KABUL PERSIAN VERB MORPHOLOGY[1]

MICHAEL M. T. HENDERSON

UNIVERSITY OF KANSAS

The morphological structure of the Modern Persian verb can be described as a six-slot string of constituents, with each slot rewritten as a pair of features or as a phonological matrix. Some of the features, all of which are necessary for semantic and syntactic rules, are not interpreted as phonological matrices, and emerge as phonetically null. Inflectional affixes are spelled directly as phonological matrices. Regular morphophonemic alternations in verb stems are described by general rules which operate on phonological and/or morphological features. Irregular alternations are described by minor rules for which formatives must be marked, their unmarked values being "minus;" the unmarked value of general rules is, of course, "plus." Minor rules describe suppletive as well as unique but non-suppletive ones; verbs which undergo both minor and general rules are so marked.

0. THE DIVISION OF THE LEXICON of Modern Persian into morphological subclasses or "parts of speech" is rather fuzzy. Verbs, however, are clearly marked by the inflections peculiar to them, and by certain morphophonemic alternations exhibited nowhere else in the language. There are remarkably few verbs in Persian: excluding derived causatives, only 234 verbs appear in Haïm's *Shorter Persian-English Dictionary* (Tehran, 1963). In Tehran Persian (TP) new verbs are added to the *î*-class (see section 2.3); in Kabul Persian (KP), the variety of which this paper is a partial description, new verbs usually take the form of a substantive plus an empty verb *kardán* 'to do' or, for intransitives, *šudán* 'to become.' Thus in TP 'puncture' would be *pančarîdán* (transitive or intransitive), while in KP it would be *pančar kardán* 'to puncture' vs. *pančar šudán* 'to be punctured.'

KP differs from TP in several other ways, chiefly in the conservatism of its vowel system. The symbols that I use for the phonemes and the distinctive features that 'spell' them appear in Table (1) below.

(1)	p	b	f	w	M	t	d	s	z	l	r	n	č	ǰ	š	y	k	g	q	x	ɣ	'	î	i	e	a	â	o	u	û
Cons	+	+	+	–	+	+	+	+	+	+	+	+	+	+	+	–	+	+	+	+	+	–	–	–	–	–	–	–	–	–
Syll	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–		+	+	+	+	+	+	+	+
Son	–	–	–	+	+	–	–	–	–	+	+	+	–	–	–	+	–	–	–	–	–	–	+	+	+	+	+	+	+	+
Nasal	–	–	–	–	+	–	–	–	–	–	–	+	–	–	–	–	–	–	–	–	–		–	–	–	–	–	–	–	–
High	–	–	–	+	–	–	–	–	–	–	–	–	+	+	+	+	+	+	–	–	–		+	+	–	–	–	–	+	+
Low	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–		–	–	–	+	+	–	–	–
Back	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+		–	–	–	–	+	+	+	+
PLACE[2]	L	L	L		L	C	C	C	C	C	C	C	A	A	A		V	V	U	U	U									
Round	–	–	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–		–	–	–	–	+	+	+	+
Voice	–	+	–	+	+	–	+	–	+	+	+	+	–	+	–	+	–	+	–	–	+	–								
Tense																							+	–	+	–	+	+	–	+
Cont	–	–	+	+	+	–	–	+	+	+	+	+	–	–	+	+	–	–	–	+	+	–								

In Table 2, blanks represent inapplicable features rather than redundancies. Reference will often be made below to three styles of speech: Formal (F), Deliberate (D), and Colloquial (C). These

[1] An earlier version of this article appeared in the *Journal of the American Oriental Society 98.4 (1978),* pp 375-388.

[2] In the earlier version of this paper, referred to in fn. 4, the point of articulation of vowels and semivowels was described with the features [±coronal] and [±anterior], necessitating the use of [-minus] to mean different things, depending on the major class of the segment described. Using PLACE with unitary features avoids this messy mode and phonetically meaningless mode of description.

three styles, which are part of the competence of any educated native speaker of KP, are integral parts of a description of the phonology and morphology of the language.[3]

The most exhaustive study of Persian verb stem alternations is that of Hodge 1958, in which six classes (and many subclasses) of verbs are distinguished, ranging from invariant to suppletive stems. While Hodge correctly observes that this classification can best be made on the basis of the present stems, he also states that "there is no predictability in the absolute sense." This is rather extreme: it is absolutely predictable, for instance, that no Persian verb stem will be found whose past stem ends in a voiced obstruent, and that no alternation between a vowel and an obstruent will be found. In general, though, the regularities are few, and many can be contradicted by one or two exceptions.

The first section of this paper is a description of the morphological structure of the KP verb, and the second is an attempt to account for stem alternations, with suggestions for the treatment of suppletions.

1.0 Structure of the verb

The complex Old Iranian tense system has been almost completely replaced by periphrastic constructions (Darmesteter 1883). What is of interest in this morphological study is the six-slot constituent of those constructions which contains the verb stem and is separated from other constituents of the verb phrase by word boundaries. It is arranged as follows:

(2) (NEG) ASPECT + STEM + (PART) AFFIX

A causative morpheme whose phonological shape is *ân* is inserted before TENSE in some verbs and after it in others. This is unpredictable, and verbs must be marked for the placement of this infix (Farhâdi). The causative is of little morphophonemic interest, and will not be discussed further here. The remainder of this section is a discussion of the six slots shown in (2) above.

1.1 NEGATIVE

The first slot is "NEG," which has the phonemic shape *n* if the next segment begins with a low vowel in style C, and *na* everywhere else including before low vowels in styles D and F. F and D *náâmad* 'didn't come' is *nâmad* in C, but *nabud* 'wasn't' is the same in all three styles, as is *náistâd* 'didn't stand.' The underlying form of NEG, which will be 'spelt' by a rule triggered by the syntactic-semantic feature [+negative], can be represented as |na|, the |a| being deleted by rule DKP (7) which deletes one |a| next to another low vowel in just the way described here.[4]

1.2 ASPECT

The two overt aspect markers in KP are the nonperfective and the optative (called the 'imperative' in many, if not most, grammars, but its use in KP includes nonimperative as well as imperative meanings). Rather than saying that the perfective aspect, which is not represented phonologically, is realized as zero, it seems better to include in the grammar rules such as (2):

---

[3] For a description of the phonological correlates of these three styles, see Henderson 1975. In the text, I refer to this article as DKP, and to the various rules in it as, e.g., DKP(2), meaning the second rule given in that article. For a discussion of inapplicable features, see Henderson 1976.

