LEXICAL DOUBLETS AS A DERIVATIONAL DEVICE IN PERSIAN: THE ARABIC FEMININE ENDING

JOHN R. PERRY  
(Chicago)

0 Introduction

0.1 Twenty-four years ago, Zsigmond Telegdi summed up the basic problems posed by the massive incorporation of Arabic loanwords into Persian: "Quels facteurs ont déterminé leur sélection? Et comment se sont-ils intégrés au vocabulaire? Il est à peine douteur qu'avec la rénovation de sa substance, le vocabulaire a été réstructuré. Comment ce process s'est-il accompli...?" Telegdi argued on the basis of earlier studies by Lazard that the majority of Arabic borrowings into Persian are in origin mots savants, incorporated in the writings of bilingual Iranian scholars over the course of several centuries after the end of direct Arab rule in Iran — i.e., taken not from a living Arabic vernacular, but from an already "classical" literary Arabic that continued to serve as a source for neologisms (such as Latin, after the demise of the Roman empire, continued to enrich French and the other languages of medieval and modern Europe). Much of this initially specialized, learned vocabulary then "trickled down" into everyday usage.

The semantic processes by which this lexical assimilation was effected — which in the Latin-script documents of European vernaculars are typically signaled by orthographic changes, generally indicative in turn of phonological evolution — are for the most part unacknowledged in the case of Arabic loanwords in written Persian. Whatever their stylistic register (nonce Fremdwort, assimilated Lehnewort, colloquialism, etc.), they appear in the Perso-Arabic script in precisely the same form as their Arabic atyama. This convention gives the erroneous impression of a timeless, monophonic, monomorphic and monosemantic lexical superstratum, and restricts our recognition of the ways in which Arabic mots savants were incorporated, sorted and redefined in their semantic and sociolinguistic ranges as they rubbed shoulders with the native vocabulary.

0.2 In only one morphological class of Arabic loanwords is there a systematic, orthographically-signaled, phonetic change triggered by the fact of incorporation into Persian and sensitive to semantic shift or other diachronic processes. Those
Lexical Doubles as a Derivational Device in Persian

sleep are instance nouns, sleep or sleeping are action nouns; or it refers to a particular manifestation of a quality: a kindness is an instance noun, kindness is a quality noun. A product noun denotes the outcome of an action, whether tangible (a script) or intangible (an analysis); it differs from an instance noun in that its referent continues in existence after the action that produced it has ceased. As the English examples illustrate, distinct categories may be formally differentiated (sleep/nap) or they may be isomorphs and their usage distinguished only syntactically ([a] kindness); this is no less true in Persian. Arabic grammatical terminology recognizes some of these as morpholexical classes (e.g., the masdar, or action noun, the ism al-marra or instance noun), but this designation cannot automatically be applied to loanwords: Persian 'emdrat 'a building, edifice' is not an action noun, even if its Arabic etymon technically (and historically) was.

To summarize the -at-e dichotomy as a system: abstract or intangible concepts of relatively low imageability, realized frequently as action nouns and quality nouns, tend to be affiliated in -at; thus esgaldan 'presence', fassiyat 'personality'. Concrete or tangible entities, and concepts of relatively high imageability, tend to be affiliated in -e: thus khane 'bowl', esfere 'sign, reference', and personal and place names (nadije, soltanaye). In the case of more abstract or less imageable referents, the overriding criterion is generally that of mass noun vs. count noun, i.e. non-pluralizing loans tend to end in -at (rotubat 'dampness', mokhamevat 'resistance'), while pluralizing loans tend to end in -e (esfere-ha 'indications, signs', tajrobe-ha 'experiences, experiments', estfede-ha (as nap) 'the uses of (oil)').

Action nouns frequently used independently, i.e. as nominalizations of a verb phrase (and thus as mass nouns), tend to end in -at, whereas those used as a verb (in conjunction with an auxiliary such as kardan) tend to end in -e: thus esgat 'scold', esteftehar (-e doba) 'granting (of a prayer)', but estfede kardan 'to use', esfere kardan 'to refer, show', mosafat kardan 'to study'. It will be noticed that the latter forms often coincide with count-noun derivatives of instance or product as mentioned above (esfere 'sign, indication, gesture', etc.

Mots savants that have remained in the higher literary registers tend to end in -at (e.g., hekdast ' anecdote'), whereas loans that have assimilated into the everyday vernacular tend to end in -e (e.g. gese 'story'). The migrants that have shifted from -at to -e have ipso facto demonstrated a degree of assimilation, and may often be shown to have changed or expanded their semantic or syntactic scope in accordance with the criteria outlined above. Thus tajrobeh, as the word was written before the thirteenth century, was used only in the singular, 'experience, testing'; maqdar extended its meaning from 'speaking, discussing' to 'discourse, treatise' during the twelfth century before definitively shifting to maqdas (2.2).

It will be obvious that there are many apparent exceptions to the individual tendencies enunciated here, but also that the criteria are to a great extent

---

2 Similarly, I have transcribed the third vowel of loanwords such as mardeo as e, reflecting the usual Persian pronunciation, rather than a, the canonical Arabic vowelizing; this mixed system of strict transription for consonants and broad transcription for covert vowels, a common and useful device for Persian, does not misrepresent the Arabic-script form of the words discussed.


4 Cf. ibid., 8-10.
interdependent. A greater degree of assimilation into the vernacular often implies some semantic extension or specialization, such as the acquisition of instance or product noun status, which automatically entails count noun status and a higher imageability; from a functional standpoint, several distinct explanations may often be adduced for one and the same instance of shift. The comparative chronology of incorporation or assimilation has also exercised an influence. Some 150 common count nouns in -at, both tangible and intangible (e.g. mamlakat ‘kingdom, country’, emarat ‘building, edifice’, modest ‘disaster’, daghat ‘consequence, sequel’) were borrowed in the earliest period of literary Persian, and their count-noun metonymy either already existed in Arabic or was acquired soon after their incorporation into Persian (by the early eleventh century); these have evidently resisted the shift to -e undergone by such loans as maqalat and tajrobat.5

0.3 The one hundred or so doublings among these feminine-ending loans are a valuable set of guinea pigs with which to exemplify and test the various criteria and processes of -at/-e affiliation and shift. The wonder is that it took so long for this intriguing lexical class to excite the curiosity of Persian scholars. During the 1950s, Mojtaba Minovi in a study of the subset of loans in -iyat/eye6 and Ahmad Bahmanyar in the ‘Orthography’ section of the introduction to Dehkhoda’s lexicon were the first to point to a systematic dichotomy at work: the latter perceptually commented that Persian speakers had put this orthographic happenstance to creative use, and identified six pairs of the doublings treated below.7 While my Form and Meaning... was still in press, two Iranian scholars published discussions of the feminine-ending doublings (one as a critique of the other, and neither as yet having read my work).8 The present article is an attempt to further the discussion in a theoretical direction by incorporating or critiquing observations of A’lam and Sadeghi, revising some of my earlier analyses, and comparing other types of doublings in Persian and other languages.

