

ARABIC LANGUAGE V. ARABIC ELEMENTS IN PERSIAN

ARABIC LANGUAGE

v. ARABIC ELEMENTS IN PERSIAN

Since the Arab conquest of Iran in the seventh century and the subsequent conversion of a majority of the population to Islam, Arabic, as the language of contact, of the Muslim scripture and liturgy, and of a large volume of wide-ranging scholarly literature for more than a thousand years thereafter, has exercised a profound influence on the Persian language. Apart from the writing system, this influence is evident chiefly in the large Arabic vocabulary that has been incorporated into the Persian lexicon. The following will survey the topic under the rubrics of Lexical statistics; Phonology and orthography; Loanword classes; Grammatical elements; Semantics; History and evolution.

Lexical statistics.

A dictionary-based sample yields an inventory of approximately 8,000 Arabic loanwords in current use (Rāzi) or about forty percent of an everyday literary vocabulary of 20,000 words (not counting compounds and derivatives). In corpus-based inventories, the frequency of use of Arabic vocabulary per text will obviously vary with stylistic register, individual style, and topic of discourse, and can be seen to have risen and peaked over the course of time. Thus, a sample from the versified national epic, the *Šāh-nāma* (completed ca.



400/1010), yields an Arabic vocabulary of only 8.8 percent and a frequency of 2.4 percent (Moïnfar, esp. pp. 61-66); Firdowsi's younger contemporary 'Onşori, in his eulogies modeled on the Arabic *qaṣida*, yields ca. 32 percent and 17 percent respectively (see ARABIC (iii), p. 234). In a sample of Sufi verse from about the 14th century these proportions rise to 51.8 percent and 24.3 percent respectively (Utas, esp. pp. 75-102, 121ff.); and in the prose fiction of Bozorg Alavi from the 1950s they drop to 46.5 percent and 19.7 percent respectively (Koppe, pp. 590-93; see also Perry, 1991, pp. 203-205, and ARABIC (iii), pp. 234-35).

Since Arabic lexical morphology is highly systematic, certain prefixed and suffixed formatives of Arabic are salient in the Persian dictionary, as are certain assonant word patterns. Thus the letter *mim*, the initial of three highly productive Arabic prefixes, accounts for about 1,800 loanwords, or almost a quarter of the Arabic vocabulary in modern Persian; *alef*, a carrier of several prefixes, accounts for ca. 1,220 words, or one-seventh; and *tā'* provides ca. 815 words, or more than one-tenth. The overall Persian inventories under these letters are correspondingly inflated: about 13.5 percent of Persian vocabulary begins with *mim*, which is four times that of the average letter. Loanwords terminating in the Arabic feminine ending (either *-at* or *-a*) account for at least 1,500 items, or 18.75 percent (almost one-fifth) of the Arabic loanword inventory (Perry, 1991, p. 6 and Appendix). The Persian inventory ending in the productive suffix-group *-i* (originating in two distinct New Persian suffixes, and further augmented by Arabic relative adjectives with the suffix *-iyyun*) comprises more than 14 percent of current Persian vocabulary (Kešāni, Table 1). As for pattern assonance, the *m*- inventory of modern Persian contains at least 140 Arabic loanwords of the lexical pattern *mofā'ala* and 70 of *maf'ala*.

Phonology and orthography.

With a few exceptions as noted below, Arabic loanwords in Persian are written exactly as in Arabic. They were incorporated directly from Arabic by bilingual scholars who had no need to vernacularize them; doubtless the sanctity of Arabic script as the vehicle of the Qur'ān also militated against any alteration. A number of Arabic characters represent consonants alien to Persian, which are therefore assimilated to the closest Persian phonemes: thus *s*, *ṣ* and *ṣ* are all realized as /s/, *z*, *ḏ*, *ḏ* and *ḏ* as /z/, *t* and *ṭ* as /t/, *h* and *ḥ* as /h/ (the voiced aspirate).

In writing Persian and other non-Arabic words the default variants are *s*, *z*, *t*



and *h*. Exceptionally, *ṣ* is used to spell the Persian *ṣad* ‘hundred’; it was originally written as *sad* but later changed to avoid confusion with homographs—the noun *sadd* ‘dyke, dam’, according to the *Ġiyāt al-loḡa* (Dehḡodā, s.v. *ṣad*), though a more likely (as being more frequent) source of ambiguity would have been the verb *šod* ‘went, became, etc.’, since the distinguishing dots of *šin* were often omitted in early manuscripts. Anomalously, both *t* and *ṭ* have been used for the Persian epic hero Tahmāsb (and his Safavid namesakes). The *ṭ* in a few Persian place names, such as *Ṭus* and (formerly) *Ṭehrān*, preserve early records in Arabic geography books. Other accepted arabicizations of Persian words involve a phonetic change, notably *fārs*, *fārsi* for *pārs*, *pārsi* ‘Fārs (province), Persian’, and *fil* for *pil* ‘elephant’, though some writers have always preferred the variants in *p*.

