On the Origin of Qur’ānic Arabic

Mark Durie, Arthur Jeffery Centre for the Study of Islam
Melbourne School of Theology

Abstract

Previous research into the origins of the dialect of Arabic which provided the standard for the Qur’ānic consonantal text has thrown up two puzzles: a puzzle from above, and a puzzle from below. The puzzle from above is that the Muslim philologists who searched earnestly for the Qur’ānic dialect were unable to identify any of the dialects known to them as the source of Qur’ānic Arabic. The puzzle from below is why, among the many thousands of pre-Qur’ānic ancient Arabic inscriptions, spread over a vast region, there are so few inscriptions which could reflect a precursor to Qur’ānic Arabic.

The solution to both puzzles, it is argued, is that Qur’ānic Arabic, as reflected in its rasm, or consonantal skeleton, developed directly from the Arabic of the Nabataeans. The solution to the puzzle from below is that the Nabataeans left many inscriptions, but in their preferred written language, Aramaic, not their mother tongue Arabic. The solution to the puzzle from above is that the Muslim philologists were looking for a Bedouin source for the Qur’ānic text, but Nabataean Arabic’s linguistic legacy was to be found in dialects spoken by peasants in the Southern Levant.

Recent research by Al-Jallad has documented pre-conquest Southern Levantine Arabic dialects in the Nabataean direct sphere of influence. These findings confirm the hypothesis that the Nabataean Arabic supplied the linguistic variety of the Qur’an.

The Puzzle from Below: Pre-Islamic Evidence

For a millennium leading up to and including the 6th century CE, tens of thousands of inscriptions and graffiti in a variety of South Arabian scripts attest to the existence in the Arabian Peninsula of a variety of early Arabic dialects. A key question is: which of the attested ancient Arabian linguistic varieties, if any, was the precursor to Qur’ānic Arabic? And if there was no written precursor, how was it that a dialect with no written history came to supply the Arabic of the Qur’an?

A salient feature of the Qur’ānic Arabic consonantal skeleton, is the use of the definite article, ’l-. In contrast, most of the attested ancient early Arabic inscriptions, sometimes referred to as ‘Ancient North Arabian’, are in varieties which use a variety of other definite articles, hn-, h-
or 'l- (Macdonald, 2004). Only a few inscriptions use 'l-. Evidence for the al- form appears very early. Herodotus, writing c. 440 BCE, almost a millennium before the Qurʾān, reported that the deity of the Arabs was known as Άλιλάτ, which must reflect Arabic al-ʾilāt ‘the goddess’.

M.C.A. Macdonald (2009c:179) considered it to be a riddle of Old Arabic studies that tens of thousands of inscriptions and graffiti are found over most of the Arabian Peninsula and beyond, in a wide variety of scripts, of dialects which rarely if ever use 'l-. In contrast the number of inscriptions using the 'l- variety dialect(s) is so tiny that Macdonald (2008:464) commented, ‘… it is virtually certain that Arabic [here Macdonald means the precursor dialect to Qurʾānic and Classical Arabic] was a purely spoken, rather than a habitually written, language for most of its pre-Islamic history’. He puzzled over the lack of inscriptions:

> There seems to be no practical reason why Arabic should have remained an unwritten language for so long, particularly in areas which witnessed massive epigraphic activity in related dialects and other languages. (Macdonald, 2009d:21)

It is argued here that the answer to Macdonald’s puzzle is that Nabataean Arabic, in which the 'l- article had become standard, was the precursor to Qurʾānic Arabic, and the Nabataeans were so accustomed to writing in Aramaic that they used this script and language when they left graffiti behind them, rather than their spoken Arabic vernacular. It is consistent with this explanation that a significant majority of the Arabic inscriptions which do use 'l- (except for

1. Al-Jallad (2017) set the definition of Old Arabic on a proper footing by identifying shared distinguishing innovations. Using these identifying features he concluded that Safaitic and Ḥismaic, two of the varieties formerly known as ‘Ancient North Arabian’, can be considered varieties of Old Arabic, but he excluded other ‘Ancient North Arabian’ varieties such as Taymanitic and Thamudic from the group of Old Arabic dialects. Ernst Knauf (2011:207) have previously objected to ‘Old Arabian’, suggesting ‘Old Arabic’ was the better label. Retsö (2013:438), also pointed out that some modern dialects of Arabic spoken in Yemen have an- and m- articles and on this basis concluded that the al- article is ‘not a panarabic feature’.