The majority of Persian -at/-e doublings result from the diachronic shift alluded to; instead of completing the shift to -e, processes such as govāt > govve (2.1) have left a lexical trace in -at which preserves a distinct part of the meaning or usage. Other pairs, such as sarārat ‘wickedness’ vs. sārdār ‘spark’ (1.1), were borrowed independently of one another; though their affiliation reflects the appropriate semantic criteria, it distinguishes not an acquired semantic change but an original homonymy in Arabic resulting from one of three processes: (i) a coincidence of roots (‘radical homonymy’), or (ii) a coincidental

---

5 Ibid., 114–5, 198–9.
7 Lajinīmey-e Dehkhoda, Mqsid-e 172; mentioned by Sadeghi, 42–3 (see following note). For others who recognize (or failed to recognize) this phenomenon, see also Perry, Form and Meaning 5–6 and A’lam, 22–24 (see following note).
change of vowel in modern Persian voslat (voslat is also recorded in earlier usage) is perhaps due to contamination by the cognate loan vasti ‘joining, connection’ (Arabic wasta, which is the unitary noun derived from wasti, is an Arabic grammatical term, ‘post-juncture quiescent alif’, irrelevant to and not found in Persian). Essentially this doublet pair represents a disambiguation of separate though homonymous derivatives of one and the same root in Arabic, though in view of the transparent metaphorical relations between these cognates, and the likely influence of voust, it seems possible that some blending has also taken place.

A more straightforward case is resālat/resēlī (A’lam, 26; Sādeqi, 33; Perry, 57–8, 129–30), at least synchronically speaking: the endings neatly distinguish between the quality noun ‘apostleship, prophethood’ and the homonymous product noun from the same root ‘missive, (written) dissertation’. Diachronically, however, these common and culturally freighted loans have been nicely entangled. Both words were originally inducted in the one form resālat (10th–11th centuries), and each had a broader range of meaning, reflecting the metonymic continuum of their Arabic etyma: ‘prophetic status/mission; (secular) diplomatic mission; diplomatic correspondence’ (as in divid-e resālat ‘the chancellery, court secretariat’); ‘a formal letter, essay’ (divid-e resālē, the Sarnānī and Ghznavī chancellery, using the Arabic ‘broken’ plural of resālat ‘letter’; cf. ‘a r 30 resālat-hā ‘poetry and essays’ in Nāṣer-e Ḫosrow). Not until Saʿdī does the form resēlī ‘letter, treatise’ assert itself consistently to distinguish the tangible court noun from the abstract or collective mass noun. In present-day usage resālat ‘apostleship’ is a specialized religious term, resēlī ‘dissertation’ is a staple of academic and student jargon.

Similarly, eqāmat/eqāme (A’lam, 25; Sādeqi, 34; Perry, 71, 73, 106, 107) are recorded earlier (15th century) in the single form eqāmat, and were not spelled contrastively before the seventeenth century. The derivational homonym here is particularly subtle, involving ‘merely’ a difference of verbal aspect in the Arabic base verb; however, as illustrated in the derivational tree (figure 1),9 idioms evolving from Theme IV (which adds a causative or declarative dimension to each aspectual sense) are sufficiently disparate to justify a separate lexical entry. This is not an option in traditional, morphology-bound Arabic lexicography, since the masdār of Theme IV (unlike that of Theme I) can have only one form. In Persian, however, the shift to -e by the imperative-causative/declarative sense ‘initiation, performance of a formal or ritual act; declaration “qad qāmat al-sulād” as a bidding to prayer’ and the resulting exclusivity of the stative sense to the -at reflex (‘residence, sojourn’) has accomplished this useful taxonomic task. The result is two action nouns, each of which readily verbalizes with the addition of a Persian auxiliary (eqāmat dāštān ‘to reside, dwell’; eqāme kardan ‘to initiate, perform; announce the bidding to prayer’). This may

9 See Perry, Form and Meaning 71; for the suppression of the conventional object (‘flocks’) in Theme IV, see A. de Biberstein Kazimirski. Dictionnaire arabe-français (2 vols.) Paris 1850.

LEXICAL DOUBLETS AS A DERIVATIONAL DEVICE IN PERSIAN

thus be termed a ‘pragmatic’ split, i.e., not consonant with the ‘schematic’ semantic characteristics of -at and -e affiliates (unlike the four other examples in this section).

\[ \text{QWM} \]

‘stand up; be standing’

\[ \text{QWM}_2; \text{eqāme} \]

‘be standing, stay’ (static, resultative)

\[ \text{IV} \]

eqāma[1]

‘set up, initiate, institute, perform; declare to be initiated’

\[ \text{IV} \]

eqāma[2]

‘halt (one’s flocks); stop; settle; reside’

\[ \text{P. eqāmat₂} \]

‘initiation, performance; bidding to prayer’

\[ \text{P. eqāme P.} \]

‘residence, stay’

Fig. 1 Derivation of eqāmat/eqāme

The pair seḥba/sehbe (A’lam, 24; Sādeqi, 34; Perry, 83) constitutes a near schematic split between a mass noun of state or quality and a tangible count noun of instance or product. The origin and status of sehbe, however, are questionable. The meaning ‘signature, endorsement, validation’ is not found for Arabic sikha[1], being more appropriate to the ‘declarative’ Theme II infinitive, taskih. Persian sehbe is probably a fabrication, using the well-known concretizing function of Persian -e added to the conventional Arabic notation sikkah, it is valid’ (with sympatric vowel change) to produce a hybrid seḥba ‘signature, instance of notation of endorsement’.  

1.2 nas’t ‘genesis, growth’

nās’e ‘drunkenness, a “high”; drunk, high’

marārat ‘bitterness’

marāre ‘gall bladder’

maymanāt ‘fortune, prosperity’

maymane (a) ‘right wing’ (of army)

(b) name of a town
Among the ‘asymmetric’ homonymic pairs, naš’at/naš-e (A’lam, 29; Şâdeqi, 34; Perry, 81, 88) barely qualifies for inclusion on a number of technicalities. The Arabic roots are not identical (NS ‘genesis, growth’ for naš’ and NŚ ‘drunken dizziness’ for naš-e; this last form, with hamza, does not occur in Arabic, and even in Persian is only an alternative to naš-e, which does recapitulate one of the Arabic forms naš’at/niswa/naswa). The pair martrat/mardare (A’lam, 28; Şâdeqi, 32, 35; Perry, 57, 59) displays a minimal homonymy by dint of metonomyny in Arabic; the schematic split is original, not the result of shift. This is true also of the Classical Persian pair maymana/maymana (A’lam, 29; Şâdeqi, 34, 35; Perry, 45–7, 59), a rather more interesting case, as figure 2 will illustrate.

The root YMN encodes a locational/directional concept (YMN₁), with a metaphorical extension (YMN₂). The physical, locational sense surfaces in a

locus noun (Arabic īsm mašār, īsm šarg) ‘right wing’, a specialized military term (maymana₁, cf. its counterpart, maysara ‘left wing’) which was borrowed into early literary Persian. The metaphorical sense surfaces in a quality or state noun (Arabic masdar mi’īm), which is of the same form (mâyûs) as the locus noun; this maymana₁ was also borrowed into Persian, and the resulting homonymous clash was automatically avoided by schematic application of the -e/- split. There is, however, a second maymâne in Persian, viz. Mainmâne, a toponym in north-western Afghanistan. This conforms with the application of the -e ending to toponyms, and might fairly be interpreted as a locus noun (cf. Arabic marâdā ‘pasture, grazing ground’ as the Iranian name of Maragheh). If so, the history of the town’s name would claim a derivation through the metaphorical sense (maymana₁, a word not found in the Arabic dictionaries), not the (pre-empted) locational sense; i.e., ‘place of good augury’, not ‘place on the right/to the south.’

