Creating Standards

Studies in Manuscript Cultures

___

Edited by
Michael Friedrich
Harunaga Isaacson
Jörg B. Quenzer

Volume 16

Creating Standards

Interactions with Arabic Script in 12 Manuscript Cultures

Edited by
Dmitry Bondarev
Alessandro Gori
Lameen Souag

DE GRUYTER

Contents

Paola Orsatti

Persian Language in Arabic Script: The Formation of the Orthographic Standard and the Different Graphic Traditions of Iran in the First Centuries of the Islamic Era

Abstract: This paper offers a critical review of the orthography of the most ancient original New Persian texts written in the Arabic (11th century), as well as in other scripts (Hebrew, Syriac and Manichaean), to discover any indirect evidence about the beginnings and the formative period of Arabo-Persian orthography. On the basis of the New Persian original documents in non-Arabic scripts here examined, dating to an earlier period than documents in the Arabic script, we can tentatively date and localise the beginning of the Arabo-Persian orthographic influence on the other written traditions of Iran: northeastern Iran, end of the 9th– beginning of the 10th centuries. By then, Arabo-Persian orthography appears as already fixed. Though it cannot be excluded that some scattered and unsystematic attempts at adaptation of the Arabic script to Persian were accomplished here and there in different places of Iran, the hypothesis of a multi-centric origin of the Arabo-Persian orthographic canon seems less probable in the light of our documentation.

1 Introduction

The aim of this study is to carry out a critical review of the orthography of the most ancient original New Persian texts written in the Arabic, as well as in other scripts (Hebrew, Syriac and Manichaean), to discover any indirect evidence about the beginnings and the formative period of Arabo-Persian orthography.[1] By New Persian, or simply Persian, I am referring, in a broad sense, to the Persian language of the Islamic period.

Though agreeing with Giorgio Banti's criticism expressed during the Hamburg conference in 2013 about the use of the term 'standard' for pre-modern times, I retain it to underline the fact that New Persian orthography in Arabic script seems to follow a well-established norm from the time of its most ancient original attestations.

1 A general study of the most ancient New Persian documents in alphabets different from Arabic is given by Orsatti (2007, 107–172, with reference to previous studies), and – from a historico-cultural perspective – by Orsatti (2008). For the layout of the page in Islamic manuscripts

2 The adaption of the Arabic script to New Persian

As is well-known, towards the mid-7th century western Iran, the heart of the Sasanid empire, was conquered by the Arabs. This historical event had many important consequences: most Iranians, more or less gradually, gave up their ancient Zoroastrian religion and converted to Islam; and – given the strong link existing between writing and religion – they abandoned the ancient Pahlavi script, in which what we call Middle Persian was written, and adopted the Arabic alphabet to write their language.

The first original written records of Persian in Arabic script (I will speak of 'Arabo-Persian'),[2] both literary and not, go back to the 11th century: therefore they are not really so ancient. We do not know when, where and for what purposes (administration, literature, private documents etc.) the Arabic script was first adapted to write the Persian language. And we do not know in which way and through what stages Arabo-Persian orthography was fixed. From the extant documentation we get the impression that an orthographical norm was well-established from the very beginning. On the other hand, at least for literary texts which have undergone a long manuscript transmission, we do not know to what extent the normalising intervention of copyists may have contributed to suggest the existence of an early orthographical norm.

Historical sources can help find out when and for what purposes New Persian emerged as a written language, prevailing over Middle Persian, written in the Pahlavi script, and over Arabic, the language of religion and science. In the first centuries of Islam the prestige of Middle Persian was still strong, especially in south-western Iran. From a piece of information provided by the Arab historian Balāḏurī (9th century) we know that, in western Iran, Middle Persian in Pahlavi script was used for administration until the end of the 7th or the beginning of the 8th century, and in eastern Iran even longer, before being substituted by Arabic.[3] For administration, therefore, Middle Persian was replaced directly by Arabic, and only later was New Persian used as the language of administration. The same can be said of epigraphy (Bivar 1986): for inscriptions, Middle Persian in Pahlavi

---

and its relationship to the text copied, manuscript typology and destination, I again refer to some observations in Orsatti (1993, 281–282, 319–323). On the historical orthography of New Persian texts in Arabic scripts, apart from scattered remarks on the orthography of manuscripts in the introductions to the editions of texts, see Šīrvānī 1974 and Hashabeiky 2005, 79–85 (remarks based on an early study by Jalāl Matīnī).

2 The term 'Muslim Persian' is less appropriate because different minorities in Islamic Iran, especially Christians, also commonly used the Arabic alphabet (see below).

3 See Xānlarī 1986, I, 307-314.

script was gradually replaced by Arabic, written in the so-called Kufic script, one of the most ancient inscriptions in New Persian being the Kufic inscription in Persian verses in the palace of Masʿūd III at Ghazna (Afghanistan), dated 505/1111 (published by Bombaci 1966).

What is sure is that, when in the 9th century, during the period often referred to as the 'Iranian renaissance', New Persian lyric poetry of the literary type – i.e. patterned after Arabic poetry – came into being in the Persian courts of eastern Iran, it could not be written down other than in the Arabic script. Indeed, New Persian lyric poetry arose as an experiment in the Persian courts, by then independent or semi-independent from the caliphate of Baghdad; an experiment that consisted of substituting Persian for Arabic within the pattern of Arabic poetry.[4] We can suppose that the establishment of New Persian orthography was a part of this experiment.

The anonymous author of a local chronicle, the *Tārīkh-i Sīstān* 'History of Sistan' (11th century, with later additions), recounts an interesting story about what he presents as the first Persian poem of the Islamic era, which – he says – was told for Yaʿqūb b. Lay_t_ of the Saffarid dynasty of Sistan (south-eastern Iran). In 251/865 Yaʿqūb gained an important victory over one of his enemies; on this occasion poets recited poems in Arabic in front of him, extolling his victory. Yaʿqūb, who was a man of arms and not a learned person, said then the famous phrase: 'Why should a poem be composed which I cannot understand?' Muḥammad b. Waṣīf, his correspondence secretary, was in the audience; he had the idea of composing a poem in Persian, of which six lines, full however of Arabic words and expressions, are quoted (*Tārīkh-e Sistān*: 166–167).

The story of Yaʿqūb the Saffarid and his secretary and court poet Muḥammad b. Waṣīf is an important piece of information about when, where and in which way the new Arabicised poetry in the 'vulgar tongue' of the time, that is New Persian, came into being: during the 9th century, in the Persian courts which flourished in eastern and north-eastern Iran.[5] Prose began later: the first prose texts in literary New Persian, in Arabic script, go back to the mid-10th century – the most ancient dated New Persian manuscripts of literary works in Arabic script, however, belong to the middle of the following century.

---

4 On these themes cf. in particular Bausani 1960, 307–311 and the studies by Lazard 1971 and Lazard 1975.

5 For an evaluation of this and similar accounts concerning the birth of New Persian literary poetry, cf. in particular Lazard 1975, 607–610.

2.1 Persian in Arabic script: general characters

The following table shows the most important characteristics of the adaptation of the Arabic alphabet to the Persian language (mid-11th – mid-13th century):

A. Conservation of the 'only Arabic' letters, i.e. letters found only in Arabic loanwords:

Arabic letters	Arabic value	Phonetic value in Persian
ث ‹ṯ›	θ	s
Arabic ذ ‹ḏ›	δ	z (perhaps pronounced δ in Early New Persian)
ح ‹ḥ›	voiceless pharyngeal fricative ħ	h
ص ‹ṣ›	pharyngealised s	s
ض ‹ḍ›	pharyngealised δ or d	z
ط ‹ṭ›	pharyngealised t	t
ظ ‹ẓ›	pharyngealised z or δ	z
ق ‹q›	q	introduction of the new phoneme /q/, in later times merged into /ɣ/ ‹ġ›, but retained till now in Afghani Persian
ع ‹'›	voiced pharyngeal fricative ʕ	not pronounced or pronounced as a glottal stop (merging with the Arabic *hamza* [ʔ])

B. Introduction of new letters, mainly by addition of diacritical points

Letters	Phonetic value	Commentary
پ	p	already introduced, as the three other 'Persian' letters (see below), in the most ancient manuscripts, especially of fine workmanship, but normally seldom used
چ	č	see above
ژ	ž	see above
گ ,ڲ	g	see above
ڤ ‹ḇ›	β	spirant allophone of /b/. This letter is used only in the most ancient manuscripts
'Persian' ذ ‹ḏ›	δ	spirant allophone of post-vocalic /d/, used in manuscripts until the middle of the 13th century

C. One digraph, which represents the phoneme /xw/ of Early New Persian (this being the only case of historic orthography in the New Persian writing system, conserved till now):

خو	xw	later merged with /x/ with dropping of the labial articulation of the ancient phoneme

One of the main differences between the Arabo-Persian writing system and the Arabic one is the regular spelling of final short -*a*, which in Arabic is not written, by means of <h>. This is well explained by a passage from the *Muʾǧam* by Šams-i Qays (first half of the 13th century), where the great Persian philologist says that: 'In Persian orthography (*dar xaṭṭ-i pārsī*), whenever a word ends in -*a*, you have to add a *hā* to it [...] These *hā*s in the Arabic language are clearly pronounced [i.e. in their value as *h*] [...] Instead in Persian (*pārsī*) they are by no means pronounced except, by poetic license, if they are in rhyme position; (only) in this case they are counted as a quiescent letter and are feebly articulated' (Šams-i Qays 1981, 243–244).

2.2 Written and not-written morphemes

An important characteristic feature of Arabo-Persian orthography is the way two important and very frequent morphemes of the Persian language are represented: the *iḍāfa*, that is the particle -*i* of the *status constructus*, and the coordinative conjunction *u* (English 'and'). Both consisted in a (probably) short vowel, and therefore – according to the Arabic writing system – should not have been written; but, while the first one – in contrast with the pre-Islamic graphic traditions of Iran – is not represented in the Arabo-Persian orthography, the second one is written <w> as a stand-alone word.

The *iḍāfa* particle is an enclitic -*i* vowel, originally long but probably shortened from the beginning of the New Persian period, which follows the head of a nominal phrase: *kitāb-i muʿallim* 'the book of the teacher', *kitāb-i buzurg* 'the big book'. In Arabo-Persian orthography this short -*i* is not written, except after words ending in a vowel: only in this case is the *iḍāfa* represented as <y> (the letter *yā*) attached to the preceding word – what is represented in such cases is the glide *y* which develops between the final vowel of the word and the *iḍāfa* vowel: *banda-yi tu* 'your slave'. In the most ancient Persian manuscripts this orthographic rule is always observed, apart from very rare cases in which the *iḍāfa* is represented by <y> also after a word ending in a consonant (Lazard 1963, 200, § 162). A couple of such cases, pointed out by Minorski (1942, 188), seem to be present also in one of

the most ancient non-literary Arabo-Persian documents, the deed for the sale of land found in Khotan (Central Asia) (line 2), dated 501/1107. Though it is generally admitted that in New Persian the vowel representing the *iḍāfa* had already been shortened, these occasional spellings, as well as the fact that in poetry the *iḍāfa* can count as a short or as a long vowel, seem to point to the presence – in Early New Persian – of a long variant of the *iḍāfa* (Meier 1981, 131–132).[6]

An ancient spelling of the *iḍāfa* particle by means of an isolated or proclitic *alif*, attested in the quotation of Persian phrases in works of some Iranian authors (Abū Nuwās, Hamza al-Iṣbahānī, the author of the *Taʾrīx-i Qumm*) of the 9th–10th centuries writing in Arabic, may represent an ancient spelling, soon fallen into disuse, preceding the establishment of Arabo-Persian orthography (Henning 1958, 88–89); this is even more relevant in the light of the fact that the authors in which this spelling is to be found are from central and western Iran.

