcastle upon Tyne: Cambridge Scholars Publishing.
- Karimov, H. et al. (1970) *Zabon-i adabi-i hozira-i tojik: sintaksis*, Dushanbe: Nashriēt-i Irfon.
- Key, G. (2008) 'Differential object marking in a Medieval Persian text', in Karimi et al. (eds) *Aspects of Iranian Linguistics*, pp. 227–247.
- Khānlari, P. N. (1988/1366) *Tārikh-e zabān-i Fārsī*, Tehrān: Enteshārāt-e Now (fourth printing).
- Kieffer, C. M. (1983) 'Afghanistan v. Languages', *EnIr* 1.5, pp. 501–516.
- (1989) 'L'avancée du persan vers l'Est: le cas des isolats persanophones les plus orientaux en Afghanistan,' *StIr* 18.2: 221–236 (vernaculars of Laghmān, Sorkhrōd, Kandībāgh, Ūrgūn).
- (2004) 'Hazāra iv. Hazāragi dialect,' *EnIr* 12, pp. 90–93.
- Koutlaki, S. A. (2002) 'Offers and expressions of thanks as face enhancing acts: *tæ'ārof* in Persian', *Journal of Pragmatics* 34.12: 1733–1756.
- Lazard, G. (1956) 'Caractères distinctifs de la langue Tadjik', *BSL* 52: 117–186.
- (1963) *La langue des plus anciens monuments de la prose persan*, Paris: Klincksiek.

- (1968) 'La dialectologie du judéo-persan', *Studies in Bibliography and Booklore* (Cincinnati) 8: 77–98.
- (1970) 'Études quantitative de l'évolution d'un morphème: la postposition *rā* en persan', in Cohen, D. (ed.) *Mélanges Marcel Cohen*, The Hague: Mouton, pp. 381–388.
- (1982) 'Le morphème *rā* en Persan et les relations actancielles', *BSL* 77.1: 197–208.
- (1985) 'L'inférentiel ou passé distancié en persan', *StIr* 14, 1985, pp. 27–42.
- (1990a) *Dictionnaire persan-français*, Leiden, New York: E. J. Brill.
- (1990b) 'Lumières nouvelle sur la formations de la langue persane: un traduction du Coran en persan dialectal et ses affinités avec le judéo-persan', in Shaked, S. and Netzer, A. (eds) *Irano-Judaica II, Studies Relating to Jewish Contacts with Persian Culture Throughout the Ages*, Jerusalem: Ben-Zvi Institute, pp. 184–198.
- (1992) *Grammar of Contemporary Persian*, Costa Mesa, Calif: Mazda Publishers (tr. of *Grammaire du persan contemporain*, Paris: Klincksieck, 1963).
- (1993) *The Origins of Literary Persian, The Annual Norūz Lecture by Distinguished Scholars of Iranian Studies*, Bethesda, Maryland: The Foundation for Iranian Studies.
- (1995) *La formation de la langue persane*, Paris: Peeters.
- (1996) 'Darī', *EnIr* 7.
- (1999) 'Pahlavi, Pārsi, Dari', in *EnIr* 9, pp. 158–162.
- (2002) 'Le médiatif: considérations théoriques et application à l'iranien', in Johanson, L. and Utas, B. (eds), pp. 209–28.
- (2002) 'La formation du persan sous l'influence de la langue arabe au IV/X siècle', *CSI* 26: 11–19.
- (2003) 'Du pehlevi au persan: diachronie ou diatopie?', in Paul, L. (ed.) *Persian Origins*, pp. 95–102.
- Lorentz, M. (1964) 'Die partizipialen Wendungen – ein Unterscheidungsmerkmal des Tağikischen gegenüber dem Persischen', *Mitteilungen des Instituts für Orientforschung* 10: 133–139.
- Mahootian, S. (1997) *Persian. A Descriptive Grammar*, London: Routledge.
- Matras, Y. (2002) 'Kurmanji complementation: Semantic-typological aspects in an areal perspective', *STUF* 55: 49–63.
- Meunier, A. and Samvelian, P. (1997) *La postposition rā en persan. Ses liens avec la détermination et sa fonction discursive, Cahiers de Grammaire 22*, Toulouse: Université de Toulouse–Le Mirail.
- Meyer-Ingversen, J. (1974) 'Fekr. Ein Beitrag zur Transformationsgrammatik des Neupersischen', in Frye, R. (ed.) (1974) *Neue Methodologie in der Iranistik*, Wiesbaden: Harrassowitz, pp. 131–282.
- Nercissians, E. (2001) 'Bilingualism and diglossia: Patterns of language use by ethnic minorities in Tehran', *International Journal of the Sociology of Language* (Berlin) 148: 59–70.
- (1985/1364) 'Dozabānegi dar Irān' (Bilingualism in Iran), *Majalle-ye Zabān-šenāsi* 2.2: 73–80.