The next two pairs (yabduyayat-e, mâyânavât-e; Perry, 24, 28, 39) illustrate the disambiguation of nouns of quality (here, social category, i.e., religious affiliation) from toponyms or collectivities. The same principle is at work in the archaic ziyâyâyat/zâyâyîyî (Perry, 34), distinguishing the quality from an anatomical element characterized by that quality, and amniyat/amniyyî (A’lam, 25; Şâdeqi, 33; Perry, 31), the latter term being a modern (but now obsolete) toponym in the analogy of terms for collective bodies such as mânûviyy or sufiyy ‘the Sufis’.

The structural origin of such terms in -iyat lies in the ‘stranded epithet,’ i.e., the adjectival member of a noun phrase such as Arabic al-qâwâl al-amniyya (>Arabicate Persian gowda ye amniyye) ‘security forces’; the head noun (feminine, or a grammatically feminine plural) is suppressed and the adjective, complete with its feminine concord, remains as the collective term. The quality noun in -iyat has been particularly productive for zoobilogists and pseudo-Arabic loans of the type irdiyiyat ‘Iranian-ness, Iranian (national) character’ (now irdigan) and xarâyiyat ‘asshinity’. The last three pairs listed here all illustrate the interplay of these propensities in different ways.

The Classical and literary raviyyat ‘reflection, consideration, thought’ is a genuine Arabic loanword (root RW), logically affiliated in -ar as a mass noun of action. However, it is attested as early as ‘Awfī (ca. 1220) in the senses ‘manner, way, method, rule, course (of action), policy’; these uses are semantically close, but are not justified by the Arabic etymon and are in no case an obvious extension of the basic meaning. Most likely this ‘metaphorical extension’ is actually an Arabicoid compound from the Persian stem raw ‘go’ (cf. a modern near-synonym, the Persian action/instant noun ravêl). As a count noun of instance, raviyyat in this sense shifted to ravye, creating the ‘doublet’ which is found in several Persian dictionaries produced in Iran a few decades ago, not only in the sense ‘way, method, etc.’, but also in the meaning ‘reflection, etc.’ which is proper to raviyya; this popular confusion is reflected in the entry ravye in, e.g., Miller’s Persiske-russkii slovar’ (Moscow 1960) and in the entry raviye.
2. Semantic change

2.1 āvāt ('divine') sign, portent
qovvat 'strength, might'
mordā'ebat 'observation, supervision
parāt 'Sufi brotherhood
mordā'ebat 'companionship, society'
mordā'ebat 'assistance, aid'
hāvālat 'transfer, handing over
hamdāsī 'gallantry, knight-errantry
ketābat 'art, style of writing
mas'ālāt 'asking, begging
is-kāyāt 'condolence, mourning
nawāb 'turn, watch'
żarbāt 'a beating, blow
martabat 'status, rank, position'

āvāt 'scriptural verse'
qovvāt 'power, force, faculty
mordā'eb (devout) contemplation
tarīqā 'way, method, manner
mordā'ebat 'interview'
mordā'ebat 'pecuniary aid, advance
hāvālat 'bank draft'; +V 'to transfer, change'
hamdāsī 'epic (poem)
ketābat 'piece of' writing, motto, inscription
mas'ālāt 'problem, matter
is-kāyāt 'Moharram passion play
nawāb 'intermittent fever, malaria
żarbāt 'a blow, stroke, lash
martabat (repeated) time, occasion

erādat 'goodwill, devotion'
mordā'ebat 'returning('
erādat 'will(power), resolve, intent'
mordā'ebat 'reference'

2.1 Whereas the majority of doublings in the preceding section were borrowed (or coined) separately, only three involving a partial shift (resātul-ā, ẓādmat-ā and raviyat-ā), most of the examples in section 2 are the result of this historical shift and its residual traces in -at. The exceptions are mordā'ebat and mordā'ebat, which are fourteenth-century neologisms coined independently of their analogues in -at. In every case there is a transparent (or historically traceable) semantic relation between the pairs. The first eleven listed here may be classed as cases of straightforward lexical specialization, where the term in -e represents a more specialized or concretized (tangible or highly imageable) instance of the quality, activity, etc. expressed by the historically earlier term in -at. Thus āvāt (ālam; Ẓādeqi, 34; Perry, 82, 83, 105, 107, 131) is a specific kind of sign or wonder from God (āyāt), mordā'eb (ālam; 27; Ẓādeqi, 38; Perry, 66, 209) is a specific and tangible kind of assistance. Often this development is accompanied, schematically, by the attainment of counting status in the -e reflex: hāvā-āt, hāvā-āt 'bank drafts, etc.'; ketābāt-āt 'mottos, inscriptions, etc.' - fourteen of the -e reflexes in this section are countable, as against five of the -at reflexes. There is one evident exception to the semantic polarity, in that tarīqā 'Sufi brotherhood, dervish order' would appear to be more specialized (and concrete, and imageable) than tarīqā 'way, manner' (ālam; 27, Perry, 54, 59, 199).

A clear illustration of the historical process at work, well documented in texts, is the pair qovvat/qovvāt (ālam; 27; Ẓādeqi, 34; Perry, 84, 85, 106, 107, 108-9, 150-1). The original form was qovvat, used not only in its general sense of (physical) strength, might but by Avicenna and al-Ghazali in the eleventh century as a specialized philosophical term, (physiological or mental) faculty, power, as qovvāt-āt 'dandāne 'facultas intelligens', qovvāt-āt 'faculty of discernment'. By the middle of the twelfth century, however, the shifted form was being widely used for these scientific senses (qovvāt-āt 'rebellious force', in Nezāmi 'Aruži'), and exclusively so by the end of the sixteenth century; the base form in -at, meanwhile, became exclusively associated with the mass noun of quality 'strength, might'. In modern Persian the count noun of instance qovvat refers additionally to military and industrial power: qovvat-āt qovvāt-āt 'military forces, qovvāt-āt berāq 'electric power'. The distinction is maintained in derived verbs, as qovvat dáddan 'to revive, vigorize' vs. qovvät dáddan 'to motorize, electrify'.

The pair ta'zayat/ta'ze (ālam; 26, Ẓādeqi, 32, 35; Perry, 63, 106, 107, 145, 148, 182-3) introduces a further dimension. The synchronic contrast is schematic, between an action noun in -at (+V, i.e. belonging primarily to the verbal system, as in combination with gofün to express 'to offer condolences') and a fully nominal substantive in -e (+N); either a specialized instance of the action (the dramatic representation of the Kerbelaa tragedy among the Shîa of...
Iran) or a concrete symbol of the mourning for this tragedy (the cenotaph of Husayn as carried in procession among the Shi‘i of South Asia). The shift -ar- to -e (as in Turkish by retell +V) ‘to condole, etc.’, ta‘zeyat as ‘Moharram passion play’ has a Persian synonym šabih-zudni. Diachronically, however, ta‘zeyat was not acquired through a straightforward doublt split internal to Persian, but seems to have followed a route outlined in figure 3.