A group of fifteen Judaeo-Persian tombstones edited by Gnoli (1964) found in the Ghūr (Afghanistan) and dated – according to the Seleucid era – to the years between the second half of the 12th till the beginning of the 13th century CE, show the gradual falling into disuse of the writing of the *iḍāfa*, perhaps due to the influx of coeval Arabo-Persian orthography, and possibly also as a consequence of the disappearance of the long *iḍāfa*. In the initial formula *wafāt-i ʿadan* 'Edenic death', the *iḍāfa* is written as <-y> joined to the preceding word in tombstones IV and VIII (dated 1484 Seleucid/1172 and 1502/1190 respectively), and is not written in tombstones III (also dated 1484 /1172), V, VI, VII, IX, X, XI (this being the most recent one, dated 1526/1214), and XIV.

Besides the *iḍāfa* particle – which, being a (short) vowel, very soon disappeared from orthography though remaining well alive in the language – other morphemes are not represented in the Arabo-Persian orthography. One of them, consisting in an originally long *ō* afterwards shortened, is the directional Middle Persian preposition *ō* 'towards, to', well-known for the New Persian linguistic period from the most ancient Judaeo-Persian documents originating from south-western Iran (Lazard 2009; see below, Group C). Its disappearance from Arabo-Persian orthography was perhaps due to the fact that this morpheme was considered too dialectal, colloquial or archaic to be allowed into the writing; or, rather, its disappearance may well correspond to its gradual falling into disuse from the north-eastern variety of Persian which was at the basis of literary New

---

6 See also an inverse spelling in a New Persian fragment in Manichaean script, pointed out by Provasi (2011, 162, 164): in fragment M 595a+, *dāng*, 'a small coin', followed by *iḍāfa* is written <d'ngy<u>h</u>>, with the abstract suffix -*ī* (spelled -yh) representing a probably long *iḍāfa*.

Persian, the so-called *darī* or *pārsī-yi darī*.[7] Indeed, in the extant New Persian documentation in Manichaean script, mainly originating from Samarkand and eastern Iran between the 10th and 11th centuries, there is no trace of it; and it is rarely attested in the ancient Judaeo-Persian texts from north-eastern Iran (Paul 2003, 183; Shaked 2009, 452). It is possible therefore that the preposition *ō*, shortened in New Persian and not represented in Arabo-Persian orthography, was gradually replaced in common usage by other more fortunate prepositions, as the polysemous *ba* 'in, to, on, by, with, for'; and that it was then expunged from literary New Persian, though surviving in many Persian dialects especially in the western part of the Iranian plateau (Browne 1895; Filippone 2011, 198). However, this preposition seems to have been transmitted as a legacy to the common language based on literary New Persian. An indirect trace of it – no longer written nor pronounced – can be detected in a number of cases, typical of the contemporary spoken language (ex. *miram xune* 'I am going home', instead of *be xāne miravam*), where a preposition is lacking or has fallen away (Lazard 1986, 252; Lazard 1990, 189).

The same can be said of the use of the *iḍāfa* particle as a relative pronoun, well-known from New Persian texts in scripts different from Arabic. In this case too, more modern forms replaced the outcome of the old Middle Persian relative pronoun *ī*; but scattered traces of it can be detected in literary Early New Persian texts (Lazard 1963, 490–491, §§ 855–856), as well as in the modern spoken variety (Lazard 1990, 188–189).

Arabo-Persian orthography does not distinguish between some originally different morphemes: the Middle Persian conjunction *kū* 'that, than', the relative/interrogative pronoun *kē* 'who, which', and the temporal/conditional conjunction *ka* 'when, if', all three merged in what seems to be a single new form *ki*, spelled <ky> or <kh>, or simply <k> joined to a preceding pronoun (<ʾnk> *ān-ki* 'that which') or to a following word beginning with a vowel (<kʾmd> *k-āmad* 'who came') in early Arabo-Persian orthography.[8] These morphemes were still distinct

---

7 For the meaning of these glottonyms, and the distinction between *pārsī* and *darī*, see the studies by Gilbert Lazard reprinted in the book *La formation de la langue persane*, Paris 1995. Roughly speaking, *pārsī* means generically 'the Persian language', while *darī*, '(language) of the court', indicates – at the beginning of the Islamic era – the north-eastern variety from which literary New Persian sprang.

8 Šams-i Qays (1981, 249) says that both spellings, with final <y> or <h>, are only a graphic device intended to represent a preceding short *i* vowel (*kasra-yi mā-qabl*). When *ki* is used with an interrogative and abstract (*istifhām-i mujarrad*) value, as in the expression *ō kī-st?* 'Who is he?', however, Šams-i Qays recommends the spelling with final <y> which, he says, 'is well perceptible in pronunciation too'. Perhaps he meant that, when this form had the value of an interrogative

– at least graphically – in some of the most ancient Judaeo-Persian texts, but often appear to be used interchangeably as far as their function is concerned; and *ka* 'when, if', spelled <k'>, has been retained only in the text of religious controversy referred to as Argument, published by MacKenzie in 1968 (Paul 2013, 151–152, § 185c, and 168–169, § 207). Manichaean New Persian too, despite its adherence to the graphic tradition of Middle Iranian languages, especially Sogdian and Middle Persian, in Manichaean script, shows some traces of confusion (see de Blois 2006, 106, *s.v.* k'; Provasi 2011, 166-167, *s.vv.* k', kw, ky). This means that – during the formative period of Arabo-Persian orthography – such morphemes were losing, or had already lost, their distinct meaning due to a possible formal coalescence caused by the shortening of the vowels (*kū > ku*, *kē > ki/ke*). This probably helped their merging into one single form.

Some of the earliest Judaeo-Persian documents – from both south-western Iran (as for example *Argument*; but not the two dated documents referred to as Group C below) and Central Asia (for example the two letters from Dandan Uiliq in Central Asia: Group B below) – give evidence of the existence of two different prepositions, corresponding to the sole Arabo-Persian preposition *ba(d)* written <b-> (or <bd-> before a vowel):[9] *pa(d)* from Middle Persian *pad* 'to, at, in, on', spelled <pd>, <p'>, or <p-> joined to the following word; and <by>, to be probably read *bē*, with a directional value 'to, towards'. These two prepositions are also attested – respectively spelled <b-> and <by> and with the same distinct meaning – in an interlinear translation of the Qur'an, in Arabic script, written in a particular variety of New Persian rich in dialectal features attributable to Sistan (south-eastern Iran). The manuscript of this text, referred to as *Qur'ān-i Quds* and dated by Lazard to the end of the 11th century (1990, 188, 192), has been edited, in facsimile and in a diplomatic edition, by Rivāqī (1985).[10]

As to the genesis of the Arabo-Persian preposition *ba(d)*, apart from a possible coalescence between the two prepositions *pa(d)* and *bē* mainly evidenced by *pa(d)* also acquiring a directional value,[11] contamination with the almost synonymous Arabic preposition *bi-* 'with, in, by, at, near' can also be cited, at least as

---

pronoun, it was still articulated with a long *ē* or *ī*, as it is today. On the form of this/these morpheme(s) in the most ancient New Persian prose texts in Arabic script see Lazard 1963, 237, § 253 (pronoun); 473, § 809 (conjunction).

9 Lazard 1963, § 145, p. 191 shows that vocalised manuscripts also attest an occasional *bi* or *bu* (in labial context) pronunciation.

10 A clear study of the repartition of the two prepositions' usages has been given by Lazard (1986, 245–247 and 1990, 187–188).

11 Paul (2003, 179–185, and especially the table on p. 185) shows that it is only in New Persian that ancient *pa(d)* also acquired a directional meaning.

far as its spelling with initial <b-> is concerned (de Blois 2006, 109, n. 8).[12] Indeed, the Arabo-Persian preposition is spelled with initial <b-> – instead of <p-> attested by coeval texts in scripts different from Arabic – even in the most ancient manuscripts which make at least occasional use of the four 'Persian' letters. The Manichaean New Persian fragment M 595a+ edited by Provasi shows a curious inverse spelling for the verbal prefix *bi-*, written <pd> as the preposition (Provasi 2011, 161–162, 166). This might indicate that the rather inaccurate scribe of this fragment confused the two morphemes, perceiving them as homophonous (at least as regards their initial consonant); or that he was influenced by the Arabo-Persian orthography, in which the verbal prefix and the preposition were both spelled <b->.

2.3 The rule of *dāl* and *ḏāl*

As seen above, a clear normalizing aim seems to have presided over the establishment and development of Arabo-Persian orthography. This ended up with the exclusion of all forms considered too dialectal or colloquial from writing, and therefore from literary New Persian; and with the simplification of a number of forms which had lost their distinct original form or meaning.

Among the normalizing choices connected with the development of early Arabo-Persian orthography there is – to my mind – the so-called 'rule of *dāl* and *ḏāl*' well-known from the most ancient literary manuscripts. According to this rule, all *d*s after a vowel or a diphthong, within the same morpheme, were spelled <ḏ>, i.e. as an interdental voiced fricative (ex. <pḏr> *piδar* 'father'); instead, after a consonant or at the beginning of a word (ex. <drd> *dard* 'ache') or after morpheme boundary (ex. <bd'n> *bi-dān* 'know'), they continued to be spelled <d>. The complementary distribution of *d* and *δ* implied by this rule concerned only the words of Persian (or Iranian) origin and the loanwords, such as those from Greek, entered into Persian at an ancient date. By contrast, the development of *δ* from postvocalic *d* did not concern Arabic loanwords, probably because in Arabic /d/ and /ḏ/ were (and are) two distinct phonemes. As manuscripts follow this orthographic usage approximately until the middle of the 13th century, scholars have supposed that, around that time, the complementary distribution of the outcome of historical /d/ in Persian words, as *d* or as *δ*, came to a halt and all

---

12 Martin Schwartz (personal communication) suggested another possible explanation for *pa(d) > ba(d)* in New Persian: alignment with the initial *b-* of many other prepositions (*abar*, *abāg*, *abē*, *abāz*), that by then had lost their initial vowel.

'Persian' *δ*s (but not Arabic ones) became *d* again, as they are today, with the exception of a small number of words which retained *δ* – afterwards pronounced *z* – from historical *d*: *guδaštan* 'to pass', *guδāštan* 'to put', *paδīruftan* 'to receive', and a few others.