[4] I cite underlying forms between vertical bars and surface phonemic forms in italics. When necessary, phonetic differences are shown using IPA symbols between square brackets [ ].

(2)

$$\text{ASPECT} \rightarrow \left\{ \begin{array}{l} [\pm \text{ perfective}] \\ [\pm \text{ optative}] \end{array} \right\}$$

The brace notation indicates that the two aspects are mutually exclusive. Spelling rules are triggered by the features as follows:

(3)

(a) [–perfective] → |mE|
(b) [+optative] → |bV|

Since [+perfective] and [–optative] are not mentioned in (3), they have no phonetic manifestation. See Harris 1969 for a discussion of this approach to verb morphology, which allows syntactic features to operate in the ways required by the syntactic features while allowing the phonological component's rules simply not to mention the spelling of non-overt categories. The cumbersome use of zeroes surrounded by formative boundaries is thus avoided.

It is interesting to note that nonpast forms rarely occur without the nonperfective marker. Persian grammars list forms such as *bâfam* 'I weave,' but remark that forms such as *mebâfam* 'I am weaving' have assumed perfective as well as imperfective meaning, and that the latter have replaced the former almost everywhere. (Lambton 1957:25).

The nonperfective marker |mE| contains a vowel which appears as *i* before stems beginning with two consonants, and *e* elsewhere, as in *mífrušam* 'I sell' vs. *mésazam* 'I build.' This alternation cannot be predicted on the basis of stress, since the marker is always stressed; the underlying representation must contain a vowel unspecified for the feature [high], which is then specified by the rule that spells the formative. (The feature [tense] is redundantly "plus" in KP mid vowels.) (3a), then, should be revised as follows:

(4)

$$[\text{–perfective}] \rightarrow m \quad \left[\begin{array}{l} \text{V} \\ \alpha\text{high} \\ \text{–low} \\ \text{–back} \\ \text{+stress} \end{array} \right] \quad /\_\_\_ + \alpha\langle C \rangle C\ V_0C_0$$

That is, if the verb stem begins with two consonants, the vowel in the imperfective marker is [+high] *i*; otherwise it is [–high] *e*.

The optative marker |bV| appears as *i* before stems in which the first vowel is [–back], and *u* before stems in which the first vowel is [+back]. Note that unlike the epenthetic vowel described by rule (2) in DKP, this vowel is not conditioned by the roundness of the next vowel. Given the stem |štâb| 'to hurry,' the epenthetic vowel will be *i*, as in *šitâftam* 'I hurried', since |â| is [–round] at the underlying level; but the vowel of the optative marker will be *u*, since |â| is [+back], producing *buštâbim*, 'we should hurry.' (3) must then be revised as follows:

(5) [+optative] → *b* $\begin{bmatrix} \text{V} \\ +\text{high} \\ \alpha\text{back} \end{bmatrix}$ /___+ C^2 $\begin{bmatrix} \text{V} \\ \alpha\text{high} \end{bmatrix}$

1.3 STEM

Nonsuppletive verb stems are listed in the lexicon as a single underlying form.[5] The surface allomorph is conditioned by (a) the presence or absence of the past marker and (b) the general phonological rules of the language. The stems of Persian verbs and their alternations are discussed in detail in the second section of this paper.

1.4 TENSE

Persian verbs are either past or nonpast in form. A rule for this will look like the following:

(6) TENSE → [±past]

A further rule spells [+past] as |T| (what *t* and *d* have in common): *t* appears after obstruents and *d* elsewhere:

(7) [+past] → $\begin{bmatrix} \text{T} \\ \alpha\text{voice} \end{bmatrix}$ / [αsonorant]]Verb + ___

1.5 PART(ICIPLE)

Persian verbs have two participles, past and present. It seems best to view "PART" as one morpheme whose form is determined by the presence or absence of a Tense morpheme. The participle used to be *ta*, added directly to the verb stem (Darmesteter, p.219), but in a synchronic analysis there seems to be no reason but to regard it as |a| added to |T|. The present participle is |ân|, added directly to the verb stem, since there is no [–past] morph. The rule will look like this:

(8) PART → $\left\{ \begin{array}{lll} \text{a} & & [+\text{past}] \\ & / & \\ \text{ân} & & [-\text{past}] \end{array} \right.$ /___ $\left. \right\}$

The present participle is of limited interest in this study: first, it has no affixes added to it, and so plays an inactive role in the morphophonology of the verb. Secondly, its use seems to be quite restricted, and it may be a frozen form in combination with only a few stems, in which case the lower line of rule (8) can be deleted. Farhâdi puts it thus (p.108):

---

[5] It is not clear from evidence found in child language acquisition that nonsuppletive stems are actually stored in the brain in a single form; it is for **descriptive** economy that linguists regard them so, recognizing that suppletive forms must be acquired separately, as in English *go/went*.

> Parfois la part. Prés. Fait, en tant qu'adjectif adverbial, partie du vocabularie de la langue, bien que celle-et n'ait pas gardé le verbe correspondante.

The past participle, on the other hand, is productive, affecting the affixes added to it, as will be shown in the following section.

1.6 AFFIX

The inflectional affixes of Persian verbs, representing three persons and two numbers, vary according to the style of speech as well as according to the formatives to which they are attached. In this section the F and D variants will be discussed first, followed by the C style.