The original ta‘zeyat ‘mourning, condolence’ was borrowed from Arabic through Persian into Turkish, where in most or all dialects (notably Ottoman, Azerbaijani, Uzbek and Tatar) it soon shifted to ta‘seye, with the same meaning. In some cases (as in modern Azerbaijani) it evolved the instance noun sense ‘funeral (obsequies)’, among the Shi‘i Turkmen of the proto-Safavid period this was further specialized as referring par excellence to the ritual commemoration of Kerbelā, and with the establishment of the Shi‘a in Iran the word entered Persian usage alongside šabih-zudni. In terms of immediate provenance, Persian ta‘zeyat is thus a shifted Rückwanderer from Turkish.

Arab.

\[
\begin{align*}
\text{ta‘ziyyat} & \quad \text{‘condolence’} \\
\text{Classical Pers.} & \\
\text{ta‘zeyat} & \quad (+V) \quad \text{‘condolence’} \\
\text{Turk.} & \\
\text{ta‘zey} & \quad (+V) \quad \text{‘condolence’} \\
\text{Pers.} & \\
\text{ta‘zeyat} & \quad (+V) \quad \text{‘to condole’} \\
\text{ta‘zey} & \quad (+N) \quad \text{‘funeral; Moharram rites’} \\
\text{Pers./Indo-Pers.} & \\
\text{ta‘zey} & \quad (+N) \quad \text{‘Moharram drama/cenotaph’}
\end{align*}
\]

Fig. 3. Derivation of Persian ta‘zeyat/ta‘zeyat

The last five pairs listed in this section involve either lexical specialization in both reflexes, or a selective metaphorical usage exhibiting a pragmatic, rather than a schematic, distinction. Thus nawbat and nawbe (A‘lam, 29; Perry, 83, 129) earlier shared both current meanings, as well as a variety of quite concrete specialized uses, such as ‘sentry; ceremonial drumroll; drum’, before each reflex stabilized. The subtle distinction between šarb and šarbe (Šādeqi, 42; Perry, 86) originates in a metaphorical extension of the basic meaning, as in govaat/gove and jariay tiyare, but is not adequately distinguished in Persian dictionaries: both forms, from an Arabic instance noun (and their plural šarabat, and the basic action noun šarbat) may be used for a physical blow or shock, but šarbat (as the more literary, high-style reflex?) is preferred in figurative use, as šarbat-e qāde ‘deceitful decisive’, while šarbe (as the more vernacular reflex?) is preferred for specialized or technical versions of the physical sense: šarbe-yé fannī ‘a knockout’ (in boxing), bit šarbe-yé šalā ‘twenty lashes’ (a flogging, as a statutory punishment). It is tempting also to see an example of the latter in phrases such as motur-e do-sarbe/fārābī-sarbe ‘two-stroke/four-stroke engine’, but here the ending -e is probably the adjectivalizing suffix as in, e.g., ṣīl-e do-nafar-e ‘double room’.

Martab has transferred the spatial reference of martabat to a temporal one (A‘lam, 28; Perry, 45, 120); though both are count nouns, the usual plural mardesteb refers only to the spatial martabats, as in ba-mardaste behar ‘considerably (i.e., by degrees) better’; mardesteb is generally seen only as a singular after a numeral (do mardesteb ‘twice’, etc.). In erdadeerad (A‘lam, 25; Šādeqi, 53; Perry, 70-1, 131-2), Persian distinguishes two modes of volition that had been differentiated idiomatically in Arabic (cf. eqdan/eqdame). Arabic ardada bihi xayar ‘he wished/demanded X for him’ generates an action noun irda[t] expressive of a conventional benediction, which becomes Persian erdāt as in erdātamand ‘(your) wellwisher; yours sincerely,’ in closing a letter; Arabic ardāda xinhu X ‘he wanted/demanded X from him’ generates the same form irda[t] in the sense of imposing the speaker’s intention for himself, which becomes Persian erdāe. Similarly, the mass action noun mordješar and the count noun mordješ ‘e (A‘lam, 28; Perry, 66, 104, 125) originate one in an intransitive, and the other in a transitive, use of the same Theme III verb in Arabic. Both may combine with an auxiliary, but whereas mordješar kardan ‘to return, go back’ is the primary context for the -ar reflex, putting it squarely in the verbal system, mordješ-e tarden ‘to refer, have recourse (to)’ is a denominal derivative from the specialized count noun reflex in -e (pl. mordješ ‘references’).

2.2 xetābat ‘office of preacher; address’

\[
\begin{align*}
\text{maqālat} & \quad \text{‘discourse’; +V ‘to speak’} \\
\text{moldiyeŠ ‘kindness, favor’} \\
\text{eŠnam ‘aid, assistance’} \\
\text{eŠbeŠ ‘expertise’} \\
\text{eddā ‘period of probation’}
\end{align*}
\]

maqāle ‘treatise, article’
moldiyeŠ ‘letter, note’
eŠnam ‘alms, charitable gift’
exŠbeŠ ‘an expert’
eddā ‘number’
as, e.g., an attributive or predicative adjective, in a fixed collocation, or in a lexical composite.

The first three pairs illustrate the designation of the -e reflex as a predicative adjective. Sōhre (A’lam, 27; Perry, 84, 85) occurs both in the absolute and in ezāfə phrases such as Sōhre-ye zāhe o ‘Jamīn ‘the talk of all classes’, and is evidently calculed on the analogous idiom in Arabic (zāhe šubāran; he became a notoriety’, i.e. notorious). We shall see other examples in these sections of an Arabic ezāfə phrase as the apparent model for an Arabic loanword in Persian, either alone or in a Persian ezāfə phrase, with a contrastively used -e ending. A similar derivable could be argued for gelle (A’lam, 26; Sadeqi, 33) and zeyāde (Sadeqi, 44; Perry, 121). In the latter case, it can be shown that the unshifted -at form first evolved an adjectival usage, both from literate instances (in Sāfī) and by the existence of the derived, morphologically Persian, quality noun zeyādat ‘excess, superabundance’ (cf. raḥa‘ and raḥāmāt, replacing quality nouns in -at which have evolved into adjectives). The general-purpose predicative adjective and adverb zeyād is a truncated form of zeyādat, and zeyāde is restricted to predicative phrases such as zeyāde az hadd ‘exceeding the bounds’.

In kolliyat-e (A’lam, 27; Sadeqi, 33; Perry, 28, 31-2, 39, 111) and manzelat-e (A’lam, 29; Sadeqi, 34; Perry, 45), the archaic abstract nouns in -at are adapted as -e in everyday ezāfə or prepositional phrases following Arabic models. In mokhebsaat-e (Sadeqi, 42; Perry, 66-7) the -at reflex is used for the quality/state noun only, while the -e reflex is exclusively verbal; in fact mokhebsaat-e is additionally transitive and involves a subtle shift in meaning, not unlike the more obvious change in mordhez-e (2.1). The distinction expressed between dreyat and dreye (A’lam, 27; Sadeqi, 34; Perry, 112) is similar, with the extra complication that both reflexes may be used adjectively, though with subtly different senses or at least in different stylistic registers: dreyat (seen as a predicative adjective, ‘borrowed’, in Classical and literary Persian) has evolved a modern all-round adjective dreyat ‘artificial’, of which dreye ‘borrowed; artificial’ is a vernacular variant. Mohāfšeṣat (A’lam, 28; Perry, 56, 82) is both nominal and verbal in usage, while the shifted form occurs only in a specialized nominal compound mokhebsaat-kārd-e (‘political’ conservative) and its derivatives; hamāsde (2.1) is similar in occurring most often (though not exclusively) in the compounds hamāsde-kārd ‘epic poet’, hamāsde-kārd-e ‘epic poetry’.