To explain this change in the manuscript tradition, Pisowicz (1985, 109) formulated another hypothesis, namely that from the second half of the 13th century the allophonic development of the ancient postvocalic *d* to *δ* – which he considered characteristic of central and western dialects – was ousted from the literary language due to the influence of north-eastern dialects, where this phenomenon had not occurred. Indeed, in a famous passage from the *Muʿjam*, Šams-i Qays says that 'in the language of the people of Ghaznīn, Balkh, and Transoxiana there are no *ḏāl*s, all of them being pronounced as the letter without points [i.e. *dāl*]' (Šams-i Qays 1981, 221); this – according to Šams-i Qays – would explain a number of irregular rhymes, which do not comply with the distinction between *dāl* and *ḏāl*, in the works of poets coming from these regions. Meier (1981, 104), however, had already shown that the spirantisation of postvocalic *d* was not unknown to the north-eastern dialects, considered as a whole; according to him the boundary between the areas affected and not affected by this phenomenon passed between Marv in Khorasan (spirantisation), and Bukhara in Transoxiana (without spirantisation). Pisowicz's hypothesis appears problematic also because the influence of north-eastern dialects on the literary language was certainly relevant from the very beginning of the history of literary New Persian; therefore it is unlikely that such an influence had had consequences only in the 13th century.[13] When Šams-i Qays says that 'in the language of the people of Ghaznīn, Balkh, and Transoxiana [i.e. Afghanistan and Central Asia] there are no *ḏāl*s', he is simply noting that not all north-eastern dialects have a spirant allophone of /d/, this causing some difficulties for the poets coming from these regions; but he is not stating anything about the complementary distribution of *d* and *δ* implied by the orthography of the most ancient manuscripts.

Recently Filippone, in the context of a study of the language of the *Qurʾān-i Quds*, where no instances of spirantisation of postvocalic /d/ in words of Persian origin are to be found, has conducted research on Persian dialects to find out the dialect or the dialects from which the complementary distribution of *d* and *δ* represented in literary ancient manuscripts could have originated. She concludes: 'In analyzing these phenomena, I think that one should avoid reconstructing

---

13 Recent summaries of this complicated question have been given by de Blois 2006, 94–96; Orsatti 2007, 94–98 (with translation of the relevant passage from Šams-i Qays); Filippone 2011, 185–186.

highly standardised realities, with a homogeneity in time and place which has probably never existed. The tendency to a certain (contextual or free) variability of *d/δ/z* appears as a constant factor throughout the history of West Iranian. [...] But cases of *d/z* fluctuation are mostly bound to single words. In many cases they remain unexplained' (Filippone 2013, 186). This is exactly the situation of modern standard New Persian, where *ḏāl* (pronounced *z*) from historical postvocalic /d/ is only found in a limited number of words.

At this point, a possible solution would be to consider the complementary distribution of *dāl* and *ḏāl* attested in early literary manuscripts as the result of the application of a (mainly orthographic) rule intended to set order into the multiform realisations of /d/ in the spoken or dialectal varieties of Persian. Šams-i Qays, in the previously mentioned passage, attributes the complementary distribution of *dāl* and *ḏāl* to the 'correct *darī* language', that is to literary New Persian, and not to Persian *tout court*. Throughout Chapter Two of Section Two of the *Muʿǧam*, devoted to 'The letters of the rhyme' (where – concerning the rhymes of the letter *ḏāl* – the passage in question is to be found), he distinguishes carefully the Persian language, called *pārsī* or *luġat-i pārsī*, from the *luġat-i darī*, 'the *darī* language', often qualified as *ṣaḥīḥ* 'correct'. For example, he says that in the Persian language (*pārsī*) most words end with a quiescent letter, i.e. – roughly speaking – in a consonant (Šams-i Qays 1981, 209); and that among the peculiarities of the Persian language (*luġat-i pārsī*), there is the fact that the clause is not complete without the copula (Šams-i Qays 1981, 215). When he notes that in Persian the final long *-ā* is normally shortened, he refers this linguistic notation to the colloquial form of Persian (*muḥāwarāt-i pārsī*) (Šams-i Qays 1981, 211). When the author uses the term *darī*, instead, he refers to a rule or a canon, often even complaining of the lack of a clear criterion on which to rely (Šams-i Qays 1981, 205).

That the rigorously complementary distribution of *d* and *δ* was a phenomenon only affecting the literary language is proved by the fact that it seems to be unknown in New Persian texts of non-literary character. Neither the ancient New Persian texts in Hebrew script coming from south-western Iran (see Group C below), nor the non-literary *Qurʾān-i Quds* in Arabic script, from Sistan, bear any trace of a spirant pronunciation of postvocalic /d/ (Lazard 1995, 136), and even less of a complementary distribution of *d* and *δ*. Among the New Persian texts written in scripts different from Arabic only one shows a clear complementary distribution of *d* and *δ*: it is the fragment of the bilingual (Syriac, and New Persian in Syriac script) Psalter from Central Asia (edited by Müller 1915, Sundermann 1974, and Sims-Williams 2011). Sundermann, however, ascribes the orthographic usage in this manuscript to the influence from the coeval Arabo-Persian orthography (1974, 450). The Manichaean New Persian orthography,

though having at its disposal a distinct letter <δ>, attests only occasional renderings of Persian postvocalic /d/ as δ in <ʼryδ> āraδ 'he brings' and <nbwδm> nabūδam 'I was not' in the fragment of the *qaṣīda* published by Henning (1962, ll. 10 and 34); and in <gwδrʼndg> *guδaranda* 'passing' in the so-called 'Catechism' (*Lehrtext*) published by Sundermann (2003, c2, where this form occurs twice, this verb being one of the few Persian words that have retained the letter *ḏāl* till now): a total of only three words.

Therefore, the complementary distribution of *d* and *δ* should probably not be considered as a genuine linguistic phenomenon originating from one or more Persian dialects and hence entering into the literary language, but as a rule – in fact, ancient authors often speak of 'rule' – that was supposed to be applied to literary New Persian (*darī*) and had its main scope and field of application in orthography and in the scholastic and artificial pronunciation characteristic of poetry (see also Section 2.5). Probably around the mid-13th century such a rule ceased to be extensively applied in the copying of literary manuscripts.

2.4 The most ancient New Persian texts in Arabic script: the Codex Vindobonensis and the marriage contract from Bāmiyān

One of the most ancient dated Persian literary manuscripts is the Codex Vindobonensis (cod. A.F. 340 of the Österreichische Nationalbibliothek in Vienna), containing a pharmacological tractate by Abū Manṣūr Muwaffaq b. ʻAlī al-Hirawī (therefore from Herat in modern Afghanistan), copied by the poet Asadī of Ṭūs in Šawwāl 447 / 24 December 1055 – 21 January 1056 (facsimile editions: Muwaffaq 1972, 2009).[14] A relatively short span of time separates the composition of the work, which can be dated to the second half or end of the 4th/10th century, from this copy. Both the author and the copyist were from Khorasan, in the east of the historical Iranian territory.

The orthography of the manuscript is not too different from today's. The Arabic loanwords are spelled as in Arabic, though some orthographic usages of Arabic are not retained: a principle which has presided over Arabo-Persian orthography over the years. The *tā marbūṭa* alternates with the *tā ṭawīla* 'long *tā*', both read *-at*, or with final <h>, read *-a*. In particular, taking into consideration the second double-page of the manuscript (fols 2v–3r), the *tā marbūṭa* is used in *ǰihat*

---

14 For a critical evaluation of the manuscript for linguistic studies and a bibliography until then cf. Lazard 1963, 45–48, N. 4.

Fig. 1: Codex Vindobonensis, fols 1v–2r (from Muwaffaq 1972).

Fig. 2: Codex Vindobonensis, fols 2v–3r (from Muwaffaq 1972).

(2v 9), *saʿādat* (2v 10), *ḥaḍrat* (3r 1 and – probably with loss of the two points – 2v 5), and *ziyādat* (3r 1); the *tā ṭawīla* is found in *quvvat* (2v 1, followed by the plural suffix *-hā*), in *quvvat u maḍarrat u manfaʿat-aš* (3r 5, before the suffix pronoun *-aš*), and in *qismat* (3r 7); instead, the writing with final <h>, certainly read *-a* (in fact, it is followed by the <y> indicating the *iḍāfa* after a vowel), is found in two Arabic words: *xizāna* (3r 2) and in *daraja* (3r, 5, 8–12). *Hamza* is not written in *muʾayyad* (2v 5), *taʾammul* (2v 7) and *taʾlīf* (3r 2). The latter spelling could actually represent a linguistic feature of the language of the text: the dropping of the glottal stop with compensatory lengthening of the preceding vowel (Meier 1981, 128, 133–134).

In this manuscript, final <y> – besides its normal form(s) – also has a small form, mostly used to represent the *iḍāfa* after a word ending in *-a* (see 1v 4,5; 2r 2; 3r 2, 5, 8–12). This small *yā* can be considered as the origin of the 'Persian' *hamza*, i.e. a *hamza* placed over or by a final <h> indicating *-a*. In fact, in fol. 1v, a second hand seems to have begun to replace these 'little *yā*s' with more modern *hamza*s (ll. 4 and 5).

The manuscript is fully vocalised and provided with orthographic signs, and complies with the 'rule of *dāl* and *ḏāl*'. As to the so-called *majhūl* 'unknown' vowels, i.e. the vowels *ē* and *ō* unknown to the Arabs, a vocalisation for *ē* different from *ī* is well attested (Meier 1981, 86–87), whereas I found no examples for *ō*. A massive use is also made of the distinctive signs (additional diacritical points and little letters written above or below the main letters) intended, in addition to the usual diacritic points, to differentiate letters of the same form:[15] three points under *sīn* and *kāf* to differentiate them from *šīn* and *gāf*; one point under *rā*, *dāl*, *ṣād* and *ṭā*, to differentiate them respectively from *zā*, *ḏāl*, *ḍād* and *ẓā*; a letter underwritten, to differentiate *ʿayn* from *ġayn* and *ḥā* from *xā*. Final <y> has often two subscribed points, as in Arabic writing, a graphic usage afterwards abandoned in Arabo-Persian orthography.

Among the other more relevant orthographic usages of this manuscript there are the following:

– the 'Persian' letters <p>, <č> and <g> are occasionally used (I have found no instances of <ž>), all three written with three points below, but often replaced respectively by <b>, <j>, <k>;
– letter ڤ representing the spirant allophone of /b/ occurs quite regularly in verbal forms from *aβzūdan* 'to add' and *aβgandan* 'to throw', and in the suffix *-βām* 'color';

---

15 Grohmann 1971, 42–46, § 4.

– the preposition *ba-* 'in, to, by' (<*pa/paδ*>) is always written with< b> – and not with the more ancient <p> – attached to the following word, a usage continued until recent times (now discouraged in favor of the 'separate' spelling <bh>);

– *īn* 'this' and *ān* 'that' are often written, without initial *alif*, joined to a preceding preposition (*az-īn* 2v 3; *andar-īn* 3r 3) according to a usage frequent until very recent times;

– the conjunction 'that'/relative pronoun 'who, that, which' are always written <ky>;

– the verbal durative prefix has the form *hamē* and is always written as a stand-alone word;

– *alif-madda* is only occasionally used; initial *ā-* is often written as a simple *alif*, or as two *alif*s next to one another.