The glottal stop of Arabic (written as *hamza*) is retained after a consonant, but in speech is generally realized before a consonant as a prolongation of the vowel, and between vowels as a glide or a bilabial fricative, though in careful enunciation it may be sounded as in Arabic (/sowāl/ or /so’āl/ for *so’āl* ‘question’). Final postvocalic *hamza* is not usually written or pronounced in Persian of today: *’olamā-ye Qom*, earlier *’olamā-e Qomm*; *ḡaḡdōā* ‘food’ (< Ar. *ḡidōā*). The peculiarly Arabic sound of *’ayn* is ignored in initial (and, colloquially, in final) position; it is realized between vowels as a glide or a glottal stop, and before a consonant as a prolongation of the vowel (/ba:d/ for *ba’d* ‘after’; in Afghan Persian, the quality of the vowel is also changed, as /bā:d/). The sounds of *qāf* (native to Arabic and Turkish, but not MPers.) and *ḡayn* (probably approximated in MPers.; see Pisowicz, pp. 135, 139-40) are pronounced alike in Standard Persian (initially as a voiced velar stop or affricate, elsewhere as a voiced velar fricative; cf. arabic (i), p. 230), but are distinguished in most other dialects, including Afghan and Tajik Persian. Arabic *w* is realized as labio-dental /v/ in Standard Persian, though in other dialects it may occur as a bilabial or semi-vowel. The other Arabic consonants have Persian counterparts or close approximations.

The three “short” vowels of Persian were equated with those of Arabic, and not represented in the orthography; the three “long” vowels were equated with those of Arabic, and represented by *alef*, *wāw* and *yā*’ as *matres lectionis*. Two other vowels of Middle and early New Persian, *ō* and *ē* (the so-called *majhul*, i.e., non-Arabic, vowels), were also represented (ambiguously, until they collapsed with *u* and *i* in Persian of western Iran) by *wāw* and *yā*’. Sounds of Persian that did not occur in Arabic (*p*, *č*, *ž*, *g*) were represented in the



Perso-Arabic script by letters representing similar sounds (*b, j, z, k*), and some time later were provided with the familiar diacritics.

Vowels in Arabic loanwords are subject to assimilation, dissimilation and syncope in certain environments, and to analogical changes (cf. ARABIC (i), pp. 230-31). Thus *nahār* x *nāhār* ‘lunch’ (one of very few such changes to be registered orthographically); *ṣadā* x *ṣedā* ‘sound’ (/a/ is raised in proximity to a sibilant); *ḥaraka(t)* x *ḥarekat* ‘movement’, but *šarika(t)* x *šerkat* ‘partnership’. *Ma’derat* ‘excuse’ and *ma’refat* ‘knowledge’, however, correspond to canonical forms in Arabic. The change *mosāfara(t)* x *mosāferat* ‘journey’ (/a/ is raised in an open penultimate syllable), which applies to the whole form class of about 140 such loans in Afghan and Tajik, as well as Standard, Persian, perhaps rests on morphological analogy more than phonetic law, i.e., by contamination with the corresponding (active) participial loanword, such as *mosāfer* ‘traveler’, *mobārez* ‘fighter’, *monāseb* ‘suitable’, etc. This kind of change—psychologically to be seen as an attempt to harmonize evident cognates on familiar (Indo-European) principles of suffixation instead of the alien non-concatenating morphology of Arabic—can clearly be seen in the Persian pronunciation of *šojā’at* ‘bravery’ (Arab. *šajā’a[t]*), by analogy with the borrowed adjective *šojā* ‘brave’.

If Arabic hardly influenced the phonetics of Persian, it had a noticeable effect on the phonotactics, in introducing a number of alien consonant clusters (especially word-final, as in *rabṭ*, *feqh*, ‘*adl*’; cf. ARABIC (iii), p. 234). Some dialects of Persian (and other languages endowed with these loanwords) deal with the problem of pronunciation by inserting an epenthetic vowel, as /húkəm/ for *ḥokm* or /qábəl/ for *qabl*. Standard Persian, in contrast, tends to de-emphasize or elide one of the two consonants, as /vaxt/ or /vax/ for *waqt* ‘time’ and /so:b/ for *ṣobḥ* (with compensatory vowel lengthening).

Loanword classes.