1.6.1 In nonpast forms, the affixes are attached directly to the verb stems; in past forms, the affixes are attached to |T|; and in (past) participial forms, they are added to |Tá|.[6] The following paradigm shows the F variants, attached to the stem |bâf| 'to weave':

(9)

	Present	Past	Participial
	1 *mébâfam*	*bâftam*	*bâftá'am*
Singular	2 *mébâfti*	*bâfti*	*bâftá'id*
	3 *mébâfad*	*bâft*	*bâftá'ast*
	1 *mébâfim*	*bâftim*	*bâftá'im*
Plural	2 *mébâfted*	*bâfted*	*bâftá'ed*
	3 *mébâfand*	*bâftand*	*bâftá'and*

The forms in (9) show that all of the suffixes except the third singular are the same no matter what they are attached to; but the third singular is *ad* after the stem, null after |T|, and *ast* after |Ta|. The lack of a third person singular affix in past forms is the result of a Middle Persian process by which affixes were attached to original participles in all but the third singular (Darmesteter, p. 223). The Pahlavi (Middle Persian) paradigm of the past tense of 'to weave' had *waft* for all persons; later, affixes were added to all but the third singular forms in the continuing change from "ergative" to "nominative" agent forms (Geiger 1893:1-5). Similarly, the participial forms were (and are) made up of the participle and the reduced form of the copula: the latter is identical to the other affixes in all persons except the third singular, in which it is the same as the full form, *ast* (Darmesteter, p. 231). It seems from these facts that no single base form can be posited for the various third person singular affixes, but that the two overt forms should be spelt directly, as follows:

(10)

$$[3\ \text{Pers}] \rightarrow \begin{array}{ll} |\text{ad}| & / [-\text{past}] + \left[\overline{-\text{plural}}\right] \\ |\text{ast}| & / [+\text{past}] + \text{PART} + \left[\overline{-\text{plural}}\right] \\ |\text{and} & / \left[\overline{-\text{plural}}\right] \end{array}$$

Since the environment [+past] with no "PART" after it is not mentioned in (10), no rules will apply to it, and there will be no phonological realization of this category.

Now underlying forms can be posited for the affixes:

[6] Stress is marked on the past participial form because, unlike in all other verb forms, it attracts the otherwise initial stress, as if for all the world it had turned into an adjective: *mékufam* 'I chop' vs. *kuftá kabâb* 'chopped kabob.'

(11)

	Sing	Pl
	1 *am*	*îm*
	2 *î*	*ed*
	3 *ad, ast*	*and*

Farhâdi (p. 78) lists forms identical to these as the classical phonemic shapes of the affixes. I interpret this as further support for the analysis given here in which, of course, the affixes are not present in the lexicon but are spelled by transformational or "spelling" rules such as those given above.

1.6.2 The C(olloquial) variants of the affixes are different from the F and D variants. The differences are shown here:

(12)

	F, D	C
	1 *am*	*um, em*
Sing	2 *î(d)*	*î*
	3 *ad, ast*	*a*
	1 *îm*	*îm*
Plural	2 *ed*	*en*
	3 *and*	*an*

The third person forms, singular as well as plural, could perhaps be derived from the non-Colloquial varieties by an extension of DKP (13), which deletes for word-final coronal stops. The second person plural affix is *ed* in F and D. Farhâdi suggests (p.78) that the C form *en* may be derived from an earlier *-end*, a formative analogous to *-and*. In any case, this affix must be derived either from |ed| or from a different underlying form. Similarly, the C first person singular affixes, *um* after non-participial forms and *em* after participial forms, must be derived from non-Colloquial |am| or from some other underlying affix. Farhâdi says (p.79) that *em* is a combination of the participial *a* and the affix *–am.* I doubt that this is correct, because it would be the only instance in the language of two |a|'s coalescing to *e* rather than to *â*, but Farhâdi's suggestion (p. 13) that "la langue évite ainsi les confusions semantiques" should perhaps not be dismissed out of hand.

The positing of different underlying forms for stylistic variants was rejected in DKP (p.653) on the grounds that the variant forms can be heard uttered by the same person in close juxtaposition, so I conclude that the derivation of C forms from the same forms that underlie the F and D variants is the correct analysis, and that somewhere in the grammar rules of the following type occur:

(13)

F, D		C
(a) *am*	→	*um, em*
(b) *ed*	→	*en*
(c) *and*	→	*an*
(d) *ad, ast*	→	*a*

The participial vowel must be deleted in C; this rule must be different from DKP (7), because the participial vowel is stressed and therefore escapes said rule. This rule, which must refer to morphological features, looks like this:

(14)

PART + AFFIX			∅	[+stress]
1	2	⇒	1	2

2.0 STEM ALTERNATIONS

Persian verbs typically exhibit alternations in the last one or two segments of their stems. Reference has already been made to the most thorough treatment of Persian verb stem alternations to date, Hodge (1958). The usual treatment is to divide the verbs into several classes: the "regular" class consists of invariant stems; the other classes consist of alternating and suppletive stems. The "regular" verbs are a closed class of twenty-three members.

Whether or not Persian speakers have convenient base forms stored in their brains is, of course, undiscoverable. While it is quite possible that they have base forms for some verbs and sets of memorized allomorphs for others (cf. Stanley 1971), it is impossible to tell from their performance which are which, just as it is impossible to tell whether an English speaker has memorized *write* and *wrote* separately or derives both from a single underlying form. But a descriptive linguist, in today's models, sets up base forms that reveal as many of the phonological and morphological regularities of the language because—so the argument goes—a simple listing of allomorphs treats regular alternations no differently from suppletive ones, which are demonstrably more difficult for learners (children or speakers of other languages—cf. Chafe). This may be an overstatement. It is, of course, possible to reveal regularities simply and elegantly in an item-and-arrangement framework without abstract underlying forms, as is obvious in Hodge's treatment of Persian verb stems. In keeping with the spirit of the model in which this description is presented, however, I have posited single underlying forms for verb stems whose alternations can be described by rules that make true generalizations about the language.

2.1 CONSONANTAL ALTERNATIONS

In this section, I discuss those KP verbs whose consonantal alternations are regular enough to be described by general rules, and those whose alternations can be described by minor rules that affect more than one verb. The table in (15) contains examples of some of the verbs under discussion; each example is followed by a number in parentheses indicating the total number of verbs exhibiting that particular alternation. Since there are only 234 verbs in KP,[7] 166 of which (including the 134 invariants) exhibit no consonantal alternations, a class with five members is proportionately large. The past stems are those which appear before |T|; the present stems appear before the present participle |ân| or the affixes, all of which begin with vowels. In general, past stems can be derived from present stems more easily[8] than the reverse.