3 Selection for Word Class, Collocation or Citation

3.1 Sōhre ‘fame, notoriety’,
gellat ‘abjectness, hardship’,
zeyādat ‘excess, increase’,
kolliyat ‘totality, universality’,
mazālat ‘rank, degree, status’,
mokhebsaat ‘congruence, accord’,
dreyat ‘loan; borrowed’,
mokhefsat (+V) ‘preservation’

This category overlaps with several others. We have already seen in the case of, e.g., havāliat-e (2.1) how one reflex may specialize as an action noun and the other as a count noun and/or in a reconstituted verb or, in the case of xebārat-e (2.2), how the shifted form may become a type noun and adjective. Here I wish to designate those cases where one reflex (generally the shifted one) is exclusively reserved for a syntactically or morphologically marked role.

3.2 Efādat ‘communication,’
efāde ‘explanation; benefit; arrogation’, +V ‘to put on airs’,
nedārat ‘watching, supervision’,
sirat ‘conduct, character; honor’,
rehle ‘the Rihla (Travels) of’,
moshef ‘knowledge, wisdom’,
kāra‘ (‘large’) quantity,
ja‘m + kāra ‘plural of quantity’
6.1. Only time and usage will determine in each individual case whether the shift -at > -e will be aborted, completed, or resolved as a doublet pair.

5 Towards an Analysis of Persian Doublets

The total of sixty pairs discussed above may be broken down, in partly overlapping categories, into the following three types:

- Current synchronic pairs (1.1, 2.1, 3.1, 4.1)
- Semantically/syntactically distinctive synchronic pairs (1.1, 2.1, 3.1)
- Asymmetric/puative diachronically distinctive pairs (1.2, 2.2, 3.2)

Despite the clear contrasts presented by a core of these doublets (minimally, the twenty-nine pairs of the second type above), acceptability judgments of the remaining pairs recorded or postulated are likely to remain subjective. A'lam counts thirty-seven pairs, without aspiring to be definitive. Sädeqi adds eleven, to make forty-eight. In Form and Meaning, I originally included forty pairs as showing current synchronic contrast, examined thirty of these as being similarly (or differently) contrasted in Turkish, Urdu, etc. (cf. 6.1), and mentioned a further half-dozen possible diachronic contrasts.

4.1 *Aq'idat-e* 'belief, opinion'
*etārāt-e* 'indication, sign'
*kafarāt-e* 'atonement, expiation'
*mudābrekāt-e* 'partnership'
*māzdarāt-e* 'fear, danger'
*māzbarekāt-e* 'intercourse'
*rozbāt-e* 'rank, status, degree'

4.1.1 In these pairs, both forms are in use without any apparent difference in meaning or fixed distinction of usage other than personal preference based upon perceptions of stylistic register or other felicity conditions. The -e reflex results from a process of historical shift; where speakers judge the -at reflex to be literary or archaic, though acceptable in some contexts, we may consider the shift to be not yet completed (e.g., *'aq'idat-e,* etārāt-e, *rozbāt-e*). The distinction may be reinforced by selection for word class or usage: *etārā +V,* i.e. *etārā kardan* 'to point, indicate' would be preferred to *etārā +V* (and cf. *māzdarāt-e*).

These examples and others may be found in A'lam, 27, 29, and Sädeqi, 34, 35, 36, 40; they are not treated in Perry, Form and Meaning.

Acta Oriens, Hung, XLVIII, 1995
may no longer legitimately be 'derived' from a verb (ṣarrā, qawwāya, etc.) which does not exist in Persian. They must, on the other hand, be categorized under the semantic or syntactic roles they have assumed in Persian, which may or may not approximate to those of their etymon in Arabic: ṣarāṭ 'wickedness' and qawwāta 'strength, might' as mass nouns of quality, ḡawwē 'faculty, (source of) energy, (military) force' as a count noun of instance or instrument (0.2).

5.2. Though ostensibly confining his analysis within the technicalities of Arabic lexicology, Šādeqi uses ʾem 'noun, substantive' in its general Persian sense, as well as in its technical Arabic meaning in contradistinction to masdar, which exposes his generalizations to imprecision and apparent contradiction. Thus, having stated that ḡawwā 'noun (ʾem) that has also retained its rôle as a masdar — true, if he means that it is an instance noun that may be reversionaled as ḡawwā dādān (2.1.1) — he describes qawwāt as a noun (ʾem) that is counter-intuitively affiliated in -at,14 though as a mass noun of quality this is schematically affiliated as one would expect (as is the contrasting ḡawwē).

Šādeqi notes that some loanwords currently assimilated as -at (he lists eleven) appear in some Classical Persian texts as -e, in Form and Meaning. I also noted ten such occurrences, only one of which (mojāheđe) is common to both lists.15 Do these counterexamples invalidate the characterization of the well-attested historical shift -at > -e as a one-way loss of final /t/? I think not. There is good reason to regard them as either ghost words or fossils of evolutionary dead ends.

First, in earlier manuscripts, Arabic feminine-ending loanwords in Persian are sometimes written with the ʾa marbūtā instead of final ʾa; this hybrid graph in turn is sometimes written without its superscript dots, appearing as a final ḡd'; in which case it is impossible to know for sure whether it represents a true -e or an original -at. In texts where the same word is found elsewhere with orthographic -at, we may assume this is the norm. Unfortunately, subsequent copyists (or editors) may have normalized an anomalous -e or ʾa marbūtā as -at or -e on their own initiative. I suspect that this cycle of orthographic laxity and hyper-correction is responsible for a good many ghost words in -e.

However, there are certainly cases where an anomalous -e is obviously intended, as, e.g., in a line of verse where -at would not rhyme or scan (roxe for roxāt 'leave, furlough' in Hafez).10 A number of such words (e.g., ḡelāyjje, ṭorbē, masḏač, ṣawwhe) did in fact shift to -e in the lexicons of adjacent dialects or languages — Persian of Central Asia, Afghanistan and India, Turkish, Urdu, etc. (6.1). Their occurrence in earlier literary Persian might be termed a case of 'arrested shift', in that they evolved in some idiolect or dialect of Persian in addition to, or instead of, the -at reflex, but were not accepted into the literary standard.

5.3 Discarded doubles: More germane to the present study are early occurrences of both -at and -e variants of a later standard in -at, apparently each with a distinct meaning or usage. Such 'discarded doubles' are not unusual where the surviving form is the shifted -e (see below), but quite rare when the survivor is the 'primitive' -at. A plausible candidate is mojāheđe-e 'struggle, endeavor': this pair appears to have had two distinct connotations, the political-military mojāheđe-e 'fighting for one's liberty, religion, etc.; crusading, crusade', and the ascetic-mystical mojāheđe +N 'strenuous religious devotion, self-mortification'. Such a split would neatly differentiate the two senses (communal warfare on the infidel, vs. individual struggle against one's lusts) which the Sufis attribute to A. jihād, jehđa; this word, of course, as the alternative Theme III mašdar of the root JHD, is theoretically a synonym of mojāheđe(-i), but is incapable of a similar differentiation. The split would also parallel the distinction seen in kordeše-e 'observation, supervision' vs. mordeš +N 'devout contemplation' (2.1); in each case, a Grundbedeutung in -at would have evolved a specialized connotation in -e for an ascetic, Devotional, or Sufi context. However, the evidence presently hand to hand does not support a consistent distinction in usage: though Atār uses mojāheđe in the ascetic sense, Rumi has it in its military connotation, and Sāhī in its basic or neutral sense. Subsequent examples show the -at reflex in all senses.