2. Inscriptions from the same period as Herodotus have been found in northeast Egypt attesting to the name of the goddess as hn- 'lt, with the same meaning, and referring perhaps to the same goddess, but using an alternative hn- form of the article (Macdonald, 2004:517).
Graeco-Arabica, which is discussed below) are in the Nabataean script (Macdonald, 2008). There are only a small handful of inscriptions with ‘l- in South Arabian scripts (Macdonald, 2008; Al-Jallad 2015).

Nabataeans and nabaṭī
Srabo described the Nabataeans as the people whose capital was Petra (Greek ‘rock’), which was known as Raqmu to the Nabataeans (Avi-Yonah and Gibson, 2007; Retsö, 2011). The terms Naḥbatāi and Nabataei are borrowed from Arabic nabaṭī (pl. anbāṭ). This ancient title derives from the Arabic root n-b-ṭ, which refers to water welling or oozing from the ground, and, by extension, to the extraction of ground water by digging a well. Thus nabaṭī originally meant ‘well-digger’. Inscriptions have been found, not only in Nabataean Aramaic, but also

3. Cf Strabo’s Geography, Book 16, Chapter 4.

4. A related meaning seems to have persisted well past the creation of the Qur’ān. Lane (1863:2759-2760) reported that nabaṭī was described by medieval Arab lexicographers as a derogatory term, referring to a ‘mixed’ or base people who made a living by agriculture, i.e. to sedentary peoples, in contrast to nomads, which is consistent with its derivation (Joukowsky, 2007:716). Ammianus Marcellinus, writing in the 4th century CE, reports in Arabs’ own words their distaste for agriculture: ‘the worst evil that can befall a people, and after which no good can come, is that their necks are bent’ (Ammianus Marcellinus, 1950:§14.4). Jeffery (1938:27) observed that among the Arabs the term nabaṭī was also used for ‘many communities in Syria and Iraq’, namely those who lived by agriculture, and Arab lexicographers, in addition to suryanī, used for Syriac speakers in the north, used the term nabaṭī to refer to southern varieties of Aramaic. Bar Hebraeus, writing in the 1200’s CE, distinguished three varieties of Aramaic. It was the Aramaic spoken in the mountains of Assyria and southern Iraq which was identified as al-nabaṭiyya. (Hoyland, 2008:52). H. Fahd reported (1993:836) that to be ‘nabatised’ (istanbaṭī) meant to become sedentary and practice agriculture while to be ‘arabised’ (isat ṭabū) meant adopting a Bedouin (nomadic) pastoral lifestyle. This reflects an important distinction between the two main kinds of Arab-speaking communities: the desert dwellers or ‘true’ Arabs, and the sedentary Arabs who had settled and pursued agriculture.
in early Arabic, in South Arabian scripts, in which people are identified as nbṭ, nbṭy or nbṭw5 (Macdonald, 1993:307, fn.28). It is not always clear whether this is intended as an ethnic or political designation, i.e. ‘Nabataean’ or a description of someone’s mode of subsistence through agriculture. One inscription which must refer to the Nabataean kingdom is the Safaitic w g{l}s¹ mn ‘ dmt s‘nt mt mlk nbṭ ‘and he halted on account of the downpour the year the king of Nabataea died.’

Macdonald has argued that Arabic existed – and indeed flourished – for centuries among the Nabataeans of Petra in a stable diglossic environment in which first Aramaic and then Greek were used for formal written communications, while Arabic was the mother-tongue of the community with a vital oral tradition, including liturgical texts, a tradition which was not committed to writing.7 A variety of evidence supports this conclusion. Strabo referred to the Nabataeans as Arabs,8 as did Diodorus Siculus.9 Josephus also called the Nabataeans ‘an Arabian nation’ (Josephus, n.d. :1.43), and Petra, the Nabataean capital, the ‘royal seat of the king of Arabia’ (Josephus, n.d. :3.395). Moreover Josephus called the region to the south of Judea ‘Arabia’,10 which included the southern reaches of the Dead Sea where Petra was located. Trajan renamed Nabataea Provincia Arabia when he annexed it in 106 CE and it was