(15)

Present	Past	Gloss	Number
raw	*raf*	'go'	(2)
yâb	*yâf*	'find'	(5)
afrâz	*afrâx*	'elevate'	(2)
sâz	*sâx*	'build'	(18)
šnâs	*šinâx*	'know'	(1)
furoš	*furox*	'sell'	(1)
xâh	*xâs*	'want'	(4)
dâr	*dâš*	'possess'	(9)

[7] As noted above, KP brings new verbs into the language by putting *kardan* 'to do/make' or *šudan* 'to become' after a substantive, while TP adds *–idan* directly to the stem of the substantive, essentially adding more verbs to the lexicon.

[8] For the analyst, even if not for the native speaker.

In the past stems, only the voiceless spirants *f s š z* occur stem-finally. In the present stems, *b w s z š r h* appear. Table (16) summarizes these alternations:

(16)

Present (/___V)	Past (/___T)
b, w	*f*
h	*s*
s, z, r	*š*
s, z, š	*x*

Ignoring the present stem-final *r* for the moment, DKP (9) showed that surface *w* could be derived from several different underlying segments, including |b| and |f|. If |v| is the final segment in the form underlying *raw – raf* 'go', all of the stem-final consonants except *r* are obstruents. Rule 7 above spells the past marker |T| as *t* after obstruents and *d* after sonorants. After rule 7 has applied, the obstruent cluster assimilation/dissimilation rule DKP (8) devoices and spirantizes |b v| to *f* and |z s| to *s*.[9] More changes are necessary, of course, because most of the obstruents in (16) change not only their manner but also their point of articulation; this will be dealt with directly. Rule DKP (8), which does not apply in F, must be annotated so that it applied to verb stem-final obstruents in all speech styles.

In some cases, present stem-final *s z š* must be velarized to *x*, and in others *s* and *z* are palatalized to *š*. The first of these alternations (which is more general than the second by a ratio of 20:3) is the relic of old stems in final *č*: intervocallically, the *č* was voiced to *ǰ* and then depalatalized to *z* (Darmesteter, p. 203); in the past forms, the *č* was velarized to *x* (ibid. p. 203). In a synchronic analysis, nothing appears to be gained from positing *č* as the final segment of these verbs. Accordingly, the present stems in final *s z š* are taken as the base forms, with a rule given to palatalize them in certain grammatical/phonological environments:

(17)

$$s\ z\ š \rightarrow x \ / \ V\_\_\_+ |T|$$

("V___+ |T|" stands, in these rules, for the environment in which these alternations are found: the end of a verb stem plus the past marker. The latter will be spelled *t* or *d* as the circumstances warrant.) In the case of verbs in *z*, the output of (17) will be ɣ, devoiced to *x* by rule DKP (8).

The three verbs in final *z* whose past stems end in *š* must not undergo (17). They can escape it by means of a lexical feature [–rule 17], which makes them more marked than the stems that do undergo (17); see section 2.4. Having escaped (17), they undergo a special rule:

(18)

$$\text{stem-final } s\ z\ š \rightarrow \begin{bmatrix} +\text{high} \\ -\text{anterior} \end{bmatrix} \ / \ V\_\_\_\_\_+|T|$$

[9] DKP (8) states that in a cluster of two obstruents, the first must agree in voicing with the second; furthermore, if the second is a stop, the first must be or become a spirant. This is an ancient rule in the language, having first entered Iranian more than one and a half millennia ago. See H. Hübschmann, *Persische Studien* (Strassburg: Trübner, 1895), pp. 114-115.

The input to (18) must be a continuant because stem-final *d* does not become *š*. It becomes *s*, regularly, by rule DKP (8).

It is now necessary to decide what to do about stems in *r – š.* The former is a sonorant, after which |T| appears as *d* after sonorants, just like eight other sonorants whose past stems end in *r*. The nine verbs in question here, however, have past stem-final *š* before |T|. They are descendants of verbs in final *rš* (Darmesteter, pp. 83. 208-209); it seems best, however, to consider their base forms to end in |r| in a synchronic analysis, since native speakers acquiring Persian have no access to Middle Persian. It is somehow necessary to "obstruentize" |r| before rule DKP (8) spells |T| as *t* in order to derive the correct forms of |T|.

One obvious tactic is to assign the feature [-sonorant] to just those |r|'s that turn into *š* before |T| and [+sonorant] before all other |r|'s in the language. But aside from the nine verb stems exhibiting this kind of "absolute neutralization" (Kiparsky 1968), there is not a shred of evidence for it. Positing a special underlying segment after doing its dirty work in the tiny number of forms that it was designed to "explain" is, to my mind, the sort of hocus-pocus that generative phonologists ought to be trying to avoid as they become more concerned with phonetic reality (in fact, in some models of phonology, absolute neutralization is completely disallowed (Kiparsky, *ibid.*).

It seems to me that the best solution to this problem is to call an exception an exception, and to have these nine verb stems be marked as having a special rule which simply changes *r* to š before |T|. In this way, the entire verb stem, rather than a single segment, is marked as the exception that it truly is, something native speakers acquire singly, as English speakers do with such forms as [rijd] ~ [rɛd], [sɪt] ~ [sæt], and [gow] ~ [wɛnt]. A different way of preventing the obstruentization rule from applying to all the other underlying |r|'s will be discussed in section **2.4**. While it would it would somehow be more elegant to have the rules that spell grammatical formatives apply before this type of surface rule, it is apparently impossible in this and other cases. Before rule DKP (8) spells |T|, then, the nine stems in *r* ~ *š* must undergo the following rule:

(19)

$$r \rightarrow \check{s} \,/\, V\_\_\_\_ + |T|$$

|h| is the final segment of four stems whose past variants end in *s*. The original stem of these four verbs in *h* ~ *s* ended in *d* (Darmesteter, pp 208-209), which spirantized to *s* before *ta* and was replaced by *h* intervocalically, as in *hvâdâmi* > *xwâham* 'I want.' (Darmesteter, p.71). The base forms should thus end in |h|, which assimilates to a dental before |T|−the reverse of the historical change of PIA **s* to Old Iranian *h* by rule 20:

(20) $h \rightarrow s \,/\, V\_\_\_\_ + |T|$

A single verb with a past stem in *š* has a present stem in *rd:* g*ašt* 'wandered'~ g*ard* 'wanders.' This verb, cognate to Sanskrit *vartati* and Latin *vertō,* is the sole survivor of a group which analogically developed a past stem in *š* and fell together, except for the curious present stems, with the other verbs in *r* ~ *š* (Darmesteter, p. 84.) Rule (20) should thus be preceded by a rule simplifying morpheme-final triconsonantal clusters:

(21) $C \rightarrow \varnothing \,/\, C\_\_\_\_ + C$

While this rule accounts for only one verb, it is more general than it looks: it also accounts for the derivation of D and C *havdá* '17' and *haždá* '18' from |haft| '7' and |hašt| + *dah* '10,' so the rule can be stated in phonological terms, with no reference to morphological boundaries beyond the obligatory one between the stem and *dah*.