The following lists the principal identifiable classes of Arabic vocabulary incorporated into Persian, with some indications of how they fit into Persian structure and usage. (A convenient summary of the Arabic element in Persian, largely in tabular form, is to be found in Elwell-Sutton, pp. 157-67).

Nouns. With the exception of the feminine-ending loans (see below), Arabic nouns (and most other classes) are inducted into Persian in their bare stem form, without inflection or other modification. To this form may be juxtaposed



all appropriate Persian affixes and clitics: *ketāb-hā-i* ‘some books’; *bi-vaḡfā-i* ‘disloyalty’.

In a few nouns ending in *alef maqṣura* this syllable has assimilated via a spelling-pronunciation (*yā*’ as *-i*, as in *ma’ni* ‘meaning’, pronounced /ma:ni/), but is pronounced in the literary register as /ma’nā/ and written before *eẓāfa* as *alef*, followed by *yā*’: *ma’nā-ye ān* ‘the meaning of it’. In the case of *da’wā* ‘dispute, litigation’ and *da’wi* ‘claim, pretension’ the different pronunciation and orthography have been lexicalized as two distinct words.

Action nouns (*maṣḍar*) and other deverbal derivatives may form Persian verbs in one of two ways: by suffixation of the Persian past stem and infinitive, as *fahm-idan* ‘to understand’ (the original way of forming denominal verbs in Persian, cf. *nām-idan* ‘to name’); or by combining with a dummy verb such as *kardan* ‘to do, make’ or *šodan* ‘to become, be done’, as *jam’ kardan* ‘to gather’ (*jam*’ ‘collecting’), *qabul šodan* ‘to be taken on, accepted, to pass (examination)’ (*qabul* ‘acceptance’). The former (synthetic) strategy was favored in earlier Classical Persian, and is still productive in Tajik; the latter (analytic) is preferred in Standard Persian. The meaning may be refined by use of an auxiliary with some semantic weight: *qabul dāštan* ‘to agree, concur (in argument)’ (*dāštan* ‘to have, hold’; here metaphorically ‘to hold to be, consider as’).

Besides the varied, unpredictable forms of the *maṣḍar* of Theme I (the basic sense) of the Arabic verb, there are ten fixed morphological patterns (*qāleb*) representing systematic semantic extensions of the meaning of the verb which have been extensively borrowed into Persian and commonly form compound verbs of the above type. Thus from the trilateral root *ṢLḤ* ‘(being) right, fit, proper, harmonious’ are derived the following Arabic verbal nouns that also appear in Persian, often as verbs or verbal idioms: *ṣolḥ* ‘peace’, *ṣalāḥ* ‘honesty, propriety, fitness’, *ṣalāḥ dānestan* ‘to deem appropriate, see fit’, *maṣlaḥat* ‘interest, expediency’, *maṣlaḥat didan* ‘to deem prudent’, *eṣlāḥ kardan* ‘to improve, correct, edit; shave’, *moṣāleḥat* ‘reconciliation’, *eṣtelāḥ* and *moṣṭalaḥ* (pl. *-āt*) ‘(technical) term, idiom’. There are also the plural *maṣāleḥ* ‘benefits, interests’ (in Indo-Persian, and hence Hindi-Urdu, ‘materials, ingredients, spices’), the adjective (originally an Arabic active participle; also a personal name) *ṣāleḥ* ‘wholesome, beneficial’, the compounds *ṣalāḥ-kār* ‘charitable’ and *eṣlāḥ-nā-paḍir* ‘irremediable’. There are many such multiple root-cognates in the Persian lexicon, conditioning the educated reader by alliteration to the connection of a particular consonant combination with a certain semantic



field, even though he may not know Arabic as such.

Other fixed patterns identify nouns of place, as *madrassa* ‘school’ (place of teaching, cf. the cognate loan *dars* ‘lesson’); of instrument, as *meẓrāb* ‘plectrum, dulcimer hammer’ (cf. *zarbat* ‘blow, beat’); and of habitual activity or occupation, as *raqqāṣ* ‘dancer’ (cf. *raqṣ* ‘dance’). They express several sorts of adjectives (*šarīf* ‘noble’, *fa‘āl* ‘active’) and derive quality nouns from adjectives (*nejāsat* ‘impurity’, cf. *najes* ‘impure’). Several patterns, notably the elative and diminutive, do not normally appear as loanwords except as names (Akbar, Ḥosayn).

