

KABUL PERSIAN VERB STEMS REVISITED

MICHAEL M. T. HENDERSON

UNIVERSITY OF KANSAS

1. Some time ago (Henderson 1978, 2009) I described the verb morphology of Modern Persian as a six-slot string of constituents, each slot rewritten as a pair of features or as a phonological matrix. The second section of that article listed the five large classes of verb stems, divided according to the differences between the present and the past stems:

- (1) invariants, such as *māndan* ‘stay’, *xordan* ‘eat’ (23 verbs).¹
- (2) consonantal alternations, such as *bast/band* ‘tie’ (70 verbs).
- (3) vocalic alternations, such as *burd/bard* ‘carry’ (15 verbs)
- (4) augmentative stems, in which the past stem is one or two segments longer than the present stem, such as *dānist/dān* ‘know’ or *nihuft/nih* ‘wear’ (18 verbs, plus the productive *î*-augment class).
- (5) exceptional or suppletive verbs, for which no responsible assertion can be made that the alternations are rule-governed, such as *bū/bāš/ast* ‘be’ and *dīd/bīn* ‘see’ (18 verbs).

Other members of class (5) may share one or more alternations with members of another class, such as *zudūd/zidā* ‘rub off’ which shares a process with *āsūd/āsā* ‘rest,’ but has a unique vowel alternation earlier in the stem.

A linguist writing rules to describe linguistic phenomena such as the above is trying to account for the knowledge that native speakers of the language have in their minds, and to do so in a manner consistent with a plausible view of a child’s ability to acquire its language. It is a plausible assumption, for instance, that confronted with a reasonably large set of examples of a particular phenomenon, a child will construct a rule which can be applied to as-yet-unheard forms. We know that this happens because we have plentiful examples of children’s overgeneralizations: English-speaking children apply the regular past tense formation rule (add *d*, devoiced to *t* after voiceless consonants, and insert *ə* first if the stem ends in *t* or *d*) to verbs with irregular pasts. The English-speaking world is full of children producing forms like **goed*, **comed*, and **sended*. Errors like these tell us that our rules are in fact descriptive of rule-governed behavior. But we also write rules, such as the ten I listed in my other publications on Modern Persian verbs, which resemble the kinds of rules a child might construct to account for a large corpus of data, but which in fact account for only one or two forms. These rules are just like the kind we could write for English to account for the many ‘strong’ verb alternations that have remained in the language, such as $ə \rightarrow ey$ for the *come/came* alternation. It looks like a rule, and it could be written with distinctive features, but a **rule** is commonly accepted to mean “a description of a regularly occurring phenomenon.” Just as it takes at least three aligned trees to make a row, it certainly takes more than one or even two examples to be called a rule-governed occurrence.

In the earlier study (p. 384) I said that “a generative grammar is written in terms of rules, but...this does not imply that everything the grammar describes is rule-governed behavior.” Nevertheless, the vexing question remains of how to produce a description that adequately gives the facts, while at the same time plausibly reflecting the kind of knowledge a Persian-acquiring child might have, which my earlier studies (while whining somewhat about it) did not.² In the time since the publication of that article a new approach to the description of morphophonology has been developed, one which helps to deal with the welter of historical relics that make up the modern Persian verb system. It is to this newer model that I now turn.

¹ The verbs are those found, as in Henderson 1978, in Haïm 1963).

² Henderson 1978 did, I still believe, give a very plausible account of the alternations in the morphemes other than the verb stems; even decades later, I see no need to revise the rules which “spell” the grammatical formatives.