More numerous are the 'discarded doubles' where the -at reflex is now obsolete or archaic, but may for a limited period have constituted a synchronic contrast with the surviving reflex in -e. The well-known edāre-e (+V, +N) 'management, administration, office' began its career as edāre, but only in the literal Arabic sense of 'passing round (wine cups, etc.)', as attested up to ca. 1280; the shifted form in -e as the figurative action and instance noun 'administration, etc.' appears some time thereafter, but was not extended to the locus noun 'office' until the late nineteenth century. Similarly, Arabic ḡāšyja 'fringe, border (of a garment); followers, retreat; margin (of a text), marginal note' was incorporated into Persian with the two figurative senses, pragmatically differentiated for several centuries as ḡāšyjā Retreat, couriers' (Bayhaqi and Nāṣer-e Qosrow, late 11th century) and ḡāšyē (pl kordeš) 'marginal note' (widespread, from the outset). By the thirteenth century, both meanings were represented by the -e reflex.17

5.4 Other non-doubles: Šādeqi includes among his doubles two pairs which, like the discarded doubles above, have once exhibited the contrasts he claims, but no longer do so. The archaic mordeš-e 'litigation', according to the dictionaries, had already shifted to -e before acquiring its additional vernacular sense 'dispute, quarrel', and both senses are now encoded in the surviving mordeš-e (but cf. da-vāšda-vi, 5.2). In ṭešēf + 'increase; relation (general or philosophical usage) vs. ṭešēf-e 'annexation, noun phrase construct' (grammatically), a past synchronic contrast is unlikely, since ṭešēf is also attested earlier as

---

14 Ibid., respectively 34 and 36.
15 Ibid., 41; Perry, Form and Meaning, 120-3.
16 Ibid., 125, citing M. Šukurov et al., Farhangi zaboni tojīč (2 vols.) Moscow 1969.
17 Ibid., 128.
the grammatical term (and survives in this form and sense in Turkish and Urdu), while āfā in modern Persian has taken over not only the grammatical sense but the base meaning 'addition, increase'\(^\text{18}\).

A phenomenon which seems superficially similar to the 'discarded doublet' syndrome, yet is interestingly different, is what we might call the 'post-shift neologism': in this case an original -at form shifts (with the same or similar meaning) to -e; subsequently an unconnected neologism homonymous with the -e form enters the lexicon through some other process. Three distinct words are involved; the fact that the -at form and the neologism in -e contrast semantically is purely incidental, and synchronically irrelevant — as is the fact that the earlier shift-reflex in -e clashes with the homonymous neologism. Thus motārebāt 'fencing, duel(ing)' is also found in a shifted form motārebē, even before the invention of a homonymous motārebē 'commander, limited company' in the nineteenth century. The latter was a backformation from the participle motārebē 'fellow-duelist, sparring partner, business partner', probably coined in the Ottoman empire and later imported into Iran together with the kind of commercial organization it referred to\(^\text{19}\). Likewise, Classical Persian jabāt (later > jabhe) 'brow, forehead', a poetical loan from Arabic, is not the source of modern jabhe 'front' (military or political). The fact that the military term employs the same anthropomorphous trope in modern Arabic and Turkish — but is not recorded as a metaphor of jabha(l)sjabhe in pre-modern Arabic, Turkish or Persian — reveals it as a calque on French front 'brow, forehead; (battle)front' which was introduced (again, probably through Ottoman Turkish) in the late nineteenth or early twentieth century\(^\text{20}\).

6 Comparison with Other Categories of Doubles

6.1 Feminine-Ending Doubles in Tajik, Turkish, etc.: Even before attempts were made to analyse the -at-e dichotomy in Persian or Turkish, it was recognized that the affixation of individual loanwords did not always correspond in the two languages. Šăđeqi likewise lists fourteen examples of differently-affiliated loans in Persian of Central Asia (Tajik) and of Afghanistan (Dari or Kabol- li)\(^\text{21}\). Does this difference reflect a continuation (even a refinement, or extension) of the functional scheme evident in Persian, is it an accident, or is it the product of a different system?

In chapter seven of Form and Meaning, I examined differences in affiliction not only in the languages mentioned but also in several Turkic languages, Urdu and Hindi, and Malay and Indonesian. I concluded, first, that the synchronically similar -at-e distinctions in the Arabic loan inventories of Malay-Indonesian (and of several African languages) have no direct connection with the -at-e split in Persian. Since these languages did not acquire the bulk of their Arabic loan vocabulary through Persian, this was to be expected. Some borrowings in -at have evidently been passed on ready-processed from Persian (more specifically, Indo-Persian) sources; others have been taken directly from Arabic in a pre-juncture context where final /l/ is articulated; while the -e loans have mostly been taken directly from an Arabic pre-pausal form in /l/, though some may have come through Indo-Persian assimilated forms in -e. The meloities, however, reflect either accidents of etymology or phonotactic requirements of the host language, not the semantic or syntactic functional split seen in Persian. The only doublet split noted for Malay-Indonesian that does indeed have a lexically disambiguating value is sunnah 'Islamic tradition, orthodoxy' vs. sunah 'orthodox rise; circumcision': the first is directly from Arabic (either in isolation or abstracted from the construct ablu'sunna 'the orthodox community'), and the second either from the construct sunna'sunna 'the normative practice of the Prophet' or, more likely, through (Indo-)Persian sonun, where the specialized sense 'circumcision' was already established\(^\text{22}\).

On the other hand, the Iranian, Turkic and Indic languages in the eastern continental Islamic ecumene did acquire virtually the whole of their Arabic vocabulary from the mots savants already assimilated into Persian; the identity of the respective -at-e meloities is more than ninety percent preserved. Allowing for some imprecision in allocating pairs to the synchronic or diachronic/asymmetrical categories, Tajik and Dari each have about ten fully distinctive synchronous doublet pairs, or one-third the number in Persian; unknown to Persian are three of these in Tajik, and at least one in Dari. Five pairs (50%) in each language are not identical. Urdu likewise has ten such pairs, three of them unknown to Persian: these are khafyar (+V) 'sufficiency; suitability, competence' vs. khayā 'enough' (as a predicative adjective and adverb; the same as Tajik); dāl 'instrument, tool' vs. dāl 'penis' (the same as Dari, though with a tendency to reverse the -at-e affiliction) and āfā/āfā (the same as Turkish, cf. below).\(^\text{23}\)

Turkish has about fifteen, or half the number of doublets in Persian, though at least six of them are unknown to Persian. Examples of the latter are atrim (+V) 'resolution, undertaking, setting out' vs. aقيم (+N) 'incantation, spell; amulet, mascot' (schematically disambiguating a radical homonym which appears only as -at in Persian); nualet (+V) 'equivalence, equalization' vs. nüalet 'an equation' (a schematic split: both are -e in Persian); həreket (+V) 'movement, start' vs. həreke 'vowel diacritic', and the similar izaf (-e) 'annexation; extra' vs. izafet 'noun phrase construction' (pragmatic splits; note that in the last two examples the -at-e afflications are reversed in respect of the

\(^{18}\) Šăđeqi, 142; Perry, Form and Meaning, 68, 70, 144.
\(^{19}\) Perry, Form and Meaning, 68.
\(^{20}\) ibid., 119, 122, 123.
\(^{21}\) Šăđeqi, 41–2.