---

5. In Old Arabic inscriptions w and y are often interchangeable.
7. In coming to this view, Macdonald reversed his earlier rejection (2003) of the proposition that the Nabataeans were Old Arabic speakers. The observation that the Nabataeans spoke Arabic as their mother tongue is of course not inconsistent with the presence in Nabataean settlements of speakers of other languages, such as Greek or Hebrew.
8. Strabo’s Geography, Book 16, Chapter 4.
10. Josephus reported that one could see ‘Arabia’ from towers of Jerusalem (n.d. :4.202), and the Herodium lay ‘so near to Arabia’ (n.d. :4.368).
the custom of the Romans to name provinces after ethnicities (Macdonald, 2009b:298).
Moreover it is clear from their Aramaic inscriptions that the Nabataeans had Arabic names.\textsuperscript{11}
Epiphanus, writing in the 4\textsuperscript{th} century CE, refers to the Petraeans singing hymns in Arabic to
their virgin goddess called ‘Kaabou’ (Hoyland, 2008:54), and Macdonald (2006:94) has
suggested that two lines of Arabic poetry recorded in an inscription in the Negev, attached to
a text in Nabataean Aramaic, was an excerpt from an oral liturgy in praise of the deified king-
god Obodas. This is evidence of the liturgical use of Arabic poetry in a Nabataean context.

There are also known to be a large number of Arabic loanwords in Nabataean Aramaic
(Macdonald, 2009b:397; O’Connor, 1986; Greenfield, 1992; Morgenstern, 1999; Beyer
2004), as well as evidence of Arabic syntactic influence on written Nabataean Aramaic
(Healey, 1993; al-Hamad, 2014; Gzella, 2004), which is even stronger evidence than lexical
borrowing for a substrate relationship between Arabic and Nabataean Aramaic. Furthermore,
the place names in 6\textsuperscript{th} century Petra Papyri, written in Greek, are predominantly Arabic
(Hoyland, 2008:57), and there are also known to be a large number of Aramaic loans in
Classical Arabic (Fraenkel, 1886).

Further evidence that the Nabataeans were Arabic-speaking is the existence of a collection of
legal papyri dating from the late first and early second centuries CE, in which Jewish
Aramaic documents include lists of Jewish Aramaic terms followed by Hebrew equivalents.
In Nabataean Aramaic documents found alongside them, Aramaic terms are followed by their
Arabic equivalents (Macdonald, 2009a:19).

It is hardly surprising that settled Arabs, who needed to develop administrative infrastructure
which demanded written texts, looked to imperial Aramaic for a language of writing.
However all the evidence suggests that those who used the ‘Nabataean’ script to write were
mother-tongue Arabic speakers:

‘… we should not assume that all those who wrote or commissioned a text in
what we call the “Nabataean” language and script thought of themselves as
ethnically or politically “Nabataean,” any more than someone who writes in

\textsuperscript{11} It may also be significant that Petra, the capital city of the Nabataeans, was located in the
\textit{Wādī ‘Arabah}, a geographical feature which runs – in its ancient denotation – from the Sea of
Galilee in the north to the Gulf of Aqaba in the south. ‘\textit{Arabah} as a place name attested as far
back as the Pentateuch (Deuteronomy 2:8). It is formed from the same radicals ‘\textit{-r-b} as the
name ‘\textit{arab}, the only phonological difference being the feminine ending ‘\textit{-ah}, which is typical
for place names.
the language we call “English” is necessarily “English” by nationality.’ (Macdonald, 2003:39)

What we also know is that Arabic orthography developed organically out of a cursive form of Nabataean Aramaic script (Nehmé, 2009; Macdonald, 2009a:21), such as would be needed by traders for their account-keeping on soft materials like papyrus (Knauf, 2011:231), and there are some important early Arabic inscriptions in the Aramaic script: a pagan inscription from ‘Ēn ‘Avdat, and the Namāra inscription, dated 328 CE.  

We conclude that an oral Nabataean Arabic substrate culture endured for centuries alongside Aramaic and then Greek literacy. Then speakers of Nabataean Arabic transitioned into using their Arabic mother-tongue as their preferred medium of written communication, adapting the Nabataean Aramaic script for this purpose. We can tell this process was a gradual one from the way in which a fully developed Arabic hand-written script emerges into the light of day over centuries. Macdonald writes:

> This means that we have to assume an extensive, and possibly increasing, use of writing on soft materials in the Nabataean script throughout the fourth to seventh centuries, since only this could produce the transitional letter forms and ligatures we see first in the “Nabataean” or “transitional” graffiti of the fifth century … then in the early Arabic inscriptions of the sixth and seventh centuries, and the earliest Arabic papyri of the mid-seventh. (Macdonald, 2009a:21; see also Nehmé, 2009)

As gradual as this process was, it was presumably the necessity to administer in Arabic rather than Aramaic or Greek which led to the emerging Nabataean-Arabic script being used to record the texts of Arab-led government administration, and of the religion of Islam.  