Adjectives. Apart from participles (see below), the largest class of morphologically salient Arabic adjectives in Persian comprises the derivatives with the *nesba* or relative suffix *-i* (< *-iyyun*), as *makki* ‘Meccan’, *šaḵṣi* ‘personal’. This suffix coincides in form and meaning with NPers. *-i* (< MPers. *-īk*), as *širāzi* ‘of Shiraz’, *kāki* ‘earthen, light brown’. The latter is highly productive, and may be added directly to any class of nouns, including assimilated Arabic loanwords: e.g., *tejārati* ‘commercial’, *šī‘i* ‘Shi‘ite’ (where the orthography shows that this is not an Arabic form). In many cases, however, it is not obvious whether an adjective in *-i* represents an integral Arabic borrowing or a Persian derivative (in, e.g., *akkāsi* ‘photographic’ the suffix is technically Persian, since the arabicate *akkās* ‘photographer’ was coined in Ottoman Turkish, whence it passed into Persian; neither word is used in Arabic). The coincidence also results in homographs such as *dudi* ‘smoky, smoked’ (Pers. *dud* ‘smoke’ + *-i*) and *dudi* ‘wormlike, peristaltic’ (< Ar. *dud* ‘worm’ + *-iyyun*).

Participles. There are eighteen Arabic participial patterns (active and passive) commonly occurring as Persian adjectives and/or nouns (see Elwell-Sutton, pp. 162-63). Thus from Theme I of the verb ‘to know’ (the *maṣḍar* of which is the loanword *‘elm* ‘knowledge, science’) comes the active participle *‘ālem* ‘knowing, learned; sage, scholar’, and the passive *ma‘lum* ‘known’, pl. *ma‘lumāt* ‘data’; from Theme II, *mo‘allem* ‘teacher’ (active). A striking systematic function of many participles is in correlating with their cognate verbal nouns to form grammatically complementary verbal idioms, thus: *enteẓār dāštan* (lit., have expectation) and *montaẓer budan* (lit., be expecting) ‘to expect, wait’; *taṣmīm gereftan* (lit., take determination) and *moṣammam šodan* (lit., become determined) ‘to determine, decide’.

Adverbs. A few dozen Arabic adverbs originating in the *tanwin* accusative



comprise the only morphologically unique class of adverbs in Persian, e.g., *rasman* ‘officially’, *wāqe‘an* ‘really, actually’. These retain the Arabic orthography of a final alef with double *fatha*; a few with the feminine ending do not end in *alef* in Arabic, but may do so in Persian by accepted solecism (e.g., in *nesbatan* ‘relatively’). This characteristic ending has become productive, forming adverbs even from native words: *jānan* ‘wholeheartedly’, *nāčāran* ‘willy-nilly, of necessity’. The very common *hālā* ‘now’ is of this class, though assimilated via a spelling-pronunciation (in Afghan Persian, further assimilated as /ālē/).

Pseudo-loans. The degree to which not only individual loanwords, but also their characteristic patterns, entered Persian consciousness is shown in a number of common Persian words coined on an Arabic morphological pattern from a native Persian or other lexical base: thus *kaffāš* ‘cobbler’ (< Pers. *kafš* ‘shoe’), *nezākat* ‘daintiness’ (< Pers. *nāzok* ‘dainty’), Tajik *partiīaviīat* ‘(Communist) Party loyalty’ (< Rus. *Partiia* ‘party’, on the analogy of *partiīnost*). Arabicized forms of Persian words borrowed into Arabic were also accepted back, as *fehrest* ‘list, register’, originally MPers. *pahrist*, and *fārs*, *fārsi* (see above), which may also be regarded as a blend of MPers. *pārsīk* and Ar. *fārisiyyun*.

Loanwords with the feminine ending. The grammatically feminine marker in Arabic is realized phonetically as either /-at/ (in pre-juncture position) or /-a/ (pausal form), according to the contextual syntax of Arabic, but written with a single graph (the *tā’ marbuṭa*). The syntactically determined variation in Arabic (though in context it may initially have suggested a model) was irrelevant to Persian, where these loans needed to be lexicalized in stable forms: accordingly, some were written with regular final *t* (e.g., *hekmat* ‘wisdom, philosophy’) and others with non-linking final *h* (as in *kerqa* ‘rag, dervish’s cloak’). In current Persian of Iran there are at least 810 ending in *-at*, and 640 in *-a* (realized in Standard Persian as /-e/), including some 80 items lexicalized with both endings (i.e., 40 pairs of doublets). As the only class of loanwords that have been systematically sorted orthographically, an analysis of the rationales behind this dichotomy affords some insight into the process of loanword incorporation from Arabic into Persian.