\(^{22}\) Perry, Form and Meaning, 143–5, 157.
\(^{23}\) ibid., 144–5.
basic meaning and the grammatical term, and cf. 5.4).24 Glancing at several other Turkic languages (Azerbaijani, Chuvash, Kazakh, Kirghiz, and Volga Tatar) we find eight widely-shared feminine-ending loan doublets unknown in Persian, three of which are different from the non-Persian pairs in Turkish; seven of those occurring in Uzbek are shared with Tajik.25 Two other widely-shared splits not occurring in Persian (or Urdu) are worth mentioning as clear semantic distinctions, functionally well-motivated in accordance with the ‘Persian’ system: 'ebdrat-e (Tajik, Dari, Turkish, Azerbaijani, Tatar, Uzbek) and g'dayt-e (Tajik, Dari, Turkish, Uzbek). Calqued on the Arabic idiom 'ibdratun 'an expression of', i.e. ‘which is to say, comprising’, 'ebdrat + ablative pre- or postposition ('ebdrat ac. -dan 'ebdrat, se 'ebdrat) are used adverbially and predictively to mean 'consisting of, consists of'. In Persian and Urdu, 'ebdrat is also used as the count noun of instance 'phrase, expression, idiom'; in Dari, Tajik and the Turkic languages, however, this slot is occupied by the shifted doublet 'ebare. Similarly, g'dayt 'limit, extremity; utmost, extreme(ly)' is virtually universal, and in Persian and Urdu also represents the figurative sense 'end (in view), purpose, aim, ideal'; in Tajik, to an extent in Dari, and in the Turkic languages where this sense is represented, it takes the shifted formg'diya.25

It seems evident that literary Persian, though the originator and disseminator of the -at-e system, has not adopted or preserved all the splits that were realized at various times, in different regions of the Persian-writing (but Islamic-, Turkic-, Indic-speaking) world. In the case of Turkish, at least, it is possible to go beyond this conservative observation and argue that the system was actually extended. Evidence is furnished not only by the relatively large number of such doublets foreign to both Persian and Urdu, but also by the general propensity for doublet-formation in the Turkish languages, from both native stock and loanwords, which exceeds that of Persian and Urdu in quantity and ingenuity: examples are Turkish (and Azerbaijani) kišlak (kišlak) 'winter pasture' vs. kišla (kišla) 'barracks'; Azerbaijani nagi 'transportation' vs. nagil ('fairy')'sule' (both < Ar. nagil), widespread pabdah 'emperor' vs. pādah 'chief, governor' (Turkish), pasā 'part of name' (Uzbek), etc. (< Persian pādah, though pādah or other vernacular variants do not surface in literary Persian).27

A singularly elegant indication that the system was intuitively understood and exploited at least as well by Ottoman lexicographers as by their Persian counterparts is furnished by mazāfart-e (4.1). None of the Persian dictionaries I have consulted admits a semantic distinction between the reflexes. However, in

Redhouse’s Turkish Lexicon (1890; essentially compiled from traditional native dictionaries), there is a systematic identifi cation of mazāfart from the subjective meaning ‘fear’ (from the Arabic mazāf māt, or augmented action noun) and mazaf with the objective meaning ‘danger’ (i.e., ‘source of fear’, from the Arabic lūc noun); the plurals mazāfat and mazafv are attributed, logically, only to the count-noun sense mazaf. It may be doubted whether such schematic clarity can be applied to actual usage in modern Persian (or even in Ottoman Turkish), but the entry, even if pedantic, demonstrates recognition of a case of derivational homonymy which could, ideally, be resolved by the -at-e dichotomy.

6.2. Other forms of doublets in Persian: Apart from the -at-e pairs, there are hardly any doublets of either native words or borrowings in Persian. Only one synchronic semantic contrast of the type illustrated under 2.1 comes to mind: da'vā ‘litigation, lawsuit; quarrel’ vs. da'vā ‘claim, pretension’ (both < Ar. da'vā, written with allif māḏāra, 'claim, lawsuit'). Da'vā (written with yd) and so pronounced in modern Persian: a spelling- pronunciation) is apparently the older form and meaning (‘claim’ as both pretension and demand), appearing abundantly in texts (especially poetry) from Daqiqi onward. Unless the final syllable is somewhere attested in rhyme position, there is no proof that it was pronounced as today, and on the contrary every reason to suppose that it was originally pronounced as dāvā, in emulation of its etymology. Da'vā (written in Persian with allif, a pronunciation-spelling) does not appear much before Safavid times; it would seem to have evolved as a bureaucratic-legal specialization (‘claim for a debt or other right, lawsuit’), extending into the Persian and Turkic vernaculars of the time (along the same semantic lines as English argument) to mean ‘quarrel, row’, with verbal idioms ‘to quarrel’ and ‘to scold’. With this last development we may compare literary mardafā 'litigation, lawsuit, complaint' > vernacular mardāf 'dicto, quarrel' (5.4), which seems not to have become a synchronic doublet pair.

Three other pairs represent a registral contrast between a literary, Arabicate standard and a vernacular, assimilated reflex. These are abu-bhu- (< abī-, the Arabic kunya) as used in personal names; amir/mir 'emir' (< Ar. amīr), used as substantive, appellative, and anthropomorphically; and mawla- /molla- ('mullah), also used in names. The members of this last pair are not distributed evenly: mawla- only occurs in combination, as mawlaānā (a title, with vernacular doublet mollaānā) or with a name, whereas molla (in both written and spoken Persian) is also an independent appellative and substantive. The three vernacular doublets are rare examples of early Arabic loans where the phonotactic assimilation into Persian has been maintained orthographically. Examples of usage are: Abu 'Ali ebn-e Sīnā is Avicenna, the philosopher and physician of high Perso-Islamic culture; Bu 'Ali is his vernacular alter ego as a popular magician in a cycle of Persian folktales. 'Mulla Aly Shaw', an eighteenth-century official who is known only from European documents (i.e., in transcription
from the spoken form of his name) would undoubtedly have appeared in a Persian document as Mowlā ‘Ali Șāh; conversely we may assume that his contemporary Mowlā Mottaleb, who appears only in Persian chronicles, would be called Mollā Mottaleb in popular speech. Mir was regularly used to refer to the ruler in Samanid and Ghaznavid literature (and amīr could be used as a sobriquet for less exalted personages such as the poet Mu‘ezzī), but thereafter the more ‘cultivated’ form amīr was preferred for rulers, and mir was applied more often to popular saints, as Șāh Mir Hamzeh. Literary osṭād ‘master, professor’ contrasts in some contexts with vernacular osṭād ‘master (craftsman)’, but not so systematically as in Tajik or Turkish.29

An interesting ‘discarded doublet’ of Classical Persian, the distinction in which is maintained in Turkish, is tāmilīz ‘discernment, discrimination’ vs. tāmilic ‘clean, decent’: modern Persian has, unusually, established the vernacular variant tāmilic as the single form for both senses.

On the fringe of doublehood are variants such as pāst(0/1)farst(0), pildf; the early Classical form with p was generally superseded by an Arabized form with f, which in turn is now liable to be replaced by p in nativist or archaizing contexts.