12. See Mascitelli, 2006: 121-29, 152 for references to the considerable literature on these inscriptions.

13. There is an interesting footnote to this account of the rise of Arabic. When Ptolemy famously referred to the Sarakenoi ‘Saracens’ of North West Arabia this term was mostly likely derived from the root sh-r-q, which in Classical Arabic means ‘east’ or ‘go east’ (Macdonald, 2009e). In Ancient Northern Arabian (and in some Bedouin dialects still today) the cognate ʾṣrq meant ‘to migrate to the inner desert’, irrespective of the compass direction (Macdonald, 2004:529). Traditionally interpreted as meaning ‘of the east’, in fact sh-r-q mostly likely had as its original meaning ‘of the inner desert’. The Saracens were not
The Puzzle from Above: Evidence of Muslim Philologists

In the discussion to this point we have considered the emergence of Arabic in the light of pre-Qurʾānic evidence, and particularly the comparative scarcity in inscriptions of evidence for a dialect precursor for Qurʾānic Arabic. The suggested explanation is that the precursor to Qurʾānic Arabic was the Arabic spoken by the Nabataeans, and the reason Nabataean Arabic is not widely attested by inscriptions is that the Nabataeans usually left written records in Aramaic rather than their mother tongue. We shall now consider the implications of the ‘view from above’ evidence from later, post-Qurʾānic Arabic sources, namely the inability of later investigators to locate the dialect in which the Qurʾān was written.

The literature devoted to the problem of the origin of Arabic on the basis of later Islamic sources was summarized by C. Rabin (1955). Medieval Muslim philologists had devoted considerable efforts to studying and codifying Arabic, including seeking out evidence from Arabic dialects, and their evidence was carefully investigated by scholars in the first half of the 20th century. Muslim philologists took the language of the Qurʾān as the ʿarabīya gold standard. At the same time they also believed that Qurʾānic Arabic was the same variety as the Arabic of pre-Islamic poetry (Rabin, 1955:21-22). Early Muslim scholars drew upon this poetic corpus to help standardize classical Arabic, supplementing it with information drawn from speakers of bedouin dialects, whom they regarded as speakers of ‘pure’ Arabic (Rabin, 1951:12). Rabin commented wryly that the Bedouin ‘could not speak wrong Arabic, even if he had wanted to’ (Rabin, 1951:18).

Pre-Qurʾānic Arabic poetry was performed in court centres, such as the Lakhmid court at Ḥira in Mesopotamia (Rabin, 1955:31), and the Ghassanid court at Jābiyah in the Levant. We know also that poets came from a variety of different Arabic dialect areas, and their poetic

---

14. Islamic poetry is only known from manuscript sources which date after the birth of Islam and the form we have them in shows signs of extensive editing (Rabin, 1955:21).
idiom was standardised. For this reason scholars have called the variety used by the poets a poetic koiné (Blachère, 1952: 66-82; Rabin, 1955:24; Versteegh, 1984:1).15

The medieval Muslim philologists naturally assumed that the language of the Qurʾān was the dialect of Muhammad’s tribe, the Quraysh. However Rabin (1955:21-22) reported that by the 1940’s a consensus had developed among western scholars that although the language of the Qurʾān was accepted to have been based on the poetic koiné, this could not have been the Meccan dialect: ‘there is substantial agreement among European scholars that to most or all of those who employed it for writing poetry, Classical Arabic was to some extent a foreign idiom which had to be acquired’ (Rabin, 1951:17; see also Fleish, 1947:97-101; Blachère, 1947:156-169; and Nöldeke and Schwally, 1919:57-58). This conclusion was based on a close examination of extensive reports in Muslim sources concerning Bedouin dialects (Rabin, 1951:6-24). The discrepancies were well known to medieval Muslim scholars. For example Ibn ʿAbdalbaṭr (d. 463/1071) had pointed out that certain linguistic features of the Qurʾān, such as the preservation of the glottal stop, were inconsistent with Ḥijāz dialects (Rabin, 1951:19).