- Öhl, P. and Korn, A. (2008) 'Performanzbasierte und parametrische Wandel in der linken Satzperipherie des Persischen. Der Subordinationsmarker *ke* und die Interrogativpartikel *āyā*', *Die Sprache* 46: 137–202.
- Oranskij, I. M. (1975) *Die Neuiranischen Sprachen der Sowietunion*, translated by W. Winter, The Hague, Paris: Mouton & Co.
- Paul, L. (1999) 'Neuere Tendenzen in der Entwicklung des Persischen am Beispiel der Teheraner Freitagspredigten', *StIr* 28.2: 275–290.
- (2000a) 'The language of the *Shahnameh* in historical and dialectal perspective', ILEX Foundation, *Papers Presented at the Second Ferdowsi Conference*,

- Center for the Great Islamic Encyclopedia, Tehran, Iran, August 21–28, 2000*, http://www.ilexfoundation.org/pdf/ferdowsi_2/ludwigpaul300.pdf
- (2000b) ‘A linguist’s fresh view on “Classical Persian”’, *CSI* 26: 21–34.
- (ed.) (2003a) *Persian Origins – Early Judaeo-Persian and the Emergence of New Persian. Collected Papers of the Symposium, Göttingen 1999*, Wiesbaden: Harrassowitz.
- (2003b) ‘Early Judaeo-Persian in a historical perspective: The case of the prepositions *be*, *u*, *pa(d)*, and the suffix *rā*’, in Paul, L. (ed.) *Persian Origins*, pp. 177–194.
- (2008) ‘Some remarks on the Persian suffix *-rā* as a general and historical issue’, in Karimi, S. *et al.* (eds) *Aspects of Iranian Linguistics*, pp. 328–337.
- Perry, J. R. (1985) ‘Language reform in Turkey and Iran’, *International Journal of Middle East Studies* 17.3: 295–311; reprinted with corrections in Atabaki, T. and Zürcher, E. J. (eds) *Men of Order: Authoritarian Modernization in Turkey and Iran*, London, New York: I. B. Tauris, pp. 238–259.
- (1989) ‘Uzbek influence on Tajik syntax: The converb constructions’, in Clyne, P. R., Hanks, W. F., Hofbauer, C. L. (eds) *The Elements: A Parasession on Linguistic Units and Levels. Papers from the Conference on Non-Slavic Languages of the USSR*, Chicago: Chicago Linguistic Society, The University of Chicago, pp. 448–461.
- (1991) *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending*, Costa Mesa, Calif.: Mazda Publishers.
- (1996) ‘From Persian to Tajik to Persian: Culture, politics and law reshape a Central Asian language’, in Aronson, H. I. (ed.) *NSL.8. Linguistic Studies in the Non-Slavic Languages of the Commonwealth of Independent States and the Baltic Republics*, Chicago: Chicago Linguistic Society, The University of Chicago, pp. 279–305.
- (1997) ‘Script and scripture: The three alphabets of Tajik Persian, 1927–1997’, *Journal of Central Asian Studies* 2.1: 2–18.
- (2000) ‘Epistemic verb forms in Persian of Iran, Afghanistan and Tajikistan’, in Johanson, L. and Utas, B. (eds) *Evidentials*, pp. 229–257.
- (2002) ‘Grammaticalization in process: The evolution of the Afghan Persian adverb *xāt*’, in Cavoto, F. (ed.) *A Linguist’s Linguist: A collection of Papers in Honour of Alexis Manaster Ramer*, München: Lincom Europa, II, pp. 317–27.
- (2003) ‘Persian as a homoglossic language’, *CSI* 27: 11–28.
- (2005) *A Tajik Persian Reference Grammar*, Leiden-Boston: Brill.
- (2007) ‘Persian’, in Kaye, A. S. (ed.) *Morphologies of Asia and Africa*, 2 vols., Winona Lake, Ind.: Eisenbrauns, vol. 2, pp. 975–1019.
- (2009) ‘The origin and development of Literary Persian’, in Bruijn, J. T. P. (ed.) *A History of Persian Literature, Vol. I. General Introduction to Persian Literature*, London, New York: I. B. Tauris, pp. 43–70.
- Pisowicz, A. (1985) *Origins of the New and Middle Persian Phonological Systems*. Kraków: Nakładem Uniwersytetu Jagiellońskiego.
- (2003) ‘The possible influence of Kurdish dialects on colloquial Persian: the definiteness suffix *-é*’, *FO* 39: 177–180.
- Rajā’i, A. Alī (1964/1343). *Yāddāshti dar bāre-ye lahje-ye Bokhārā’i* (French title page: *Remarques sur le dialecte de Bukhara*), Mashhad: Chāpkhāne-ye Dāneshgāh-e Mashhad.