Apart from minor fluctuations continuing until recent times, this already normalised orthography is attested not only in the most ancient literary manuscripts, but also in the two most ancient private documents in Arabo-Persian: a marriage contract dated 470/1078 found in Bāmiyān (Afghanistan),[16] and a deed for the sale of land found in Khotan, dated 501/1107.[17]

Among the orthographic features of the marriage contract, written only about 20 years after the Codex Vindobonensis and originating from Afghanistan, the following should be noticed:

– the orthography of the text is not too different from modern orthography, with the exception that here diacritical points are often omitted;

– all Arabic loanwords maintain their original spelling;

– the four 'Persian' letters are not used: cf. for ex. <sbyd'r> *sapēdār* 'white poplar' (ll. 2, 3, 12) and <krftn> *giriftan* (l. 25);

– the 'rule of *dāl* and *ḏāl*' is not respected; in particular, two Persian words which even today are spelled with *ḏāl* are written with <d>: <bbdrfth> *bipadrifta* 'he has accepted' (l. 28), corresponding to literary New Persian *bipaḏīrufta* and <kwdšth> *gudašta* 'passed, elapsed' (l. 31), corresponding to literary New Persian *guḏašta*;

- in l. 1 the preposition *bar* 'on' is written <vr> or <fr> (it is not easily readable from the photograph published by Scarcia 1963), a spelling which probably represents a dialectal form;
- the verbal prefix *mē-* (< *hamē*), which always occurs in this more modern shortened form, is written attached to the following verbal form;
- the preposition *ba* 'in, to, at', is written as <b-> attached to the following word;
- the conjunction *ki* 'that'/relative pronoun *ki* 'who, which' are both written <kh>;
- the plural suffix, always occurring in the form *-hāy*, is written attached to the preceding word, except after words written with final <y>; e.g. <sr'y h'y> *sarāyhāy* 'houses', ll. 18, 25 and *passim* (there are no plural words ending in <h>);
- *alif-madda* is never used;
- final <y> has, though very rarely, two points below, thus indicating that the 'Arabic' writing of final <y> with two points below, later abandoned, was still in use;
- the coordinative conjunction *u/wa* 'and' is not written, apart from in eight cases (ll. 24, 25, 27, 28 twice, 29, 31 twice), mostly placed at the beginning of a new sentence.

The latter feature is very interesting and represents the sole real divergence from the orthography of literary texts (and from modern orthography). It seems to suggest that in the manuscript of the marriage contract the short *u* of the conjunction was not written, except when it was at the beginning of a sentence or after a pause. We can suppose that precisely at the beginning of a sentence or after a pause the Persian conjunction *u* (< Middle Persian *u*, *ud*) begun to be pronounced *wa*, as in Arabic. These, therefore, would be early attestations of the new form, probably influenced by the Arabic conjunction *wa*, of the Persian conjunction, even now pronounced *va* or *o* according to its syntactic position and elocution speed.

Ancient literary manuscripts too – less carefully copied than the Codex Vindobonensis – show some traces of a failure to write the coordinative conjunction. For example, in the ancient fragment of 'Unṣurī's poem *Wāmiq wa 'Aẕrā* (datable to the 11th–12th centuries) the conjunction is occasionally not written and has been integrated into the edition (Hägg and Utas 2003, 79).

Considering that the two ancient manuscripts analyzed here, that is the Codex Vindobonensis and the marriage contract, pertain to different textual typologies – a beautiful copy of a scientific-literary text and a legal private document – the orthographic differences between them are not so great. The lack of

vocalisation and of any orthographic sign in the marriage contract, as well as the defective diacritical pointing of the letters, are clearly to be connected with the practical scope of the document, devoid of any aesthetic pretension, and with the formulaic character of the text, written in a highly standardised language. The comparison between the orthography of the two texts shows a high degree of normalisation from early times, at least as far as Eastern Iran and (today) Afghanistan are concerned.

In this regard one fact is particularly meaningful: in the two most ancient non-literary documents, i.e. the marriage contract from Bāmiyān and the deed for the sale of land from Khotan, the Persian word *pānṣad* 'five hundred' (Scarcia 1963, ll. 3, 22; Margoliouth 1903, l. 12) is already written with Arabic <ṣ>, as it is now – *ṣad* 'hundred' being one of the few Persian words written with this 'only Arabic' letter. The reason for this spelling is perhaps to avoid confusion with other homophonous words such as *sad(d)* 'rampart, obstruction' or – as Perry (2002) suggests – with the very common word *šud* '(he) became' which, due to the frequent defective writing of the diacritical points, could be confused with the word for 'hundred'. The fact that both non-literary documents already offer this 'normalized' spelling is a clear proof of the high degree of standardisation of Arabo-Persian orthography from ancient times.

2.5 The spelling of the Arabic loanwords

As we have seen, one of the most striking features of Arabo-Persian orthography since its beginning is the preservation of the original spelling of the Arabic words which entered into Persian, though we can suppose that – once established in the Persian language – the Arabic loanwords were pronounced, as they are today, according to Persian phonology. The only exception are the introduction of the new phoneme /q/, in Early and Classical New Persian still clearly distinguished from /ɣ/ <ġ> (Pisowicz 1985, 111–117) and, perhaps, the pronunciation of Arabic <ḏ>: indeed, given the existence, in words of Iranian origin, of a fricative postvocalic allophone of historical /d/, the letter *ḏāl* could well have been pronounced, in Arabic loans, as an interdental voiced fricative (de Blois 2006, 94).

The preservation of the original orthography of the Arabic loanwords is certainly a consequence of their scholarly origin: the Arabic loanwords entered the Persian language mainly from books, from the written Arabic language; and only gradually, by 'osmosis from above' (Bausani 1978, 13–14), did they penetrate into the everyday language. It was mainly the Persian (or Iranian) bilingual scholars, who knew and used Arabic as a scholarly language (as was Latin for European scholars), who were responsible for the introduction of a great quantity of Arabic

learned vocabulary into their works. If the spelling of the Arabic loanwords in Persian has to be taken as referring to real pronunciation, and not only to the written form of the words, it probably represents mainly a scholarly or literary pronunciation of Classical Arabic, i.e. an artificial and altogether scholastic pronunciation. The same artificial pronunciation of Arabic loanwords is characteristic of the metrical reading of poetry. In Persian poetry, for example, the Arabic *ʿayn* is always 'pronounced', i.e. counts as a consonant, even in positions where, in normal speech, it is (and probably was) never pronounced. Likewise, the *hamza* (glottal stop) of Arabic words is counted as a consonant; and this can happen even at the beginning of a word, before a vowel, when the *hamza* is and was neither written nor pronounced. Moreover, <z>, <ḏ>, <ḍ> and <ẓ> never rhyme together, nor can <s> rhyme with <ṯ> or with <ṣ>, <h> with <ḥ>, or <ʾ> with <ʿ> (Meier 1981, 103). This seems to point to a sort of artificial and scholarly pronunciation of Arabic loanwords in Persian poetry.

The preservation of the original Arabic orthography of loanwords has an important implication (and was also probably dictated by this need): it makes the Arabic loanwords immediately recognisable, without destroying the kinship between words pertaining to one and the same Arabic root. A kind of consciousness of the original written form of the Arabic lexicon within Persian has always been maintained, and is proved not only by the fact that Arabic words can be uttered (in poetry or in scholarly contexts) with a literary pronunciation, approximating to that of Classical Arabic, but also by the fact that even now, in standard New Persian, the intervocalic glottal stop /ʾ/ can be replaced by the glide *y* only if it is represented in the original writing by a *hamza*, but never if it is represented by the homophonous (in Persian) letter *ʿayn*: for ex. /lāʾeq/ 'worthy' can be uttered as [lāʾeq], [lāyeq] or [lāeq]; but /šāʾer/ 'poet', from Arabic *šāʿir*, cannot be pronounced *šāyer, given the origin of /ʾ/ from an Arabic *ʿayn* in this word (Pisowicz 1985, 20, 102).

The orthography of Arabic loanwords in Persian has remained virtually unchanged throughout the entire history of the New Persian written tradition, remaining impermeable to any influence from the different diachronic and dialectal varieties of Arabic (a relevant exception is represented by the Arabic loanwords ending in *tā marbūṭa*, on which cf. Perry 1991 and 1995).

An accurate preservation of the original Arabic spelling for Arabic loanwords is to be found not only in Arabo-Persian orthography,[18] but also – where possi-

---

18 One exception is the Persian orthography of words with *hamza* and their phonetic realisation, on which thorough historic-linguistic research is still lacking, apart from useful remarks in the excellent work on the history of the Middle and New Persian phonology by Pisowicz (1985, 20, 47–51, 102).

ble – in the orthography of the New Persian texts in scripts different from Arabic. Indeed, for the redaction of New Persian texts, with their rich Arabic lexicon, the Hebrew alphabet had the possibility of transliterating many of the 'only Arabic' letters: *ḥēṭ*, as opposed to *hē*, was used to transliterate Arabic *ḥā*; *ṭēṭ*, as opposed to *tāw*, was used for Ar. *ṭā; ʿayin* was available for Ar. *ʿayn*; *ṣādē*, which in the adaptation of Aramaic-based scripts to the Iranian languages had already been employed for Iranian *č* (Skjaervo 1996, 516), and partially for *ǰ*, was of course also suited to represent Ar. *ṣād*; and *qōph* was available for Ar. *qāf*. Moreover, for some of the Arabic and Persian sounds not represented in the Hebrew alphabet, the Judaeo-Persian writing system resorted to the possibility of representing the Hebrew spirant allophones of the plosives by means of diacritic signs: hence *tāw* was used to represent both /t/ and, with or without diacritics, the letter *ṯā* of the Arabic loans; *dālet* was used for /d/ and, with or without diacritics, for Ar. *ḏāl*; *kaph* for /k/ and for /x/ (in both Arabic and Persian words); *pē* was used for Persian /p/ and for Arabic and Persian /f/. The only Arabic letters which had no possible graphic equivalents in the Hebrew alphabet were *ḍād* and *ẓā*, for which – as well as for Persian and Arabic /ǰ/ – a series of different solutions were adopted (amply described by Paul 2013, 30–33, §§ 11–12).

For Persian texts in Syriac script, too, the 'only Arabic' letters were easily transliterated by means of the corresponding letters <ḥ ṭ ʿ ṣ q> of the Syriac alphabet. The Syriac letters <t> and <d>, which could also represent, with a point under the letter, the spirant allophones of /t/ and /d/, were also used to transliterate the Arabic letters *ṯā* and *ḏāl*; likewise <k>, <p> and <g> with a point below were used to represent Arabic and Persian /x/, /f/ and /ɣ/. On the model of Arabo-Persian orthography, however, in some texts /f/ was represented by <p> with a point above, instead of below. For Arabic <ẓ>, for which no letter was at hand, some Syro-Persian texts used <ṭ> with a dot above, again a clear calque of the Arabic letter *ẓā*.