2. LEXICAL PHONOLOGY. It has long been assumed that while the regularly occurring, predictable parts of languages should be described by rule, the unpredictable parts belong in lists. These hold the items that a learner must memorize, consciously or subconsciously, since they are not conducive to rule formation. The list that linguists assume exists in every speaker's mind is the Lexicon. This is where all the morphemes go, at least those with any semantic content,³ and where the particular characteristics of each morpheme are listed: its form, grammatical category, semantic features such as the animacy or inanimacy of a noun, or that a verb's object be animate, as is the case for *frighten*, and any other unpredictable facts about the morpheme. Beyond these facts, it is now considered wise to put some *processes* in the lexicon, processes that look for all the world like phonological rules, such as the General and Minor rules posited in my earlier studies, but that instead of there being two types of rules, there are two types of rule *application*. Applications of rules are distinguished by the domains in which they operate (Folarin 1987:15ff). A rule applies either inside the lexicon, or after it, or sometimes both. Here are some examples of each of these types of rule application in English:

- (1) Lexical application only: addition of *-y* to stems such as *president*, *current*, or *dormant* causes the *t* to be pronounced *s* (and written *c*) as in *presidency*, etc. That this is a lexical phenomenon and not a phonological one is shown by the fact that adding the phonologically identical but morphologically different /iy/ to other words ending in *nt* does not produce the consonant change: *mint* + *y* does not become **mincy*, any more than *aunt* + *cy* becomes **auncie*.
- (2) Post-lexical application only: purely allophonic rules, such as English aspiration of voiceless stops initially in stressed syllables, or nasalization of vowels before nasal consonants; and assimilation rules such as the realization of the plural morpheme as *-d* after voiced sounds and *-t* after voiceless ones. Since these are exceptionless, predictable, and phonetically motivated, they belong not in the lexicon but in the post-lexical part of the grammar.

(3) APPLICATION TO PERSIAN VERB STEMS

A venerable rule of consonant assimilation and dissimilation in Persian states that in a cluster of two obstruents, the first must agree in voicing with the second; moreover, if the second is a stop, the first must be or become a fricative:

$$[-\text{sonorant}] \rightarrow \left[\begin{array}{c} \alpha \text{ voice} \\ <+\text{continuant}> \end{array} \right] / ___ \left[\begin{array}{c} [-\text{sonorant}] \\ \alpha \text{ voice} \\ <-\text{continuant}> \end{array} \right]$$

This rule, which produces *waxt* in Deliberate and Colloquial speech from Formal *waqt* 'time', appears to be a post-lexical rule, although it is by no means allophonic. The past tense morpheme is realized as *-t* after obstruents (stops and fricatives) and *-d* after sonorants (vowels and glides). Any stem-final consonant appearing before this *-t* will, by application of the above rule, appear as a voiceless continuant, as in *kuf/kub* 'grind' and *raf/raw* 'go'.⁴ One of the characteristics of post-lexical rules is that they do not have information about word-internal structure, such as the presence or absence of morpheme boundaries. It might therefore seem that, since it applies both word-internally and across the morpheme boundary between the stem and the past tense marker, the rule

³ Many linguists prefer to generate grammatical formatives, such as the allomorphs of the English plural or the Persian optative, directly, positing no underlying form. I subscribe to this view, which is expressed well by Harris (1969), chapter 3.

⁴ The *w* of Kabul Persian, the variety with which I work, behaves like the Iranian *v* in every way but its actual pronunciation.

should be applicable both word-internally and post-lexically. It appears, however, that despite its having been in the language for millennia (Hübschmann 114-115), to be a phonological one in Persian, one which simply applies whenever its structural description is met, without regard to the presence or absence of morpheme boundaries or past tense markers.⁵

The same consonant assimilation rule applies, but with various changes in stems ending in non-labial consonants, all of which have changes in their points of articulation. Examples are 18 verbs like *âmext/âmez*, and a number of verbs in which the past stem ends in *s* but the present stem in *h*. The voicing assimilation and fricativization is still attributable to the post-lexical rule shown above, but the changes in point and manner of articulation are definitely the kind children learn some time after they have mastered the regularly applying phonological rules.⁶ These stem alternations are historical relics and should be described with mechanisms which do not masquerade as phonological rules.