Nasal consonants are also involved in verb stem alternations. Specifically, present stems in final *mb, nd,* and *n* alternate with past stems in *f, s,* and ∅: that is, these verbs must appear to DKP (8) as if the nasal were not there; and since |T| is *t* after those that end in obstruents, and *d* after

those which do not (e.g. vowels and other sonorants) before the nasal, which is deleted. The simplest way to derive these forms is to have them undergo this deletion, after which the other changes necessary are achieved by the application of Rule DKP (8):

(22) $\begin{bmatrix} +\text{consonantal} \\ +\text{nasal} \end{bmatrix} \rightarrow \varnothing \,/\, \text{V}\_\_\_(\text{C}) + \text{T}$

Rule (22) must apply before (21).

All of the rules discussed so far have operated on base forms which are identical to present stems. As noted above, this is the way most Persian verb stem alternations work, although there are some stems, to be discussed below, whose alternations are more easily described if their past forms, not the present forms, are considered the base forms. A list of such verbs is given in the Appendix, in which it can be seen that the base forms are identical except for differences common to all the allomorphs of the stem.

2.2 VOCALIC ALTERNATIONS

Seventeen verbs whose present stems end in a low vowel sometimes followed by *r*, have past stems in final *â(r)*. Considering the base forms and present stems to be the same, we can state a rule to describe the information as follows:

(23) $\begin{bmatrix} \text{V} \\ +\text{low} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{high} \\ +\text{back} \end{bmatrix} \quad / \text{V}\_\_\_\_(r)+|\text{T}| \qquad \text{C}_0$

This alternation is the relic of an ancient causative, according to Darmesteter (pp. 194-195). The seventeen verbs that exhibit it form an unproductive class which needs to be marked with a feature [+ rule 23].

Two more rules are needed to describe the non-unique vocalic alternations of Dari[10] verb stems. Other vocalic alternations are either unique or are combined with other strange alternations, and will be described in section **2.4**. Between a present stem-final vowel and an affix, *y* is inserted, as in *mešuyad* ‘washes’ from |mE+šu+ad|. The rule for this is as follows:

(24)

$\varnothing \rightarrow y \qquad /\text{V}\_\_\_\_+ \text{C}_0$

This rule applies in environments beyond the verb system so it needs no morphological features. When a nominalizing suffix *i* is added to a stem ending in a vowel, *y* must be inserted before it, as in |bâlâ| ‘high’ + | → *bâlâyi* ‘height’ (Farhâdi, p. 101). In F forms, it changes to *ʔ*.

A present stem-final |h| in D and C changes to *y* in past stems. This is the only instance of a glottal becoming a glide between two low vowels (cf. rule DKP (3)). The variation is shown in F *meǰahad,* C, D *meǰayad* ‘he jumps.’ The rule can be stated as follows: with the morphologically environment necessarily included, since it does not apply elsewhere to |h| between low vowels:

[10] The terms ‘Dari’ and ‘Kabul Persian’ are interchangeable.

(25)

$h \rightarrow y$ / V___+ AFFIX

2.3 AUGMENTATIVE ALTERNATIONS

109 of the 234 verbs of KP have past stems that are one or two segments longer than their present stems. 100 of them are in the semi-productive *i*-class shared with TP as described in section **0** above. KP seems to prefer the clearer periphrastic Substantive + *kardan* (trans.) and *šudan* (intrans). Among the 100 verbs following the TP pattern, wherein the present stem + *i* replaces the older past tense, following (or overlapping with) the Substantive + *i* TP pattern. Examples are *raqs+i+dan* shows the substantive *raqs* 'a dance' followed by the empty *i* which precedes the *d* needed to separate the *i* from the infinitive marker *an*.[11] In addition to this class, nine other verbs have past stems with final *d, ma, uf, is,* and *es* with present stems in which these extra pieces are missing.

An easy way to deal with these nine verbs is to posit a special boundary between what is retained of the stem in both tenses (and the infinitive) and what is deleted in nonpast forms. This boundary could be symbolized as |=|; anything after it is deleted in nonpast forms. While the need for boundaries of different strengths in phonological representations is well documented in phonology (cf. Stanley), it is really too easy to posit a special boundary = whenever it is convenient to do so. The *i*-augment acts like a morpheme, even in those cases where it carries no morphological, syntactic, or syntactic load, and there is nothing to prevent the analyst from positing a special boundary; but the other "extra pieces" are in no sense morphemes, though their historical antecedents are often clear enough.[12] Positing a boundary whose sole effect is to delete segments after it in nine instances is equivalent to claiming that the rule is somehow phonologically motivated. I do not want to make such a claim: I find it doubtful in the extreme, for instance, that a KP speaker's knowledge that the *if* that appears after *gir-* in *giriftam* 'I took' does not appear in *megiram* 'I take' is knowledge about his/her phonology. There is no phonological motivation for these alternations: they do not take place because some linguist-invented boundary is there—the boundary is there *in the description* because the alternations appear in the language. I find it extremely unlikely that, except in the case of the semi-productive *i*-augment verbs, KP speakers have a single base form "stored" in their lexica for these few odd verbs; they probably learn them as practically suppletive allomorphs, as noted in fn. 11, above.