The same can be said for the New Persian texts in Manichaean script. By the time Manichaean script was being adapted to New Persian (to be tentatively placed at the end of the 9th – first half of the 10th century), six new letters had already been added to the original 22 of the Aramaic alphabet to write other Iranian languages, and were already present in Manichaean texts in Sogdian: <β γ δ f j x>. For the Arabic words entering New Persian, therefore, the letter <δ> was used to transliterate *ḏāl*; and the same letter, single or more often doubled, <δδ>, was used to transliterate *ṯā*. Two new letters introduced by punctuation (<k> and <q> with two dots above) – already used in Turkish texts in Manichaean script – were used in Manichaean New Persian to transliterate *qāf*, given that simple <q> (without dots) had already been used (as an alternative to <k>) to represent *k*. Moreover <ʾ> with two dots above was introduced to represent *ʿayn* in

Arabic words, because in Manichaean orthography <ʾ> had already been used to represent an initial palatal vowel. For <ḍ> of Arabic words different solutions were adopted: <z>, as in <zʾwbt> for *ḍābiṭ* 'commander; chaste' in the 'Catechism' edited by Sundermann (2003, c19; a different reading for this word is suggested by de Blois 2006, 114, *s.v.*; see also Shokri-Foumeshi 2014, 202–203), and <d> in <hawwd> *ḥawḍ* 'basin, cistern' (Sundermann 2003: e18).[19] On the other hand, the 'Arabic' letters <ṭ>, <ṣ> and <ḥ> could not be transliterated, because the corresponding letters of the original Aramaic alphabet had already been used to note sounds of the Iranian languages.[20] Indeed, Manichaean orthography was already so loaded with graphic habits fixed for other Iranian languages that it became impossible to render all the 'only Arabic' letters into this script. The creation of new letters had in the meantime come to an end, soon to be followed by the disappearance of the Manichaean religion from the pages of history.[21]

3 New Persian texts in scripts different from Arabic

Let us now look at the most ancient New Persian texts in non-Arabic scripts – which cover exactly the period for which we have no original documents in Arabo-Persian – in the hope of discovering some indirect evidence about time and place of the formation of the Arabo-Persian orthographic canon.

3.1 Judaeo-Persian texts

Some Judaeo-Persian texts, that is Persian texts written by means of the Hebrew alphabet, are among the most ancient written New Persian documents. The Jewish minorities living in Iran spoke Persian, or one of the various Persian

---

19 The latter spelling is interpreted by de Blois (2006, 96) as being dictated by 'the 'Persian' convention of representing a postvocalic interdental as d rather as δ', exceptionally applied also to an Arabic word. Filippone (2011, 186), instead, thinks that this spelling reflects a dialectal pronunciation.

20 No loanword with Arabic <ẓ> is attested in Manichaean documents published up till now: cf. the glossaries by de Blois 2006 and Provasi 2011, 163–168. A glossary of all Arabic loanwords in Manichaean New Persian texts has been published by Shokri-Foumeshi 2014.

21 On the adaptation of the Manichaean script to the Iranian languages cf. Henning 1958, 73–75; and for writing New Persian, cf. Henning 1962, 89–91, Orsatti 2007, 150–164.

(or Iranian) dialects spread throughout the Iranian linguistic area, as their mother tongue (Yarshater 1974); and, for written purposes, they used Persian (more or less tinted with dialectal features) written in Hebrew characters. In the past, scholars thought that the texts emanating from Jewish minorities revealed a Persian dialect different from the language of their Muslim neighbors. Recent studies, however, especially since the discovery and publication of the manuscript of the Persian dialectal translation of the Qur'an known as *Qur'ān-i Quds* (see above), have shown that this is not so, apart – of course – from the presence in Judaeo-Persian texts of some Hebrew loanwords and expressions. A number of linguistic features known until then only from Judaeo-Persian texts were also found in this Muslim text (Lazard 1990). In general, the Judaeo-Persian texts, as well as the New Persian texts in Syriac and Manichaean scripts, reflect written varieties of New Persian differing from literary New Persian. Recent studies, moreover, have focused on a difficult task: analyzing the dialectal variations within the Judaeo-Persian and other written traditions of Iran.[22]

We do not know when the Hebrew alphabet was adapted to write Persian.[23] As only texts dated or datable to the Islamic period are extant, it is generally supposed that the adaptation of the Hebrew alphabet to the Persian language occurred only in Islamic times. But some scholars think that a Judaeo-Persian literature (in particular translations of the Bible into Persian) in Hebrew characters must have already existed in the Sasanid period (Bacher 1904). An answer to this question would be essential in order to ascertain the linguistic value of Judaeo-Persian orthography and the possible presence of historical spellings. In what follows I will present the main orthographic characters of some Judaeo-Persian texts, grouped according to chronological and geographic criteria.

Group A. The most ancient dated Judaeo-Persian documents are three short inscriptions carved on a rock in a mountainous passage, Tang-i Azao, in Western Afghanistan, left by three merchants bearing Jewish names, who were coming from Kōban, the ancient name of the Qabul valley. These inscriptions (edited by Henning 1957) are dated, according to the Seleucid era, to 1064, corresponding to 752 CE.[24] They are all very short; no Arabic word is attested.

---

22 On Judaeo-Persian dialectology, after the groundbreaking article by Lazard 1968, see Shaked 2009 and Lazard 2014.

23 For a thorough analysis of the adaptation of the Hebrew alphabet to write Persian, cf. Paul 2013, 23–48.

24 Rapp (1967, 55–56) has unconvincingly questioned the dating proposed by Henning (1957, 338), proposing a much later date: 1299–1300 CE.

The main orthographic features of these inscriptions are the following:
- the *iḍāfa* particle is written with a *yōd* attached to the following word, unlike the Arabo-Persian orthography: <(y)'r y'wy> *yār-i ōy* 'his Friend' (C3);
- the suffix pronoun of the 3rd singular person *-aš* is written separated from the preceding word, i.e. with initial *'āleph*: <y'r 'š 'w b'd> *yār-aš ō bād* 'May He be his helper' (A3, B2-3);
- the letter *qōph* of the Hebrew alphabet is used to represent the Persian sound /k/: <qnd> *kand* '(he) incised' (A2, B2), <qy> *ki* 'who' (A2).

The use of <q> to represent /k/ implies that the Hebrew *kaph* could be left to represent the voiceless uvular spirant /x/ of Persian (no occurrence of *x* is to be found in these short inscriptions, however), for which sound no letter of the Hebrew alphabet was at hand. This can be taken as evidence that, in all probability, Arabic words had not entered the Persian language yet, or at least not the language represented by these inscriptions (Lazard 1968, 82).

Group B. Chronologically, after the inscriptions of Tang-i Azao there follow two letters discovered at Dandan Uiliq (Central Asia, northeast of the Khotan oasis), and referred to as DU1 and DU2, datable to the second half of the 8th century. Of the first one, a fragment of a commercial letter, a continuous reading cannot be given, as the left and right margins of the sheet have been badly damaged.[25] The other, also coming from the same area and certainly written in Khotan, has recently been published by Zhang and Shi (2008) with a study in Chinese (which I have not been able to read). Both represent the same language and the same orthographic usage.

The main orthographic features of these letters are:
1. <q> represents Persian /k/, and <k> represents /x/: <qrdwm> *kardum* 'I made' (DU1, 2), <kwdh> *xudah/xudāh* 'God' (DU1 and DU2 *passim*);
2. sometimes long *ā* is not written: <yptwm> *yāftum* 'I found' (DU1, 28), <sd hzr> *sad hazār* 'hundred thousand' (DU2, 1), <kwhrq> *xwāharak* 'sister' (DU2, 3);
3. the *iḍāfa* particle/relative pronoun is variously represented:
   - <'y>: <'z swy 'y mn> *az sōy-i man* 'from me' (DU1, 18), <'z swy 'y dwyd r'> *az sōy-i Dawīd-rā* 'from David' (DU2, 34) <q'r 'y prmwdy> *kār-i farmūdī* 'the work which you ordered' (DU1, 29, with *iḍāfa* as relative pronoun);

---

25 Published by Utas (1968), with a bibliography of the previous studies. Lazard (1988) has given a valuable contribution to the reading and interpretation of a number of passages from this text.

- not written before a palatal vowel, with which it probably blended: <pnn'm yzyd> *pannām(-i) īzid/ēzid* 'in the name of God' (DU2, 1), <tn yšm'> *tan(-i) išmā* 'yourself' (DU1, 23),[26] <n'mh yšm'> *nāma(-i) išmā* 'your letter' (DU1, 28, 33), <mrdwm'n yšm'> *mardumān(-i) išmā* 'your people' (DU2, 3);
- <y-> : <kwdh yqrbqr> *xudah-i kirbakkar* 'the beneficent God' (DU2, 1);
- <-y> attached to a preceding demonstrative pronoun *ān* 'that': <'ny ān-i> 'belonging to' (DU1, 4, 13), <wbr kwndwm 'ny nbyšt bwdy> *u bar xwāndum ān-i nibišt būdī* 'and I read what you had written' (DU2, 7, with -*i* relative pronoun and an old past participle, without final -*a*)

4. the conjunction is written <w> attached to the following word: <swd wzy'n 'y man> *sūd u-ziyān-i man* 'my profit and loss' (DU1, 14), <bzwrg wqwdq> *buzurg u kōdak* 'young and old' (DU2, 3);

5. <ṣ> is used for *č* and *ǰ*: <ṣwn> *čūn* 'as' (DU1, 13), <ṣmh> *ǰāma* 'clothing' (DU1, 10), <pnṣ> *panǰ* (DU2, 14, 27);

6. <by> representing the directional preposition *bē/bi* is well attested: <by šm'> *bē šimā* (or *bē-šmā* with contraction?) 'to you' (DU2, 27);

7. final -*a* is regularly written <h>: <prwkth> *furōxta* 'sold' (DU1, 10), <n'mh> *nāma* 'letter' (DU1, 28, 32), except than in the monosyllables <p> *pa* 'in, to' (DU1 and DU2 *passim*), <m'> *ma-* (verbal prohibitive prefix) <m' kwr> *maxwār* 'do not suffer' (DU1, 30), <n'> (verbal negative prefix) <n' d'nwm> *nadānum* 'I do not know' (DU2, 27), and in the pronoun *išmā/šimā*;

8. <'> is sometimes omitted before an initial vowel other than long or short *a*: <dwr by wpt'd> *dūr biyuftād* '(it) was delayed' (DU1, 7), <ydwn> *ēdūn* 'so' (DU1, 22, 24, 31), <yzyd> *īzid/ēzid* 'God' (DU1 and DU2 *passim*);[27]

9. in the first letter (DU1) there are no Arabic loanwords, but in the second one (DU2) a few are attested. Their spelling does not reproduce the original Arabic writing: *ḥakīm* 'doctor' is spelled <hqym> (Du2, 4, 13), *ḥarb* 'war' is spelled <hrb> (DU2, 33), without the Hebrew letter <ḥ> being used to represent the Arabic emphatic *ḥ*;

---

26 In these texts the personal pronoun for the second person singular is probably to be read *išmā* (also perhaps alternating with *šimā*), a form occasionally attested in Early New Persian (Lazard 1988, 208).