Distribution between *-at* and *-a* in the modern inventory appears to be determined primarily by semantic features, and additionally by factors of syntactic and stylistic environment or historical evolution of the words (Perry, 1991, pp. 195-224). Thus, nouns with more abstract and intangible, or less



imageable and countable, referents tend to end in *-at*: *rokṣat* ‘permission, leave’, *koṣunat* ‘asperity, roughness’, *mojānebat* ‘avoidance, non-intervention’, *mawhebat* ‘(figurative) gift, talent’; nouns with more concrete, tangible, imageable and countable referents (more likely to appear in the plural) tend to end in *-a*: *noska* ‘text, prescription’, *waṭiqa* ‘bond, security’ (document), *molāḥeẓa* ‘note, remark’, *maḥalla* ‘place, neighborhood’. There are of course exceptions, and *maṣdar* forms (by definition abstract, etc. in their basic meanings) seem to be more arbitrarily apportioned; even of these, however, the ones ending in *-a* tend to form common compound verbs in Persian and have also evolved count-noun referents (*estefāda kardan* ‘to use’, *estefāda-hā* ‘uses’; *eṣāra kardan* ‘to point out, indicate’, *eṣāra-hā* ‘indications’ (cf. the archaic *eṣārat*, still to be found as an elegant variant of *eṣāra* ‘indicating, reference’, but only in the singular as a verbal abstract). These processes are even more apparent in the doublets: *qowwat* ‘strength, power’ (general, intangible), vs. *qowwa* ‘(military) force, (industrial) energy, (physiological or mental) faculty’ (pl. *qowwa-hā*, *qowā*); *erādat* ‘wish, goodwill’, *erāda* ‘resolution, edict’; *resālat* ‘mission’ vs. *resāla* ‘message, letter, dissertation’ (Perry, 1995). The loss of final *t* (see below, under *History and evolution*) often corresponds additionally to a change of register, from literary to vernacular: thus Persian *ḥekāyat* ‘(literary) anecdote’ has remained more a literary word (in comparison with *qeṣṣa* ‘tale, story’), whereas in modern Tajik and Turkish it has dropped final *t* orthographically (i.e., an existing vernacular form in *-a* has been recognized in the written language) as *hikoya* (Taj.)/ *hikāye* (Tk.) ‘tale, story’. Thus the system of binary sorting in Persian was passed on to Turkish, Urdu and other languages of central, south and southwest Asia together with the Arabic loans that they incorporated via Persian, and was selectively expanded or modified.

Change of category. A loanword may also signal its assimilation into the vernacular by an expansion or shift of grammatical categories. Several nouns of quality of Arabic origin are now used primarily as adjectives in Persian, e.g., *kalwat* ‘private, quiet’, *rāḥat* ‘easy, comfortable’, *salāmat* ‘safe, well’; the shift was presumably achieved by way of a reanalysis of the word as predicate (in *rāḥat nist* ‘this is not (my idea of) comfort’ x ‘not comfortable’). Most such words may now be used attributively (an exception is *šohra* ‘famous, a by-word’, a doublet of *šohrat* ‘fame, surname’). Such words may derive a new quality noun by suffixing *-i*: *salāmāti* ‘health’, etc. Other nouns have become adverbs: *xolāṣa* ‘gist; in short’, *xāṣṣa* ‘specially’ (< ‘peculiar property’).



Grammatical elements.

Arabic plurals may be used instead of Persian plurals (*ketāb-hā* or *kotob* ‘books’, *mo’allem-ān*, *mo’allem-hā*, or *mo’allem-in* ‘teachers’). The choice is usually stylistic, but some plural loans have been lexicalized with a singular meaning (*arbāb* ‘landlord, boss’; the singular *rabb* ‘Lord’ is used in Persian only with reference to God). In other cases the choice of plural is lexicalized, each form denoting a part of the semantic range of the singular, e.g., *ṣāḥeb-ān* ‘owners’, *ṣaḥāba* ‘the Companions (of the Prophet)’, *aṣḥāb-e X* ‘people characterized by X’; *ḥarf-hā* ‘(spoken) words, utterance’, *ḥoruf* ‘letters (of the alphabet), written characters’. Arabic “broken plurals” have occasionally been applied to Persian and other non-Arabic nouns; some such usages were ephemeral (*dahāqin* ‘landowners’ < Pers. *dehqān*), others retain currency: *banāder* ‘the lower Gulf littoral’ < Pers. *bandar* ‘harbor’ (cf. ARABIC (i), p. 230).