Still, the analyst's job is to describe observable phenomena, not as yet undiscoverable synaptic connections. If a rule is needed to account for 100 items and that rule can be made *simpler* by generalizing it to account for nine more forms, one is duty bound to state the rule as generally as possible. In the case of the augmentative verb stems, the rule needs to be stated, as far as possible, in morphological terms, to avoid any implication that a phonetic or phonological cause is behind the overt phonological effect. The rule, then, can look like this:

(26)

$(V^1)(C^1)(V^1) \rightarrow \varnothing$ / STEM = _____ + [–past]

---

[11] It might be possible to regard *–idan* simply as one morpheme which is attached to a substantive to make the infinitival form of the corresponding verb; however, the *–an* infinitive marker has to be detached again to give a place to put the inflections shown in (12) above.

[12] See the verbs listed in section 4.1 of the appendix, and the discussion of the same verbs in Darmesteter (1883) and Paul Horn (1893).

When a substantive like |raqs| 'dance' or |xušk| 'dry' is denominalized, a transformation simply adds |=i| to it. The other augmentative verb stems simply have the phonologically null |=| boundary in their base forms.

"Augment" is, of course, a poor name for a phonological chunk from the end of a verb stem, but it must suffice for now, despite the implication that it has something to do with the Indo-European augment, which appears to be found nowadays only in Yaghnobi.

2.4 EXCEPTIONAL VERBS

There remain for consideration sixteen verbs whose alternations are unique. Three of them are suppletive (cited with the past stem first): *bû/ast* 'be'; *dî/bîn* 'see'; and *kar/kun* 'do'. The high frequency with which these appear is probably the reason for their retention of alternations lost in other verbs.

The thirteen other peculiar verbs show partial similarity in their stems, but have other alternations which are irregular: *pux – paz* 'cook', for example, has the regular *x – z* alternation, but the different vowels is like that accounted for by rule (23) above, and is unlike any other *x – z* verbs, such as *sâx – sâz* 'build'. Another, *dâ/dih* 'give', of which the C present stem is just *t,* shares no pattern with any other verb in the language.

How are these alternations best to be accounted for? On the one hand, it is clear that whenever possible, rules already present in the grammar to account for regularly-occurring phenomena should do the work of accounting for alternations like those exhibited by *pux – paz*; on the other hand, if a special rule must be posited to account for a unique alternation, why should it not be used (with, perhaps, a little tinkering) to describe the whole alternation? In addition, why should unique alternations be described by rules at all? If the verb stems exhibiting these oddities are acquired by native (and non-native) speakers as suppletive forms, as they no doubt are, perhaps both stems should be listed in the lexicon, as they are in dictionaries, and the matter pursued no further.

I think that the second question can be answered by saying that while a generative grammar is written as a list of rules, this does not imply that everything the grammar described is rule-governed behavior. Chafe (p. 117) recognizes this:

> At the moment I know of no way to determine precisely how many rules of a synchronic set represent a speaker's tacit knowledge—how deeply he has internally reconstructed in forming generalizations about his language. If individuals vary, no such determination may be possible. Perhaps a linguist can do no better than to posit the system which a maximally efficient rule user would operate with, all the while realizing that no human being is maximally efficient.

Thus a decision to describe both regular and irregular phenomena by the same rule does not entail a suggestion that all—or even any—KP speakers have internalized these rules.

Similarly, it seems unwarranted to exclude irregular or unique verbs from rules which are needed anyway to describe regular phenomena. Where there is a regularity, it should be called a regularity, and should not be penalized for appearing next to an irregularity. The problem is the same as the oft-cited one of how best to say that the past of *sit* is *sat.*

Having made the philosophical decision to describe the maximum by rules, I must suggest how this can be done. Suppose that there are two types of rules in the grammar: general rules and minor rules. General rules apply to all morphemes or strings that meet their structural description; minor rules apply to only those morphemes or strings that are marked "plus" for those minor rules. This information can be conveyed by means of marking conventions of the kind proposed in Chapter Nine of Chomsky & Halle 1968:

(27)

(a) [U rules $G_1 \ldots G_x$] → [+rules $G_1 \ldots G_x$]

(b) [U rules $M_1 \ldots M_x$] → [−rules $M_1 \ldots M_x$]

where $G_1 \ldots G_x$ is the set of general rules and $M_1 \ldots M_x$ is the set of minor rules.

Morphemes which are unmarked for all rules undergo all general rules and no minor rules morphemes marked for a general rule do not undergo it; and morphemes which are marked for a minor rule do undergo that minor rule. Only *M'*s are counted for purposes of complexity, and the exceptionality of a morpheme is measured by the number of *M's* associated with it. (This is, I suspect, one of the better ways of representing the greater number of rules needed to describe the derivation of Deliberate and Colloquial forms from Formal ones—see the end of section 1.6.2 above).

For example, consider rule (19) above, by which |r| is spirantized and obstruentized. Only nine verbs are affected by this rule; the thirteen other verbs ending in *r*, and all the other *rt rd* sequences escape it. Rule (19) is, therefore, a "minor" rule, and the unmarked value of its application is "minus;" the nine verbs referred to are [M Rule 19]. Similarly, the eighteen verbs whose stems contain an *a* or an *â* which is changed to *û* in the past are [M rule 23], the vowel-changing rule, but they are [U rule 19], because when they end in *r* they do not undergo that rule.

The three verbs in stem final *s z* whose past stems do not undergo the general rule (18) are [M rule 18], since their structural description says they should undergo it but do not, and [M rule 19] because they should not but do undergo it. A verb like *pux – paz* is [M rule 23], the rule that changes the vowel, but [U rule 18], because the *x* → *z* rule is general.

I won't state here the host of quite uninteresting minor rules necessary to account for the sixteen exceptional verbs, since from an examination of the allomorphs it is obvious in all cases what the rule must be. Aside from the special rules needed for these verbs, we see that twelve general rules and three minor rules represent the entire descriptive apparatus needed, when it is added to the devices already necessary in the language, to describe the Persian verb system as it is used in Kabul, Afghanistan.