27 See the word <myd> *umēd* 'hope' in the *Qur'ān-i Quds* in Arabic script (Filippone 2011,190) and <mwd>, with dialectal *u/ū* for *ē*, in the Judaeo-Persian inscription A (l. 3) from Tang-i Azao (Henning 1957, 342 and n. 2).

10. perhaps the ancient directional preposition *ō* (reduced to *o/u* or *a* and represented by <'> attached to the following word) is represented in <'pyš> *a-pēš* 'near, before' (DU2, 8).

Therefore – except for the coincidence with the Arabo-Persian orthography in the spelling of final *-a* as <h> – the occasional defective spelling of *ā* and of initial *ʾāleph*, the different ways the *iḍāfa* particle is written, the use of letters <q> for /k/ and <k> for /x/ are all features representing an altogether different orthographic tradition, compared to the Arabo-Persian one.

Group C. Two dated documents pertaining to the Ahvaz or Southwestern group of Judaeo-Persian documents (10th – first half 11th centuries) are considered here: a legal document dated 1262 of the Seleucid era/ 950 CE, edited by Shaked (1971); and a legal document dated Ahvaz (Khuzistan, south-western Iran) 1332 Seleucid/1020 CE. In the latter (edited by Asmussen 1965, and generally referred to as *Law report of Ahvaz*):

1. <k> represents *k*, and <q> represents *q* of Arabic or Hebrew words, in contrast to the usage in the first two Judaeo-Persian texts already discussed: <knd> *kand* 'he snatched' (l. 5), <qw'my> *qiwāmī* 'right, lawful' (l. 12);

2. the letter *ṣādē* <ṣ> is used for Persian *č* and *ǰ*, but is also used to represent the Arabic letters *ḍād* and *ṣād*:
   – <ṣ> = *č*: <bṣ'> *bač(č)a* 'baby' (l. 4)
   – <ṣ> = *ǰ* in both Persian and Arabic words: <ṣwml'> *ǰumla* 'whole' (l. 2), <pnṣ> *panǰ* 'five' (l. 5), <ṣwpt> *ǰuft* 'couple' (l. 5), <ṣw'b> *ǰawāb* 'answer' (l. 7); only once *ǰ* is written <g>: <w'gyb> *wāǰib* 'necessary' (l. 8)
   – <ṣ> = *ṣ*: <myṣr> *Miṣr* 'Egypt' (l. 5) (for Persian *s*, the letter *samek* <s> of the Hebrew alphabet is used: <s'l> *sāl* 'year')
   – <ṣ> = *ḍ*: <ṣrwr'> *ḍarūra(t)*, 'need' (l. 8), <r'ṣy> *rāḍī* 'content' (ll. 10–11)

3. <h> e <ḥ> of the Arabic alphabet are carefully distinguished: <šhwtwm> *šahwat-um* 'my desire' (l. 5), <ḥwṣt> *ḥuǰǰat* 'proof' (l. 15);

4. <t> and <ṭ> are carefully distinguished: <byst> *bīst* 'twenty' (l. 5), <slṭ'ny> *sulṭānī* 'sultanial' (l. 12);

5. the Arabic *ʿayn* <ʿ> is always written: <ʿwṣ> *ʿiwaḍ, ʿawaḍ* 'compensation' (l. 11);

6. the *iḍāfa*: is spelled as <y-> attached to the following word or is not written:
   – <y->: <kwd'wnd'n ymylk> *xudāwandān-i milk* 'the owners of the property' (l. 9), <ʿwṣ y'yn drh'> *ʿawaḍ-i īn durhā* 'as recompense for these pearls' (l. 11), <p'yn ykyrdy> *p-īn-i kirdī* 'in this that you have done' (l. 7, with written *iḍāfa*/relative pronoun)

– not written: <ʾwṣ ʾyn dynʾr> *ʾawaḍ-i īn dīnār* 'in place of these dinars' (l. 12); <bd kyrdy ʾyn kyrdy> *bad kirdī īn(-i) kirdī* 'you did wrong (what) you did' (ll. 6–7, with not written *iḍāfa*/relative pronoun);

7. the directional preposition outcome of Middle Persian *ō*, already reduced to *u/o* or probably *a*, is spelled <ʾ> attached to the following word: <ʾpyš> *a-pēš* 'near, before' (ll. 1, 3, 9, 10); it can also introduce a direct object: <ʾḥṣrʾ kyrd ʾdnyʾl> *iḥḍār kird a-Daniel* 'she cited Daniel' (l. 3);[28]

8. the directional preposition <by> is not attested.

Group D. Here only one text is considered, the *Tafsīr* of Ezekiel, that is a translation and commentary of the Book of Ezekiel. Its manuscript (ms. Firkowicz I 1682 of the St Petersburg Public Library, edited by Gindin 2007) is datable to a period between the late 10th and the early 11th century. The so-called 'Part 1' of the manuscript (pp. 1–169 and 221–226) seems to represent a northeastern dialectal variety of New Persian probably originating from northeastern Iran or Afghanistan (Gindin 2007, 23–26).[29] In Part 1:

1. the Arabic words are transliterated, and retain their original Arabic spelling:
    – <ṭ> (and not <ṣ>, as in the Ahvaz document) is used for Arabic *ḍ* (cf. Paul 2013, 33)
    – <g>, or <g> with a stroke (and not <ṣ>, as in the Ahvaz document) is used for *ǰ*, in both Arabic and Persian words: <gʿwd ʾz ʾw> *ǰud az ō* 'different from that' (5.8, Gindin 35), <gʿʿmʾyhʾ> *ǰāmayihā* 'clothes' (38.4, Gindin 73)

2. the orthography of the *iḍāfa* particle is the same as in the Arabo-Persian writing system, apart from rare cases in which it is written <-y> also after a word ending with a consonant: <šrḥy ʾn ʾydr by krd> *šarḥ-i ān ēdar bikard* 'he explained it here' (3.14, Gindin 33)

3. final *-a* is written <h> mainly in past participles, but retains the spelling with final <ʾ> in monosyllables and in many words:
    – final *-a* is written <h>: <bwdh> *būda* 'been' (35.21, Gindin 69),
    – final *-a* is written <ʾ>: <hmʾ> *hama* 'all' (*passim*), <ṣʿyhrʾ> *čihra* 'face' (38.7, Gindin 73), <pʾ zmʾnʾy mšh> *pa zamāna-yi Mōšeh* 'at the time of Moses' (5.10–11, Gindin 35)

---

28 The same form <ʾḥṣrʾ>, probably to be read *iḥḍār*, is also found in the other legal document dated 950 (Shaked 1971, l. 1).

29 Recently Lazard (2014, 91–92) has instead argued that the language of this *Tafsīr* has a northwestern or central-northern origin. Quotations are given according to the page and line of the manuscript, followed by the page in Gindin's edition.

  – monosyllables: <m' kwn> *makun* 'do not do' (5.6, Gindin 35), <n' kwnd>
    nakunad 'does not do' (*5.9*, Gindin 35)
4. the defective writing of long ā is very rare, not to say absent, and words are
   written as in Arabo-Persian.

From this brief sketch it is possible to conclude that the orthography of the
Judaeo-Persian texts taken into consideration seems to show a clear trend: as the
language becomes richer in Arabic loanwords, orthography shows an effort to
find the best way to transliterate them. The few Arabic loanwords attested in the
second letter from Dandan Uiliq are still written as if they were Persian words,
without distinguishing the 'Arabic letters' (Group B). Instead, the numerous
Arabic loanwords in the Ahvaz legal document (Group C) are carefully transli-
terated, despite the multiple values given to the letter *ṣādē* <ṣ>, which seems to
represent the weak point in the writing system represented in this text. In this
text, however, the word <w'gyb> *wājib*, written with <g> for /ǰ/, is interesting: it
shows that for /ǰ/ in the increasing number of Arabic words, a new spelling was
gaining ground, perhaps taken from Judaeo-Arabic orthography. In the *Tafsīr* of
Ezekiel (Group D) the multiple values letter *ṣādē* has in the Ahvaz legal docu-
ment have been made less ambiguous by further differentiating *ḍād* of the Arabic
words from *ṣād*, and Persian *č* from *ǰ* of both Arabic and Persian words.

   The first part of this *Tafsīr* shows an influence from Arabo-Persian ortho-
graphy. Indeed, about the way the Hebrew alphabet is used in this manuscript,
David Neil MacKenzie wrote: 'It is clearly based on a familiarity with the normal
Arabo-Persian script, and indeed an excellent knowledge of Arabic' (MacKenzie
2003, 103–104). If we consider that the *Tafsīr* manuscript is more or less coeval or
even older than the Law report of Ahvaz, it is possible to measure the importance
that the north-eastern origin of this text has in regard to its orthography.

3.2 Syro-Persian texts

At this point I will speak very briefly about the Syro-Persian documentation,
that is, New Persian texts in Syriac script, emanating from Christian milieus in
Iran. Not many texts belong to this group because – in contrast to Persian Jews –
Christians in Iran have very often used the Arabic script; it is interesting to note
that, for example, the translations of the Bible into Persian in Hebrew milieus
are written in Hebrew script, whereas the Persian Gospels are written in Arabic
script. Christians even dated their manuscripts according to the Muslim era (see
for example the manuscript of the Persian lectionary studied by Richard 1981).

Many Syro-Persian manuscripts are not dated and, even when they are datable with reasonable certainly, they have undergone a long transmission which may have exerted a normalizing effect on orthography; therefore they are not so useful for the purpose of giving indirect evidence on Arabo-Persian 'parallel' orthography.

Only the previously quoted fragment from the bilingual (Syriac and New Persian in Syriac script) Psalter from Central Asia is probably attested by an old manuscript. It shows a clear influence from Arabo-Persian orthography, in particular in the complementary distribution of *d* and *δ* (see above). Therefore, again, a text from northeastern Iran gives – through its orthographic characteristics – good evidence of how the coeval Arabo-Persian orthography was by then well-established, and exerted a strong influence on the other graphic traditions of Iran.

3.3 Manichaean New Persian texts

For the purpose of writing their texts, the followers of Manichaeism used the Manichaean alphabet, named after the founder of this religion, Mani (3rd century). Middle Persian and Parthian, two Iranian languages, were the languages of liturgy for Manichaeans. Therefore, New Persian Manichaean orthography shows an influence from both these languages' orthography, as well as from Sogdian Manichaean orthography, to the extent that New Persian in Manichaean script can be considered as the last heir of the graphic traditions of pre-Islamic Iran.

From historical sources we know that during the 10th century Manichaeans were obliged to leave Iraq and western Iran taking refuge in northeastern Iran, and especially in Samarkand, the ancient capital of Sogdiana. This was one of the capital cities of the Persian Samanid dynasty, which ruled over eastern Iran in the last quarter of the 9th century and during the entire 10th century. Persian poetry and prose had its first important blossoming under Samanid patronage. Samarkand also had strong links with the Sogdian colonies of Central Asia, in the Turfan oasis and other places in Chinese Turkestan where, at the beginning of the last century, all the extant texts in Manichaean script were discovered.