The few nouns in which the Arabic definite article *al-* is incorporated in Persian function not as nouns but as interjections or adverbs: *al-’amān* ‘mercy!’, *al-wedā* ‘farewell’, *al-’ān* ‘now’, *al-batta* ‘of course’. Arabic nominal collocations (adverbial and noun phrases), frozen and lexicalized, play a larger role: *be’l-’aks* ‘vice-versa’ (also Persianized as *bar ’aks*), *jadid al-worud* ‘newly-arrived’ (cf. ARABIC (i), p. 230). The class includes many titles and personal names (esp. servile compounds of the type *’abd(-al)-* ‘servant of —’). The feminine ending in collocations is generally written as final *h* if it occurs in the final constituent (as *fawqal-’āda* ‘extraordinary’), and with the *tā’ marbuṭa*, as in Arabic, in the preceding constituent (*dōu kamsat(-e) aẓlā* ‘pentagon’); in more familiar collocations, such as *āyat-Allāh* ‘ayatollah’, it is generally written as final *t*. In recent centuries, macaronic collocations such as *ḥasab al-farmān* ‘in accordance with decree’ (modeled on Arabic *ḥasab al-’amr*) were manufactured by self-important bureaucrats. Verb phrase collocations (interpreted as reduced relative clauses) also serve as adjectives: *lā-yanfakk* ‘inseparable’ (Ar. ‘it (etc.) is not detached’) or nouns: *mā-jarā* ‘adventure, affair’ (Ar. ‘what occurred’).

“Pseudo-concord,” the analogical addition of a grammatically feminine ending to an adjective of Arabic origin when modifying a Persian noun with a female or plural referent, originated by analogy with borrowed collocations of the type (*al-’omur* (*al-’)**kāreja* ‘foreign affairs’ (later Persianized as *onur-e xāreja*). The device was introduced in the later 12th century, but survives only in a few stylized phrases, as *kānom-e moḥtarama* ‘Dear Madam’.



Semantics.

Studies of the Arabic component of specific semantic and experiential fields are as yet few and limited. In terms of psycholinguistic categories, one's impression is that Arabic loans in Persian comprise a greater proportion of abstract, intangible, less imageable and less countable referents than of entities and other tangible, more imageable and countable referents. This appears to be confirmed by a survey of the "Sachgruppen" or experiential fields into which Koppe (after Dornseiff) sorts the Arabic vocabulary of a sample of modern Persian fiction: out of a total of 1,346 loanwords, those referring to the more vague abstracta, such as sentiment, volition and ethics, total 479 (ca. 36 percent); those referring to intangibilia with strong cultural, perceptual, social or other relations (e.g., 'point'), plus tangibilia that are systems rather than entities (e.g., 'crowd') total 731 (ca. 54 percent); and those referring to simple entities (e.g., 'rock') total 136 (ca. 10 percent) (cf. Perry, 1991, pp. 206-208). In another such experiment, comparing a random sample of Arabic loans in four languages, the vocabulary to do with material culture in Spanish was 52 percent of the Arabic loan inventory, while in Persian the total was only 14 percent; the Arabic vocabulary of general intellectual life was 8 percent in Spanish, 24 percent in Persian (J. R. Perry, "Arabic loan vocabulary in Persian, Turkish, Urdu, etc: Comparative indices," paper delivered at the 201st Annual Meeting of the American Oriental Society, University of California, Berkeley, March 1991).

Many Arabic loans have emerged from their sojourn in Persian poetry or scholarship or vernacular idiom enriched in meaning, often with an extra identity in Turkish, Urdu, or the languages beyond. One such is *ṣoḥbat*, a verbal noun meaning 'comradeship, company' in Arabic; as a staple of Persian lyrical and mystical verse (in phrases such as *ṣoḥbat-e yār* 'company of the friend'), it left much to the imagination, and has accordingly specialized in prose usage in two directions. In modern Standard Persian it denotes 'speech, conversation' (*ṣoḥbat kardan* 'to talk'), while in Indo-Persian it came to connote sexual dalliance (Urdu *ṣoḥbat karnā* 'to cohabit, have sex with'; cf. the evolution of English 'intercourse' from social to sexual).

The wholesale importation of verbal nouns of the same lexical pattern has created high-profile semantic sub-classes that are more noticeable in Persian than they are in Arabic. The large *mofā'ala* pattern, for instance, generally encodes the notion of reciprocity, which is realized lexically in three archetypal human activities: love, war and trade (or sex, conflict and



business). Thus Persian *moqārebat* ‘sexual congress’, *mojādelā* ‘dispute’, *mo‘āmela* ‘operation, deal’ each belong to a cluster of assonant near-synonyms which collectively define the greater part of mankind’s social pursuits.

History and evolution.