The large class of denominal and deadjectival verbs,⁷ formed by the addition of *î* to the present stem (often a noun) is a productive class, at least in Iranian Persian.⁸ Given a new borrowing from English such as *modem* ‘modem’, Persian speakers could construct a new verb *modemîdan* ‘to (communicate) via modem’. The extra syllable would be added automatically, though it is not a phonological process. That it is automatic in Persian is shown by doublets such as *kuştan/kuşîdan* ‘kill’.

We thus have only two classes of verbs, down from five: (1) those whose stem alternations are entirely predictable—invariants like *kandan* ‘dig’, derivatives like *raqsîdan* ‘dance’, consonant-assimilation rules such as *yâf/yâb* ‘find’; and (2) those whose alternations are so idiosyncratic as to need to be learned/acquired one at a time. This last class includes the truly suppletive verbs such as *dîd/bin* ‘see’ and the partially suppletive class, in which “phonological-look” verbs can describe the alternations, but the forms of the verbs themselves give no clue as to which rule is to be applied. Phonological rules apply to forms which meet their structural description; rules that apply otherwise are descriptive of phenomena which are not phonological, but are (in most cases) historical relics. This is easily understood in the case of very frequent verbs, like *am/are/is/was/were/be/been*, and *ast/bud/baş* ‘be’.

The advantages of the Lexical Phonology approach include (1) cutting the number of verb classes in Modern Persian from five to two; (2) avoiding the necessity of marking even invariant verbs as “exceptions” to certain rules (cf. MPVM, p. 385); (3) clearly separating, on a principled basis, those phenomena for which a child could plausibly acquire a rule from those which must be learned separately, either by correction or by observation of parents’ and elder siblings’ speech; and (4) allowing the use of rules to describe historical relics—once, but no longer, rule-governed—and other oddities without the implication that those rules have the same status as those describing predictable phonological phenomena.

⁵ There are some rules which apply *only* across morpheme boundaries. The coalescence of *a* and *i* to long *e* in Kabul Persian is attested only when the *i* is the Ezâfat, so *xâna+i+kalân* ‘big house’ is pronounced *xâne kalân*. But morpheme-internal instances of *ai* are realized as *e* only variably, as noted in Henderson 1975.

⁶ Data from children acquiring Persian natively are lacking and, sadly, likely to remain so for a long time. It is conceivable that children may acquire a rule for one of the most frequent consonant alternations, such as past *š* alternating with present *r(d)*. experiments in which past tense forms were elicited from children given made-up present tense verb forms like **melîram* could show whether they were more likely to produce an invariant such as **lîrdam* or a rule-obeying **liştam*.

⁷ Not all are derivational. There seems to be no reason to derive *xandîdan* ‘laugh’ from an underlying noun *xand*; quite the opposite. The formation, distribution, and semantics of Persian verbs have been interestingly discussed by Zand (1991).

⁸ As noted in Henderson 1978, Kabul Persian speakers seem to prefer using the separate verbs *kardan* ‘do, make’ and *şudan* ‘be, become’ with a substantive, thus preserving the transitivity/intransitivity distinction.

REFERENCES

- Folarin, Antonia. 1987. *Lexical Phonology of Yoruba Nouns and Verbs*. PhD dissertation, University of Kansas, p. 15ff.
- Haïm, S. 1959. *Shorter Persian-English Dictionary*. Tehran: Barouchim & Sons.
- Harris, James, 1969. *Spanish Phonology* ch.3. Cambridge, MAL MIT Press. Hübschmann, H. 1895. *Persische Studien*. Straßburg, pp/114-115.
- Henderson, Michael M.T. 1975. Diglossia in Kabul Persian Phonology. *JAOS* 95, 651-654.
- _____. 1978. Modern Persian Verb Morphology. *JAOS* 98.4:375-388.
- Zand, Hassan. 1991. *Aspects of Persian Intransitive Verbs*. PhD dissertation, University of Kansas.