Two fragmentary poetical New Persian texts in Manichaean script were published by Walter Bruno Henning in 1962. Henning dates the first of these manuscripts, a fragment of the poem Bilawhar and Būdīsaf, to the first half of the 10th century, during the lifetime of the first great poet of Persian literature, Rūdakī. The second, a fragment of a monorhymic lyrical composition (*qaṣīda*), could be – judging from its script – even older than the first one (Henning 1962, 99; see also below).

These texts are of great interest. From the orthographic point of view they are a proof of the historical and conservative character of New Persian Manichaean orthography. Indeed, the two poetic texts become metrically readable only supposing – behind their conservative spelling – the new forms fixed through the Arabo-Persian orthography. Regarding the second text, the *qaṣīda*, Henning made the assumption 'that the poem had originally been written in Arabic script and was then transliterated into Manichaean script by a man who did not understand it properly' (1962, 99). In fact, an occasional mistake in the spelling of some words attached to one another is evidence of poor knowledge of the language by the copyist: <cwzg'hyy> for *juz gah-i*, in the expression *juz gah-i šumār* 'except at the time of the (last) reckoning' (l. 20).

Henning's brilliant hypothesis of an original Arabo-Persian version from which the text we possess has been copied is useful to explain the metrical form of the text (or of both texts), composed according to the new quantitative prosody deeply influenced by Arabic prosody. This hypothesis, beyond permitting Henning's masterly reading of a text so full of gaps, can certainly account 'for the omission of the Iḍāfe-particle (at least four times) and the word "and"' (Henning 1962, 99); but cannot account for all the other places where the *iḍāfa* particle has been correctly inserted (ll. 9 after *bād*, 10 after *pēš* and *šarāb*, 20 after *gah*, 21 after *zēr* and 30 after *sōy*). How could a copyist unfamiliar with Persian have inserted it correctly? Moreover, how could a copyist unfamiliar with Persian have used – for *juz* 'except' in the phrase above – the historical Manichaean orthography <cwz>, with <c> being the transliteration of the ancient Aramaic <ṣ>? In the Arabo-Persian orthography he would have found, rather, <j> instead of <č> (see Table in Section 2.1 above).

Another mistake made by the copyist, the writing <rrd'> instead of <drd'> *dard-ā* 'o grief!' (l. 1), can likewise be evidence that the antigraph from which the copyist was copying was written in Manichaean script. A confusion between <r> and <d> is possible in the Manichaean writing, where the two letters only differ for a point above <r> or below <d>. Instead, though in later styles of the Arabic script <r> and <d> could be confused, it was probably not so in early styles, where these two letters had a markedly different form, and <r> could rather be confused with final <n>.[30]

---

30 For <r/z>, <d/ḏ> and <-n> in the Codex Vindobonensis, see the first line of the text (fol. 1v): *sipās bāδ yazdān-i dānā wa tawānā-rā*. The confusion between <r/z> and final <n> in the earliest Arabic writing styles is proved by some mistakes in the quotation of Persian words by Arabic authors, as <gm'z> (with <z> = <r> with a point above) for Persian *gumān* 'doubt', and <'yḏn> for *ēδar* 'here' (Tafazzoli 1974, 339, 343).

Though an original redaction of the text of the *qaṣīda* in Arabic script cannot be excluded, we can suppose that the text we possess was copied – by an inaccurate copyist who was probably not completely familiar with Persian – not from a manuscript in Arabic script, but from one in Manichaean script already showing a strong influence from the new Arabo-Persian orthography. Evidence of such an influence must be seen – as pointed out by Henning – in the cases in which the *iḍāfa* particle, and in the one case (l. 12) in which the coordinate conjunction are not written.[31] An influence of the coeval Arabo-Persian orthography on the orthography of this text is also proved by the unusual spelling of the *iḍāfa* particle in the already quoted phrase *juz gah-i šumār* (l. 20): in <cwzg'hyy> the *iḍāfa* is spelled <yy> instead of <ʿyg> or <ʿy>, which are the normal spellings for it in Manichaean orthography. Lastly, 'the scribe of the Qaṣīde allows an occasional -*δ*- for postvocalic -*d*- ('*ryδ* 'he brings', *nbwδm* 'I was not')', instead of -*d*-, given throughout by the copyist of the other poetical text in Manichaean script (Henning 1962, 90): perhaps this too is an interference from the coeval Arabo-Persian orthography. For the preposition *pa(d)* and the negative verbal prefix *nē* (or already *na*), the copyist of the *qaṣīda* prefers the joint spellings <p-> and <n-> respectively, as in the new Arabo-Persian orthography, instead of the spellings <pd> and <ny>.[32]

As already stated, Henning considers the fragment of the *qaṣīda* as even older than the other manuscript, the fragment of the poem Bilawhar and Būdīsaf in Manichaean script, which he had dated to the first half of the 10th century. Therefore the fragment of the *qaṣīda*, and even more so its antigraph, should be dated at an early date, possibly between the end of the 9th and the beginning of the 10th centuries CE. The text of the *qaṣīda* turns out to be not only one of the earliest surviving poems of the kind, but an early original document attesting to the parallel development and fixing of the Arabo-Persian orthography.

---

31 Cases of unwritten *iḍāfa*s are also to be found in other New Persian Manichaean texts (whereas they are quite rare in Middle Persian Manichaean texts): an unwritten *iḍāfa* should probably be recognised in an otherwise incomprehensible passage in the Catechism (*Lehrtext*) published by Sundermann (2003, c13–14): <q'lbd k' x'n'g (h)wm'n'g 'st ''b rwšn ny> *kālbad ka xāna humānā ast āb(-i) rōšan nē* 'the body, which is like a house, is not bright water'. Also in the New Persian Manichaean texts published by Provasi there is at least one instance of an unwritten *iḍāfa* (2011, 149). As to the leaving out of the conjunction, we already saw that this is not a rare occurrence in the most ancient Arabo-Persian manuscripts.

32 A full discussion of Henning's hypothesis about a possible Arabo-Persian antigraph for the *qaṣīda* is given by Orsatti (2007, 161–164).

4 Conclusions

On the basis of the New Persian original documents in non-Arabic scripts here examined, we can tentatively date and localise the beginning of the Arabo-Persian orthographic influence on the other written traditions of Iran: northeastern Iran, end of the 9th – beginning of the 10th centuries. By then, this orthographic tradition appears as already fixed. Though it cannot be excluded that some scattered and unsystematic attempts were accomplished here and there in different places of Iran, the hypothesis of a multi-centric origin of the adaptation of the Arabic alphabet to Persian, favored for example by Akimuškin (1987, 332), seems less probable in the light of our documentation.

References

Akimuškin, Oleg F. (1987), 'Persidskaja rukopisnaja kniga' [The Persian manuscript book], in Jurij A. Petrosjan et al. (eds), *Rukopisnaja kniga v kul'ture Narodov Vostoka. Ocherki. Kniga pervaja.* Moskva: Nauka, 330–406.

Asmussen, Jes P. (1965), 'Judaeo-Persica II: The Jewish-Persian Law Report from Ahwaz, A. D. 1020', in *Acta Orientalia*, 29: 49–60.

Bacher, W. (1904), 'Judaeo-Persian', in *The Jewish Encyclopedia.* Vol. VII, 313–317.

Bausani, Alessandro (1960), 'La letteratura neopersiana', in *Storia della letteratura persiana*. Milano: Nuova Accademia (Reprint Roma: Nuova Cultura 2011, La Sapienza Orientale – Ristampe).

Bausani, Alessandro (1978), 'Incontri linguistici in Asia. Esperienze di un orientalista', in *Lingue a contatto nel mondo antico. Atti del Convegno della Società Italiana di Glottologia, Napoli 12 e 13 maggio 1978*. Pisa: Giardini, 11–33.

Bivar, Adrian David H. (1986), 'Kitābāt. 9. Iran and Transoxiana', in *Encyclopedia of Islam*. Second edition. V, 228–231.

de Blois, François (2006), 'Glossary to the New Persian texts in Manichaean script', in François de Blois, Erica C.D. Hunter and Dieter Taillieu (eds), *Dictionary of Manichaean Texts. Vol. II: Texts from Iraq and Iran (Texts in Syriac, Arabic, Persian and Zoroastrian Middle Persian)*. Turnhout: Brepols, 89–120.

Bombaci, Alessio (1966), *The Kūfic Inscription in Persian Verses in the Court of the Royal Palace of Mas'ud III at Ghazni*. Rome: Istituto per il Medio ed Estremo Oriente.

Browne, Edward G. (1895), 'Some Notes on the Poetry of the Persian Dialects', in *Journal of the Royal Asiatic Society*, 773–825.

Filippone, Ela (2011), 'The Language of the Qor'ān-e Qods and Its Sistanic Dialectal Background', in Mauro Maggi and Paola Orsatti (eds), *The Persian Language in History*. Wiesbaden: Reichert, 179–235.

Gindin, Thamar E. (2007), *The Early Judaeo-Persian Tafsīrs of Ezekiel: Text, Translation, Commentary*. Volume I: *Text*. Vienna: Österreichische Akademie der Wissenschaften.

Gnoli, Gherardo (1964), *Le iscrizioni giudeo-persiane del Ġūr (Afghanistan)*. Rome: Istituto Italiano per il Medio ed Estremo Oriente.

Grohmann, Adolf (1971), *Arabische Paläographie*. II. Teil. *Das Schriftwesen. Die Lapidarschrift.* (Forschungen zur islamischen Philologie und Kulturgeschichte. Band II), Vienna: Hermann Böhlaus.

Hägg, Thomas and Bo Utas (2003), *The Virgin and her Lover: Fragments of an Ancient Greek Novel and a Persian Epic Poem*. Leiden, Boston: Brill.

Hashabeiky, Forogh (2005), *Persian Orthography. Modification or Changeover? (1850-2000)*. Stockholm: Uppsala University Library.

Henning, Walter B. (1957), 'The inscriptions of Tang-i Azao', in *Bulletin of the School of Oriental and African Studies*, 20: 335–342.

Henning, Walter Bruno (1958), 'Mitteliranisch', in *Handbuch der Orientalistik*, Erste Abt., Band IV: Iranistik, 1. Abschnitt: *Linguistik*. Leiden, Cologne: Brill, 20–130.

Henning, Walter Bruno (1962), 'Persian Poetical Manuscripts from the Time of Rūdakī', in *A Locust's Leg. Studies in Honour of S.H. Taqizadeh*. London: Percy Lund, Humphries & Co., 89–104.

Lazard, Gilbert (1963), *La langue des plus anciens documents de la prose persane*. Paris: Klincksiek.

Lazard, Gilbert (1968), 'La dialectologie du Judéo-Persan', in *Studies in Bibliography and Booklore*, 8: 77–98 (= Lazard, La formation de la langue persane, pp. 27–48).

Lazard, Gilbert (1971), 'Les origines de la poésie persane', in *Cahiers de civilisation médiévale* (Xe-XIIe siècles), 14/4: 305–317.