6.3 Analogies in other languages: Arabic, which has generously provided all the material for Persian doublets of every kind, and most of that for Turkish and Urdu, is itself devoid of true doublets by reason of the rigidity of its lexical morphology — in particular, the demands of ar-rī or assimilation of foreign vocabulary to Arabic morphological formulae. Languages with more elastic terms for borrowing, such as Persian, French, or English, acquire dozens of doublets — overwhelmingly of foreign origin, for ‘doubling’ vocabulary is a means of supplementing native morphological devices, and (as we have seen here) intimately tied to the several historical processes of ‘borrowing’ words. But are these doubling processes comparable in different languages?

Husain A’lam lists a selection of French doublet pairs of the type meubles/mobiles, chase/caise, noting that the first variant is the earlier vernacular descendent of Vulgar Latin vocabulary, the second is a mot savant introduced from literary Latin by scholars from the fourteenth century on; he wonders if this process has analogies in the formation of Persian doublets.30 English, too, has a similar range of doublets (and triplets) contrasting Norman French loans assimilated into Middle English with Latinate mots savants (and some Italian and French loans) of the Renaissance and later: frail/fragle, parson/person, influence/influenza, etc. This historical process of supplementing early vernacular forms with later mots savants, vernacularized in their turn, accounts for the bulk of lexical doublets in both French and English.

A close analogy can certainly be seen in several of the Persian feminine-ending doublets, at least in modern neologisms coined (or, more precisely, imported from the Ottoman realms) without direct reference to their Classical antecedents, such as jāhī ‘front’, motarēbeh ‘limited company’ (5.4), and perhaps motarēbeh ‘intercourse’ (2.1). Even the vernacular da’dā, to the extent that it began as a phonetically hypercorrected, and lexically more specialized, variant of the earlier da’ū, might be regarded as a learned suppletion. There remains the fact that in each case the earlier loan was also a mot savant in origin, and could hardly be considered more vernacular in usage than the latter one. A stronger case can be made at the other end of the historical continuum. The early loans bu, mowlā and mir were almost certainly incorporated orally, rather than as literary mots savants, as is evidenced by their assimilated form and social usage (as names and appellatives); thus their literary, Arabeulate doublets abu, mowlā and amīr (if their first appearance was indeed in written texts, perhaps several generations after the incorporation of their doublets in spoken New Persian) are to an extent learned suppletions from a foreign source.

The vast majority of (feminine-ending) doublets in Persian, however, result from the very reverse of the process seen in French or English: the earlier form is the learned term, and the later ‘shifted’ form is derived directly from it by a phonological and orthographic change which signals a more specialized or concrete sense, or a more idiomatic and vernacular usage. Even where the -e reflex is not derived from the -i form, but represents an independently inducted Arabic homonym (1.1, 1.2), the semantic relation often holds good: for an illiterate soldier in the Ghaznavid army, maymane ‘right wing’ was a more concrete, meaningful and everyday term than the pious abstraction maymanat ‘felicity, good augury.’

There is, however, a strong candidate for comparison on a historical, methodological and functional basis in the English stress patterns of certain disyllabic verb-noun (or adjective) pairs: convért/convent, convict/convict, recess/recess, perfect/perfect, affix/affix, contract/contract, refuse/refuse, and some 150 more.31 In each case the word with initial stress is a verb, the word with final stress is a substantive or adjective. Most are cognate and monosemous, but the last two examples illustrate homographs in which the meaning, as well as the lexical category, is distinguished by the accent (contract ‘business agreement’ and refuse ‘to reject’ vs. refuse ‘garbage’ — cognate, but polysemous; the last additionally distinguished by a voiceless syllabant). These two pairs may be compared with the disambiguation of derivational homonyms in maymanat-e or egdamat-e. How else is this system similar to the Persian one?

---

29 Ibid., 220.

Acta Orient. Hung. XLVII, 1995
It targets borrowed (Latinate) vocabulary. It has been in action from the early period of Modern English, and is still productive; the Persian shift (of which doublet formation is a part) has been in action from the earliest period of New Persian, and is apparently still productive (4.1). The English system works by a process of lexical diffusion, schematically splitting an original homograph by means of a sound shift, the shifted stress marking the noun or adjective (the more concrete or imageable class), and the trace marking the verb (the more abstract, less imageable class). Persian shifted (i.e., nominal) forms such as elāre, egāme, zarbe, morāje, havače, ārāye may then be verbalized by combining with an auxiliary; English shifted forms, too, may be verbalized under the same stress pattern and meaning, as in process (distinct from the original to process) and to impact (transitive; similar to the earlier impact on, different in meaning from impinged 'cramped together'). In each case, these new lexical entries are created from a single original.

The process in English would appear to be a specialized subroutine of the general tendency to ‘verb’ nouns without morphological change. In Persian, distinguishing nominal from verbal usage in original homonyms is only one of several possible functions of the -at-e split, and on the other hand its inventory is only one third the size of the English stress-class doublets. Elegant as it is, this contrastive marking is not strictly necessary to disambiguate usage either in English or Persian; the contextual morphology and syntax is generally sufficient (indeed, the English pairs remain undifferentiated in writing without great risk of confusion, and the shift-induced split in Persian may often be judged redundant, though less so in the case of homonyms where it is not attributable to a historical shift — at least two thirds of the fifteen items in section 1). In both systems, however, the taxonomic value of the dichotomy is intuitively satisfying.

7 Conclusions

The formation of lexical doublets is evidently a sub-system of another process or processes typically at work in a particular language or culturally bound group of languages. As we see from the analogies touched on above, a historical explanation plausible for one group may need to be stood on its head to apply to another. Rate of change, size of inventory and type of target words may differ not only between groups but within a group.

In what seems a plausible scenario for the Persian system here discussed, the superordinate process has been the assimilation of Arabic words in the feminine ending; this vocabulary (maximally, almost 2,000 items) included a disproportionately high number of desirable abstracta and culturally freighted words, and in its Arabic form was encumbered by a syntactically-determined alternation in pronunciation and orthography irrelevant to (and, indeed, inexpressible in) Persian. The alternates were -l-at ~ -l-āl, writable as -at ~ -a. The form -at was chosen for the majority of borrowings, which began as mots savants in the literature of bilingual Persian intellectuals. It so happens that, since Middle Persian postvocalic /l/ had become voiced by the time of the Arab invasion, terminal -l-āl was no longer in the phonotactic repertoire of Persian; this aberrant ending was thus marked from the outset as proper to a class of Fremdwörter.

The reflex -āl represented the pre-junctive pronunciation which in Arabic characterizes feminine-ending words in isolation or at the end, rather than in the middle, of a noun phrase — and was thus more frequently heard in names, glosses accompanying extensive definition, and other paraclitastic utterances. It had already been chosen for a critical mass of more tangible and imageable referents borrowed into spoken New Persian, and was assimilated into the written language as -ā. Its form is that of a vowel change, but the orthography reflects the spoken pronunciation (with lengthened e) precisely and correctly.

In this way an unwieldy chunk of useful Arabic vocabulary was sorted in accordance with Persian semantic and phonotactic principles. The formation of doublets, a special case of the same process of sorting and shift, further expanded the existing lean vocabulary. The bulk of this growing inventory, and the rules of the system, were passed on to speakers and writers of growingly adjacent languages, who further exploited them. Since little or no Arabic vocabulary is nowadays inducted into these languages, the whole process is barely productive, even in Persian; the number of shifted forms and doublets has probably stabilized, or is about to, closing a distinctive phase of Persian lexical innovation.