We have no way of documenting the first two centuries of the influence of Arabic on Persian, i.e., before about the middle of the 9th century, to which the first extant examples of Persian poetry are attributed (Lazard). Persian was long familiar with Semitic languages and their writing systems: Old Persian used a simple and efficient syllabary adapted from Babylonian cuneiform, and Middle Persian a rather less efficient adaptation of Aramaic script, with literacy in each case confined to a small class of priests and scribes. Arabic script was much better adapted to Persian, and the orthographic rigidity of Arabic perhaps suggested a matching uniformity in Persian. The nature of Islam encouraged a rapid social as well as geographical expansion of literacy in Arabic, so it is quite possible that newly literate converts, or at least the children of converts, were already writing Persian in Arabic characters in the second generation of Iranian Islam. Even before this, they were learning to speak Arabic, and many became bilingual. Persian preserves traces of this “vernacular stage” in a few early Arabic borrowings that were phonetically assimilated to Persian, and have survived subsequent orthographic normalization: e.g., *mosalmān* ‘Muslim’ (by metathesis, and perhaps modification of a plural, from Ar. *moslem*); the onomastic *bu* (< *abu*), *mir* (< *amir*) ‘commander’ and its compounds *mir-āb*, *mir-ākōr*, *mir-zā*, which parallel a tendency to apheresis in native Persian words at this time, as *(a)yār*, *(a)bā*, *(a)bar*, *(a)nāhid*, etc. Thereafter, the bulk of Arabic loanwords entered Persian as learned words in the writings of bilingual poets and scholars, most of them trickling down into spoken usage in due course (Telegdi).

Clearly it was not a paucity of technical and intellectual terminology in Middle Persian that necessitated the massive influx of Arabic. Pre-Islamic Persian enjoyed a sophisticated system of lexical derivation and compounding (MacKenzie) and a wealth of culturally specialized terms, many of which had already been transmitted to Arabic (cf. ARABIC (ii), pp. 231-33). Some of these soon came back into Persian in Arabicized form, to replace or supplement the Persian etymon (e.g., *handasa* ‘geometry, engineering’, from Pers. *andāza* ‘measure’)—showing that prestige was a factor in reversing the current. Arabic borrowings into Persian have on the whole supplemented, rather than replaced, non-specialized vocabulary, providing a wealth of synonyms, such as



mariž, *bimār* ‘sick’ and *mo’assasa*, *bonyād* ‘foundation, institute’. Nor is the field of (Islamic) religion dominated by Arabic loanwords: scores of Persian words, from *ākund* ‘cleric’ to *zendīq* ‘heretic’ (the latter in Arabicized form), are Persian, including the everyday terms for God, prophet, prayer, prayer-leader, fasting, angel, creation, creator, heaven, hell, soul, sin, worship, repent, forgive, etc. None of these facts need surprise us. The process of conversion depends for its early success on comprehension, achieved by translation into, analogy with, and use of the language of the target population. But the literature of the old religion, together with any terms without analogy in Islam (e.g., *sōšyans* ‘savior’), was swept into oblivion by the scale and rate of Islamization in Iran. Paradoxically, Iranian intellectuals then played a dynamic and leading role for over three centuries in the development of Islamic civilization by adopting Arabic as their written medium; when they came to write Persian (in part by way of translations from Arabic classics), it was easier to plug the familiar Arabic vocabulary into their native syntax than to transcribe archaic or ideologically inappropriate relics (though Ghazali and Ebn Sinā did consciously resurrect Persian vocabulary in some of their popular treatises). Since the same scholars and bureaucrats were often poets and patrons of poetry, Arabic increasingly made its way into Persian verse, and the Arabic prosodic system of *‘aruż* was adapted to scan it.

The feminine-ending loans provide some indication of the ways in which Arabic vocabulary was assimilated into Persian. Those adopted in the form *-a* were often morphologically assimilated with the large class of native substantives in *-a* (< MPers. *-ag*, such as *dāna*, *tiša*, *bačča*, *barnāma* and the active and passive participles), a class which at the time of the Arab conquest had dropped, or was in the process of dropping, final *-g* (still evident in earlier borrowings into Arabic such as *dānaq* and *barnāmaj*): by analogy with the Persian class, this consonant was supplied to many of the Arabic loans before suffixes, as in *kebragān* ‘experts’, *bi-saliqagi* ‘lack of taste’. Conversely, those loans adopted in *-at* stood out as foreign words, since by this time virtually all instances of final *t* in Persian had been voiced (*t > d*). In the course of the next several centuries, hundreds of the *-at* class shifted to the *-a* class, some leaving behind traces as doublets in *-at*. In general, the resulting *-a* words are semantically more specialized (cf. *qowwat*/*qowwa* above) and/or more firmly established in the vernacular (cf. *ešārat*/*ešāra*). This shift appears to have peaked about the 13th century, by which time the majority of the Arabic loanwords that are in general use today had been incorporated. By then, too, the stratagem for coining verbs from Arabic had changed from the suffixation



of *-idan* to the juxtaposition of a Persian auxiliary verb. About the same time, a new stratum of *mašdars* was incorporated; those from feminine-ending patterns were uniformly assimilated in *-a* (Perry, 1991, pp. 13, 191, 219).