Lazard, Gilbert (1975), 'The Rise of the New Persian Language', in R. N. Frye (ed.), *The Cambridge History of Iran, IV: The Period from the Arab Invasion to the Saljuqs.* Cambridge: Cambridge University Press, 595–632.

Lazard, Gilbert (1986), 'Les prépositions pad et bē (ō) en persan et en pehlevi', in Rüdiger Schmitt and Prods Oktor Skjaervo (eds), *Studia grammatica iranica. Festschrift für Helmut Humbach*. Munich: R. Kitzinger, 245–255 (=Lazard, *La formation de la langue persane*, pp. 149–156).

Lazard, Gilbert (1988), 'Remarques sur le fragment judéo-persan de Dandān-Uiliq', in W. Sundermann, J. Duchesne-Guillemin and F. Vahman (eds), *A Green Leaf. Papers in Honour of Professor Jes P. Asmussen*. Leiden: Brill, 205–209 (=Lazard, La formation de la langue persane, pp. 157–161).

Lazard, Gilbert (1990), 'Lumières nouvelles sur la formation de la langue persane: une traduction du Coran en persan dialectal et ses affinités avec le judéo-persan', in Shaul Shaked and Amnon Netzer (eds), *Irano-Judaica II*. Jerusalem, 184–198 (=Lazard, La formation de la langue persane, pp. 107–121).

Lazard, Gilbert (1995), *La formation de la langue persane*. Paris: Peeters.

Lazard, Gilbert (2009), 'Qu'est devenue la préposition ō?', in Eric Pirart and Xavier Tremblay (eds), *Zarathushtra entre l'Inde et l'Iran: Études indo-iraniennes et indo-européennes offertes à Jean Kellens à l'occasion de son 65e anniversaire*. Wiesbaden: Reichert, 169–176.

Lazard, Gilbert (2014), 'La dialectologie du persan préclassique à la lumière des nouvelles données judéo-persanes', in *Studia Iranica*, 43 (1): 83–97.

MacKenzie, David Neil (1968), 'An Early Jewish-Persian Argument', in *Bulletin of the School of Oriental and African Studies*, 31 (2): 249–269 [Reprinted in Carlo G. Cereti and Ludwig Paul (eds), Iranica diversa. Roma 1999, vol. II, 315–337 + Addenda et corrigenda, 671–673].

MacKenzie, David Neil (2003), 'The Missing Link', in Ludwig Paul (ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*. Wiesbaden: Harrassowitz, 103–110.

Margoliouth, D.S. (1903), 'Early Documents in the Persian language', in *Journal of the Royal Asiatic Society,* 761–765.

Meier, Fritz (1981), 'Aussprachefragen des älteren Neupersisch', in *Oriens*, 27–28: 70–176 [Reprinted in E. Glassen and G. Schubert (eds), *Bausteine. Ausgewählte Aufsätze zur Islamwissenschaft*. Istanbul 1992, vol. II. 1057–1164].

Minorsky, Vladimir (1942), 'Some Early Documents in Persian', in *Journal of the Royal Asiatic Society*, 184–194.

Müller, F. W. K. (1915), 'Ein syrisch-neupersisches Psalmenbruchstück aus Chinesisch-Turkistan', in Gotthold Weil (ed.), *Festschrift Eduard Sachau zum siebzigsten Geburtstage*, Berlin: Reimer, 215–222.

Muwaffaq b. ʿAlī al-Hirawī (1972), *Das Buch der Grundlagen über die wahre Beschaffenheit der Heilmittel: vollständige Faksimile-Ausgabe des Codex Vindobonensis A. F. 340 der Österreichischen Nationalbibliothek.* Einführung Charles H. Talbot, Franciscus R. Seligmann, Graz: Akademische Druck- u. Verlagsanstalt.

Muwaffaq bin ʿAlī al-Hirawī, Abū Manṣūr (2009), *Kitāb al-Abniya ʿan Ḥaqāyiq al-Adwiya* […]. Facsimile Copy of the Original Manuscript A. F. 340 Austrian National Library. Transcribed by ʿAlī bin Aḥmad Asadī Ṭūsī. Copied 447 Hijri. Tehrān: Mirās̲-e maktub & Vienna: Austrian Academy of Sciences Press, 1388.

Orsatti, Paola (1993), 'Le manuscrit islamique : caractéristiques matérielles et typologie', in M. Maniaci and P. F. Munafò (eds), *Ancient and Medieval Book Materials and Techniques* (Erice, 18–25 September 1992) (Studi e Testi 358, Vol. II). Città del Vaticano: Biblioteca Apostolica Vaticana, 269–331.

Orsatti, Paola (2007), *Appunti per una storia della lingua neopersiana. Parte I: parte generale – fonologia – la più antica documentazione*. Rome: Nuova Cultura (La Sapienza Orientale – Manuali).

Orsatti, Paola (2008), 'Minoranze religiose e scrittura nel primo Iran islamico', in D. Bredi, L. Capezzone, W. Dahmash and L. Rostagno (eds), *Scritti in onore di Biancamaria Scarcia Amoretti*. Rome: Edizioni Q. Vol. III, 867–885.

Paul, Ludwig (2003), 'Early Judaeo-Persian in a historical perspective: The case of the prepositions *be*, *u*, *pa(d)*', and the suffix *rā*', in Ludwig Paul (ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*. Wiesbaden: Harrassowitz, 177–194.

Paul, Ludwig (2013), *A Grammar of Early Judaeo-Persian*. Wiesbaden: Reichert.

Perry, John R. (1991), *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending*. Costa Mesa: Mazda.

Perry, John R. (1995), 'Lexical doublets as a derivational device in Persian: The Arabic feminine ending', in *Acta Orientalia Academiae Scientiarum Hungaricae*, 48: 127–153.

Perry, John R. (2002), 'Arabic Language. V. Arabic Elements in Persian', in *Encyclopaedia Iranica*. Online edition (accessed on 10/10/2014).

Pisowicz, Andrzej (1985), *Origins of the New and Middle Persian Phonological Systems*, Kraków: Nakładem Uniwersytetu Jagiellońskiego.

Provasi, Elio (2011), 'New Persian Texts in Manichaean Script from Turfan', in Mauro Maggi and Paola Orsatti (eds), *The Persian Language in History*. Wiesbaden: Reichert, 119–177.

Rapp, Eugen Ludwig (1967), 'The Date of the Judaeo-Persian Inscriptions of Tang-i Azao in Central Afghanistan', in *East and West*, 17 (1–2): 51–58.

Rivāqī, ʿAlī (1985), Qurʾān-i quds. Kuhantarīn bargardān-i Qurʾān ba fārsī? [The Holy Koran. The most ancient translation of the Koran into Persian?]. Tehrān: Muʾassasa-yi farhangī-yi Šahīd M. Rivāqī, 1364. 2 vols.

Richard, Francis (1981), 'Un lectionnaire persan des Évangiles copié en Crimée en 776 H./1374', in *Studia Iranica*, 10: 225–245.

Šams al-Dīn Muḥammad b. Qays-i Rāzī (1981), *Al-Mu'jam fī ma'āyīr aš'ār al-'ajam* [The clarified work about the criteria of the poetry of the Persians]. Ed. Muḥammad Qazwīnī & Muḥammad-Taqī Mudarris Raḍawī. Tehrān: Zavvār, 1360, 3ʳᵈ edn.

Scarcia, Gianroberto (1963), 'A Preliminary Report on a Persian Legal Document of 470-1078 found at Bāmiyān', in *East and West* N.S. 14 (1–2): 73–81.

Scarcia, Gianroberto (1966), 'An Edition of the Persian Legal Document from Bāmiyān', in *East and West* N.S. 16 (3–4): 290–295.

Shaked, Shaul (1971), 'Te'uda qara'it qeduma be-parsit yehudit' [An ancient Karaite document in Judaeo-Persian], in *Tarbiz*, 41 (1): 47–58.

Shaked, Shaul (2009), 'Classification of Linguistic Features in Early Judeo-Persian Texts', in Werner Sundermann, Almut Hintze and François de Blois (eds), *Exegisti monumenta. Festschrift in Honour of Nicholas Sims-Williams*. Wiesbaden: Harrassowitz, 449–461.

Shokri-Foumeshi, Mohammad (2014), 'Wāžahā-yi daxīl-i 'arabī dar kuhantarīn matnhā-yi fārsī-yi naw bar asās-i dast-nivīshā-yi mānawī-yi Tūrfān' [The Arabic loanwords in the earliest New Persian texts in Manichaean manuscripts from Turfan], in *Tārīx-i Īrān/Iran History* 72 (5). Autumn 1392/2013, published Spring 1393, 161–210.

Sims-Williams, Nicholas (2011), 'Early New Persian in Syriac script: Two texts from Turfan', in *Bulletin of the School of Oriental and African Studies*, 74 (3): 353–374.

Šīrvānī, Kobrā (1974), *Barrasī-yi imlāyī-yi dastnivīs-ī az* Tafsīr-i Sūrābādī [Orthographical analysis of a manuscript of the *Tafsīr-i Sūrābādī*]. Tehrān: Entešārāt-e Farhangestān-e zabān-e Irān, 1353.

Skjaervo, Prods Oktor (1996), 'Aramaic Scripts for Iranian Language', in P.T. Daniels and W. Bright (eds), *The World Writing Systems*. New York, Oxford: Oxford University Press, 515–535.

Sundermann, Werner (1974), 'Einige Bemerkungen zum syrisch-neupersischen Psalmen-bruchstück aus Chinesisch-Turkistan', in Ph. Gignoux and A. Tafazzoli (eds), *Mémorial Jean de Menasce*. Louvain: Imprimerie orientaliste, 441–452.

Sundermann, Werner (2003), 'Ein manichäischer Lehrtext in neupersischer Sprache', in Ludwig Paul (ed.), *Persian Origins: Early Judaeo-Persian and the Emergence of New Persian*. Wiesbaden: Harrassowitz, 243–274.

Tafazzoli, Ahmad (1974), 'Some Middle-Persian Quotations in Classical Arabic and Persian Texts', in Ph. Gignoux and A. Tafazzoli (eds), *Mémorial Jean de Menasce*, Louvain: Imprimérie orientaliste, 337–349.

The Tārikh-e Sistān, transl. Milton Gold. Rome: Istituto Italiano per il Medio ed Estremo Oriente, 1976.

Utas, Bo (1968), 'The Jewish-Persian Fragment from Dandān-Uiliq', in *Orientalia Suecana*, 17: 123–136.

Xānlarī, Parvīz Nātil (1986), *Tārīx-i zabān-i fārsī* [History of the Persian language], 1ˢᵗ revised edn. Tehrān: Firdaws, 1365. 3 vols.

Yarshater, Ehsan (1974), 'The Jewish Communities of Persia and their Dialects', in Ph. Gignoux and A. Tafazzoli (eds), *Mémorial Jean de Menasce*. Louvain: Imprimerie orientaliste, 453–466.

Zhang, Zhan and Shi, Guang (2008) [2009], 'Yijian xin faxian Youtai Bosiyu xinzha de duandai yu shidu' [The dating and interpretation of a newly discovered Judaeo-Persian letter], in *Dunhuang Tulufan yanjiu*, 11: 71–99, i color pl.