APPENDIX

This appendix is a list of 134 verbs which, together with the *î*-augment class, form the corpus of this study. The verbs are all of those, excluding synchronically derived causatives, in Haïm's *Shorter Persian-English Dictionary* (Tehran: Y. Beroukhim and Sons, 1963). Each verb is cited with the past stem first, followed by the present (if it is different); next comes the posited underlying form(s), and the infinitive (the traditional citation form) in Dari script.[13] This last is included for the convenience of readers who wish to look up the verbs in other sources. The order of the forms within the various subsections follows that of the Perso-Arabic alphabet. Notes on rule features appear at the end of the subsections listing verbs which undergo minor rules or do not undergo general rules.[14]

1. Invariants

âgan	|âgan|	'fill'	آگـندن
âwur	|âwur|	'bring'	آوردن
afsar	|afsar|	'freeze'	افسردن
afšân	|afšân|	'scatter'	افشن
afkan	|afkan|	'throw'	افـکـن

[13] With apologies for any misspellings.

[14] In the interest of simplicity, stress is not marked in the forms cited in the Appendix. All verbs receive stress on the last syllable of the stem except when a stressed Aspect marker is present. When the infinitive is cited, the stress is on the last syllable, *-an*.

bâf	|bâf|	'weave'	بافتن
parur	|parur|	'foster'	پروردن
xwân	|xwân|	'read'[15]	خواندن
xor	|xor|	'eat'	خوردن
rân	|rân|	'drive'	راندن
res	|res|	'spin'	ریستن
sitâ	|stâ|	'take'	ستادن
sitar	|star|	'shave'	ستردن
suf	|suf|	'pierce'	سفتن
šân	|šân|	'show'	شاندن
šikaf	|škaf|	'open'	شکفتن
šikâf	|škâf|	'split'[16]	شکافتن
kuš	|kuš|	'kill'	کشتن
gusâr	|gusâr|	'absorb'	گساردن
gastar	|gastar|	'spread'	گسترن
lumbân	|lumbân|	'gum'	لمبان
mân	|mân|	'stay'	مائدن

All of the above verb stems with final *n* are [-rule 22]. |rês| 'spin' and |kuš| are [-rule 17].

2. Consonantal alternations

2.1 Labials

âšuf	~	âšub	|âšub|	'be disturbed'	آشفتن
tâf	~	tâb	|tâb|	'shine'	تافتن
raf	~	raw	|raw|	'go'	رفتن
šif	~	šiw	|šiw|	'enamour'	شیفن
kuf	~	kub	|kub|	'chop, grind'	کفتن
yâf	~	yâb	|yâb|	'find'	یافتن
firêf	~	firêb	|firêb|	'deceive'	فریفتن

2.2 Dentals and palatals

2.2.1 x – z

âx	~	âz	|âz|	'draw'	آختن
âmox	~	âmoz	|âmoz|	'learn'	آموختن
âmex	~	âmez	|âmez|	'mix'	آمیختن
âwex	~	âwez	|âwez|	'hang'	آوختن
afrox	~	afroz	|afroz|	'kindle'	افروختن
andâx	~	andâz	|andâz|	'throw'	انداختن
andûx	~	andûz	|andûz|	'accumulate'	اندوختن
angex	~	angez	|angez|	'provoke'	انگختن
pardâx	~	pardâz	|pardâz|	'pay'	پردختان
parhex	~	parhez	|parhez|	'obtain'	پرهختن
tâx	~	tâz	|tâz|	'rush'	تاختن

[15] In D and C, the *w* is omitted.

[16] It is possible that *šikâf* 'split' is the causative form of *šikaf,* 'break.' The usual Persian method of forming causatives is to insert *–ân–* at the end of the stem. Thus, *bâft-* 'weave' might become *bâftân-* 'cause to weave.' I believe that this alternation is, if not productive, at least easily understood by Dari speakers. See also *guzar* and *guzâr* in 2.4.

dox	~	doz	|doz|	'sew'	دوختن
rex	~	rez	|rez|	'pour'	ر یختن
sâx	~	sâz	|sâz|	'build'	ساختن
sipox	~	sipoz	|spoz|	'pierce'	سپوختن
sûx	~	sûx	|sûx|	'burn'	سوختن
godâx	~	godâz	|godâz|	'melt'	ڴداختن
nawâx	~	nawâz	|nawâx|	'play'	نواختن

2.2.2 x ~ s š

šinâx	~	šinâz	|šnâx|	'know'	شئاختن
furox	~	furoz	|frox|	'sell'	فروختن

2.2.3 š ~ s z

âyâx	~	âyâz	|âyâz|	'macerate'	آپاختن
afrâš	~	afrâz	|afrâz|	'elevate'	افراشتن
sirîš	~	sirîs	|srîz|	'knead'	سریشتن
reš	~	res	|res|	'spin'	ریشن
niwîš	~	nwïš	|nwiš|	'write'[17]	ئوشتن

2.2.4 š – r(d)[18]

ambâš	~	ambâr	|ambâr|	'store'	انباشتن
angâš	~	angâr	|angâr|	'suppose'	انگاشتن
pandâš	~	pandâr	|pandâr|	'suppose'	پنداشتن
dâš	~	dâr	|dâr|	'possess'	داشتن
kâš	~	kâr	|kâr|	'plant'	کاشتن
gaš	~	gard	|gard|	'wander'	ڴشتن
guzaš	~	guzar	|guzar|	'pass'	ڴزشتن
guzâš	~	guzâr	|guzâr|	'permit'	ڴزاشتن
gumâš	~	gumâr	|gumâr|	'appoint'	ڴماشتن
nigâš	~	nigâr	|nigâr|	'decorate, write'	نگاشتن

2.2.5 s – h

ǰas	~	ǰah	|ǰah|	'jump'	جستن
xâs	~	xâh	|x(w)ân|	'want'	خواستن
ras	~	rah	|rah|	'be saved'	رستن
kâs	~	kâh	|kâh|	'decrease'	کاستن

2.3 Nasals

bas	~	band	|band|	'tie'	بستن
paywas	~	paywand	|paiwand|	'join'	پیوستن
čî	~	čîn	|čîn|	'pluck'	چیندن
rîd	~	rîn	|rîn|	'defecate'	ریندن
zâd	~	zân	|zân|	'give birth'	زاندن

[17] All of the verbs in 2.2.3 are ⊢Rule 17 and [+rule 18].