Salient among the earliest loanword classes (coined in Arabic during the philosophical-scientific heyday of Islam in the 9th-10th centuries) were the *nesba* subset of the feminine substantives, incorporated as *-iyat/-iya*, e.g., *ensāniyat* ‘humanity’, *zōjājiya* ‘crystalline lens’. During the 19th century, a wave of Arabic (and artificial Arabicate) neologisms, many calqued on French and originating in Ottoman Turkish, supplemented the technical and legal-administrative lexicon of Persian; these, too, included a large *nesba*-noun component, such as *melliyat* ‘nationalism’, *aktāriyat* ‘majority’, *eḥzāriya* ‘subpoena’, *etfā’iya* ‘fire service’ (cf. Faršidvard, pp. 61-63). The Persian vocabulary ending in *-iyat* still comprises up to 200 words (Kešāni, Table 1; Perry, 1991, p. 23), that in *-iya* about fifty. With the language purism movement of the 1930s–1940s, grammatical Arabisms were decried and Arabic vocabulary was targeted for replacement by Persian neologisms. Though this reform was not as drastically implemented as in Turkey, many of the more recent technical terms were replaced, and officially sanctioned lexical policy ever since has preferred to coin Persian terms or tolerate European loanwords (see *FARHANGESTĀN*; Perry, 1985).

The Islamic Revolution of 1979 does not appear fundamentally to have affected these trends. A few ideologically inspired Arabisms have been introduced, such as *mostaẓ’af* (pl. *-in*) ‘dispossessed, underprivileged’; but both technical and everyday vocabulary is still being expanded primarily by appeal to native Persian words and morphs (supplemented in the spoken language by borrowings from English). Writers in Afghanistan and Tajikistan since the 1980s are likewise giving prominence to native lexical funds, frequently inspired by Iranian Persian examples. Arabic is no longer a living lexical source for Persian.



BIBLIOGRAPHY

Asya Asbaghi, *Die semantische Entwicklung arabischer Wörter im Persischen*, Wiesbaden, 1987.

L. P. Elwell-Sutton, *Elementary Persian Grammar*, Cambridge, 1963.

Kosrow Faršidvard, *‘Arabi dar fārsi*, Tehran, 1348 Š./1969.

Mohammad Ali Jazayery, “The Arabic Element in Persian Grammar: A Preliminary Report,” *Iran* 8, 1970, pp. 115-24.

Kosrow Kešāni, *Farhang-e fārsi-e zānsu/ Dictionnaire inverse de la langue persane*, Tehran, 1372 Š./1993.

Reiner Koppe, “Statistik und Semantik der arabischen Lehnwörter in der Sprache ‘Alawis,” *Wissenschaftliche Zeitschrift der Humboldt-Universität zu Berlin* 9, 1959-60, pp. 585-619.

Gilbert Lazard, *Les premiers poètes persans*, 2 vols., Paris/Tehran, 1964.

D. N. MacKenzie, “Pahlavi compound abstracts,” in *Iranica Varia: papers in honor of Professor Ehsan Yarshater*. *Acta Iranica* 30, Liège, 1990, pp. 124-30.

Mohammad Dj. Moīnfar, *Le vocabulaire arabe dans le Livre des Rois de Firdausi*, Wiesbaden, 1970.

John R. Perry, “Language Reform in Turkey and Iran,” *IJMES* 17, 1985, pp. 295-311.

Idem, *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending*, Costa Mesa, 1991.

Idem, “Lexical doublets as a derivational device in Persian: The Arabic feminine ending,” *Acta Orient. Hung.* XLVIII, 1995, pp. 127-53.

Andrzej Pisowicz, *Origins of the New and Middle Persian Phonological Systems*, Cracow, 1985.

Farida Rāzi, *Farhang-e ‘arabi dar fārsi-e mo‘āšer*, Tehran, 1366 Š./1987.



Zsigmond Telegdi, “Remarques sur les emprunts arabes en persan,” *Acta Linguistica* (Budapest) 23, fasc. 1-2, 1973, pp. 51-58.

Bo Utas, *A Persian Sufi Poem: Vocabulary and Terminology*, London and Malmö, 1977.