- Rastorgueva, V. S. (1954) ‘Kratkii ocherk grammatiki tadzhikskogo iazyka’, in Eadem (ed.) *Tadzhiksko–russkii slovar*, Moskva: Gos. Izd-vo inostranykh i natsional’nykh slovarei, pp. 529–570.
- (1963) *A Short Sketch of Tajik Grammar* [= *Kratkii ocherk grammatiki tadzhikskogo iazyka*], Paper, H. H. (transl. & ed.), Bloomington, Indiana and The Hague: Research Institute for Inner Asian Studies, Indiana University.

- (1964) *Opyt sravnitel'nogo izucheniia tadzhikskikh govorov*, Moskva: Nauka.
- Rastorgueva, V. S. and Kerimova, A. A. (1964) *Sistema tadzhikskogo glagola*, Moskva: Nauka.
- Rakhimi, M. Kh. and Uspenskaia, L. Kh. (1954) (eds) *Tadzhiksko-russkii slovar'*, Moskva: Gos. Izd-vo inostranykh i natsional'nykh slovarei.
- Rzehak, L. (2001) *Vom Persischen zum Tadschikischen. Sprachliches Handeln und Sprachplanung in Transoxanien zwischen Tradition, Moderne und Sowjetmacht (1900–1956)*, Wiesbaden: Reichert.
- Sadat-Tehrani, N. (2005) 'Indifference *ke*-construction in Modern conversational Persian', in *Proceedings of the 2004 annual conference of the Canadian Linguistic Association*, <http://http-server.carleton.ca/~mojunker/ACL-CLA/pdf/Sadat-Tehrani-CLA-2004.pdf>
- Sāfi, Qāsem (2003/1982) *Bahār-e adab: Tārikh-e mokhtasar-e zabān va adabiyāt-e Fārsi dar shebħqāre-ye Hend va Pākestān*, Tehrān: Mo'assase-ye Enteshārāt va Chāp-e Dāneshgāh-e Tehrān.
- Samvelian, P. (2008) 'The Ezafe as a head-marking inflectional affix: Evidence from Persian and Kurmanji Kurdish', in Karimi *et al.* (eds) *Aspects of Iranian Linguistics*, pp. 339–361.
- Soper, J. (1996) *Loan Syntax in Turkic and Iranian: The Verb System of Tajik, Uzbek and Qashqay*, Bodrogligeti, A. J. E. (rev. and ed.), Bloomington, Indiana: Eurologua.
- Steingass, F. J. (1970) *A comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature*, Beirut: Librairie du Liban (first impression 1892).
- Utas, B. (1991) 'New Persian as an interethnic medium', in Svanberg, I. (ed.) *Ethnicity, Minorities and Cultural Encounters*, Uppsala: Centre for Multiethnic Research, pp. 103–111.
- (2006) 'A multiethnic origin of Persian', in Johanson, L. and Bulut, C. (eds) *Turkic-Iranian Contact Areas*, Wiesbaden: Harrassowitz, pp. 241–251.
- Vivier-Muresan, A. (2006) 'Le code de politesse iranien (*ta'ārof*) ou la fiction du lien social', *L'Homme* 180, pp. 115–138.
- Windfuhr, G. L. (1979) *Persian Grammar: History and State of its Study*, The Hague, Paris and New York: Mouton.
- (1980) 'Verbal deictic matrices in Persian', paper given at the 190th Meeting of the American Oriental Society, San Francisco, April.
- (1982) 'The verbal category of inference in Persian', *Acta Iranica* 22: 263–287.
- (1986) 'A spatial model for tense, aspect and mood', *Folia Linguistica* 19.3: 415–461 (verb systems of Persian, French, Russian, Arabic, Turkish, Creole).
- (1987) 'Persian', in Comrie, B. (ed.) *The World's Major Languages*, London: Croom Helm, pp. 523–546 (rev. eds. 2009).
- (1987) 'Convergence: Iranian Talyshi and Turkic Azari', in Bashir, E., Deshpande, M. M., and Hook, P. E. (eds) *Select Papers from Sala-7, South Asian Languages Roundtable Conference, Ann Arbor, Michigan, May 1985*, Indiana, Bloomington: University Linguistics Club, pp. 385–405.
- (1997) 'Persian phonology', in Kaye, A. S. (ed.) *Phonologies of Asia and Africa*, Winona Lake, Ind.: Eisenbrauns, pp. 663–677.
- (2004) 'The Iranian languages: Typology in space and time', Historical Linguistics Colloquium, Department of Linguistics, University of Michigan, two lectures, March 2004, unpublished ms.
- (2006) 'Language change and modeling modal axes: Irano-Turkic convergence', in Johanson, L. and Bulut, C. (eds) *Turkic-Iranian Contact Areas. Historical and Linguistic Aspects*, Wiesbaden: Harrassowitz, pp. 252–283.