[18] All of the stems in 2.2.4 are [+Rule 20].

zad	~	zan	|zan|	'hit'	زدن
suf	~	sumb	|sumb|	'bore/drill'	سفتن
guzî	~	guzîn	|guzîn|	'choose'	گزیدن
nihuf	~	nihumb	|nihumb|	'hide'	نیهوفتن

3. Vocalic alternations

âzmû	~	âzmâ	|âzmâ|	'test'	ازمودن
âsû	~	âsâ	|âsâ|	'rest'	آزمودن
âlû	~	âlâ	|âlâ|	'contaminate'	آلودن
afrû	~	afrâ	|afrâ|	'increase'	افرودن
afšûr	~	afšâr	|afšâr|	'press'	فشارد
andû	~	andâ	|andâ|	'coat'	اندودن
bur	~	bar	|bar|	'carry'	بردن
paymû	~	paymâ	|paimâ|	'measure'	پیمدن
sarû	~	sarâ	|sarâ|	'sing'	سرودن
sû	~	sâ	|sâ|	'rub'	سودن
sipur	~	sipar	|spar|	'entrust'	سپردن
šumûr	~	šumâr	|šumâr|	'count'	شموردن
farmû	~	farmâ	|farmâ|	'command'	فرمودن
fišur	~	fišâr	|fšâr|	'press'	فشموردن
namû	~	namâ	|namâ|	'show'	نمودن

All stems in **3** are [+rule 23]

4. Augmentative alternations: since the *î*-augment class is the productive one in Modern Iranian Persian, it would serve no purpose to list here the 100 examples appearing in Haïm's dictionary. As noted in section **0**, these may comprise a closed (list) class, the new verbs being coined by the addition in KP of the intransitive /šudan/ or transitive /kardan/ to a substantive, avoiding the semantic confusion inherent in the Iranian model.

xandî ~ xand		|xand|	'laugh'	خندیدن

4.1 Other augmentative stems

ârâ	~	ârâs	|ara=s|	'decorate'	آراستن
âmâ	~	â	|â=mâ|	'come'	آمادن
istâ	~	ist	|ist=â|	'stand'	استادن
aftâ	~	aft	|aft=â|	'fall'	افتادن
bâyîs	~	bây	|bâi=is|	'be necessary'	بایستن
pazîruf	~	pazîr	|pazîr=uf|	'accept'	پزیروفتن
tawânis	~	tawân	|tawân=is|	'be able'	توانیتن
ǰus	~	ǰu	|ǰu-s|	'boil'	چوستن
dânis	~	dân	|dân=is|	'know'	دانستن
rus	~	ru	|ru=s|	'grow'	رستن
zîs	~	zî	|zî=s|	'live'	زیستن
šus	~	šu	|šu=s|	'wash'	شستن
firistâ	~	firist	|firist=â|	'send'	فرستاندن
girif	~	gir	|gir=if|	'take'	گرفتن
mânis	~	mân	|mân=is|	'resemble'	مانستن
nigaris	~	nigar	|nigar=is|	'regard'	نگرستن
nihuf	~	nih	|ni=huf|	'wear'	نهفتن
yâris	~	yâr	|yâr=is|	'brave'	یارستن

5. Exceptional/suppletive verbs (listed in the lexicon—and learned—separately)

bû	~	ast/bâš		‘be’	بودن	است/باشد
pux	~	paz	-|paz|	‘cook’		پختن
xuf	~	xâb	-|xâb|	‘sleep’		خابیدن
xâs	~	xez	-|xez|	‘rise’		خاستن
dâ	~	dih/t		‘give’		دادن
dî	~	bîn		‘see’		دیدن
rubû	~	rabâ	-|rabâ|	‘seize’		ربودن
zudû	~	zidâ	-|zidâ|	‘rub off’		زودودن
surî	~	sur	-|sur=î|	‘slide’		سریدن
šu	~	šaw	-|šav|	‘become’		شدن
šikas	~	škan	-|škan|	‘break’		شکستن
šinî	~	šnaw	-|šnav|	‘be acquainted with’		شنیدن
kar	~	kun		‘do, make’		کردن
gurex	~	gruz	-|gruz|	‘flee’		گرختن
gusex	~	gusel	-|gusel|	‘break off’		گسختن
mûr	~	mîr	-|mîr|	‘die’		موردن
nišas	~	nišîn	-|niš|	‘sit’		نشستن
aš	~	hil		‘put’		اشتن

Stems marked with a hyphen are [+rule 25]; some of these stems, and all of the others, need minor rules to account for their unique alternations. No base form has been posited for the four suppletive stems, since one stem can be derived from the other by means of an ad hoc rule.

REFERENCES

Chafe, Wallace. 1967. The Ordering of Phonological Rules, *International Journal of American Linguistics* 34.115-136; see also Harris, p. 65.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Farhâdi, A.G. n.d. *Le persan parlé en Afghanistan: Grammaire du Kabòlì.* Paris: Klincksieck.

Geiger, Wm. 1893, Die Passivconstruction des Präteritums transitiver Verba im Iranischen, *Festgruß an Rudolf von Roth* , pp. 1-5.

Haïm’s *Shorter Persian-English Dictionary.* 1963. Tehran: Y. Beroukhim and Sons.

Harms, Robert T. 1968. *Introduction to Phonological Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Harris, James, 1969. *Spanish Phonology.* Cambridge, MA: MIT Press.

Henderson, Michael M. T. 1975. Diglossia in Kabul Persian Phonology. *JAOS* 95:651-4.

_____. 1976. Redundancy, Markedness, and Simultaneous Constraints in Phonology, *Language* 52.314-25.

Hodge, Carleton. 1958. Review of *Grammaire du persan contemporain,* by Gilbert Lazard. *Language* 33- 119.

Horn, Paul. 1893. *Neupersische Etymologie* . Strassburg: Trübner.

Hübschmann, H. 1895. *Persische Studien.* Strassburg: Trübner. 114-115.

Jazayery, M.A. 1964. Review of Rastorgueva. *Language* 41:341-2, 19

Kiparsky, Paul, 1968. *How Abstract is Phonology?* Bloomington: Indiana University Linguistics Club.

Lambton, Ann K.S. 1957. *Persian Grammar.* Cambridge: Cambridge University Press.

Stanley, Richard. 1971. *Boundaries in Phonology*. Bloomington: Indiana University Linguistics Club.