In response, Muslim philologists hypothesized that the Qurʾān combined features of a wide variety of Bedouin dialects. For example Suyūṭī commented that ‘the Qurʾān contains words from all Arab dialects’ (Rabin, 1951:19). Of course this is precisely what one might expect to find if it was not composed in any of the Bedouin dialects. Rabin remarks that ‘it is hard to understand that the scholars never seem to have realized’ that ‘the spoken language of the Bedouin was different from the Classical idiom’ (Rabin, 1951:18).16

Two questions arise from these findings. One concerns the origin of the poetic koiné, and another concerns the Qurʾān’s use of and relation to the koiné. Rabin remarked that ‘no progress seems to have been made in recent years in solving the problem of the place of origin of the poetic koiné’ (Rabin, 1955:31). This is the ‘puzzle from above’. The second question concerns Muḥammad as the reciter of the Qurʾān and why he was addressing Meccans using the poetic koiné of the Arab courts. Rabin put it this way:

What reasons caused Muhammad to address his fellow townspeople in a language which originated, and was at the time used, for narrowly

15. This is not the only attested example of a poetic koine: see Blachère (1952: 80-91) who gives numerous other examples.

16. Rabin even suggested that ‘the dialect of Quraish must have been more unlike the Classical than the present-day colloquials’ (Rabin, 1955:26).
circumscribed purposes in Bedouin society, and that mainly in regions fairly remote from Mecca? (Rabin, 1955:27)

Setting aside the question of whether the original setting of the Qurʾān was Mecca, it is proposed here that the far-reaching Nabataean trading network, established for centuries, could have provided the context for an Arabic koiné to be developed, based on the Nabataean variety of Arabic. In the context of Arabic dialect diversity, and supported by the prestige of Nabataean power, it is proposed Nabataean Arabic supplied the precursor for a poetic trans-regional standard to develop, which continued to be used by poets in far-flung Arab courts centuries after Nabataean power had waned.\(^\text{17}\) This koiné was very similar to the Arabic variety used by the Qurʾānic messenger for the recitations which became the Qurʾān. Because Nabataean Arabic was used by the traders who ranged far and wide across the Arabic speaking region, the Qurʾānic idiom would have been understood by speakers of a wide variety of dialects.

It seems improbable that Bedouin Arabs would have considered the poetic koiné a different language from their own native dialects.\(^\text{18}\) They would have simply recognized it as Arabic. This makes sense of the Qurʾān’s claim that the Messenger was communicating bi-lisānī qawmi-hi ‘in the language of his people’ (Q14:4) using ‘clear (mubīn) Arabic’ (Q16:103; Q26:195). It was clear because it was in a variety of Arabic understood by all, and not necessarily because it was necessarily in the local dialect of its first audience.

It is an intriguing footnote to this discussion that the term nabaṭī came to refer to the indigenous oral performance poetry of the Bedouins (Bannister, 2014:117-121), a usage which continues to this day. This use of nabaṭī perhaps owes its origins to a much earlier period when the poetic koiné was identified as a ‘Nabataean’ variety, and, by metonymy, the linguistic variety in which poetry was performed lent its name to the poetic art itself. P. G.

---

\(^{17}\) We recall the previously noted evidence of a Nabataean Arabic liturgical poetic tradition (Hoyland, 2008:54; Macdonald, 2006:94).

\(^{18}\) Consider for example that for more than a thousand years speakers of Germanic dialects have referred to both their local dialects and the various standard varieties – which have changed over the centuries – by the same name: in Old High German diutisc and Middle High German diutsch. (The cognate þeodisc was also used to refer to Old English.)
Emery reports (1993:838) that ‘in spite of linguistic differences, nabaṭī poetry and Classical Arabic poetry share many prosodic, thematic and functional similarities’.

The Linguistic Evidence

So far we have considered the ‘puzzle from above’, and the ‘puzzle from below’, and suggested, based on circumstantial evidence, that the best candidate for a precursor dialect to the poetic koiné and the Arabic of the Qurʾān was Nabataean Arabic, which would have been in use throughout the Nabataean trade networks. In this section we compare linguistic features of Qurʾānic Arabic and Nabataean Arabic.

There are two kinds of linguistic evidence which link Qurʾānic Arabic with the Nabataean variety. One is the very fact that the Arabic script developed from the Nabataean script. This implies that it was bilinguals who knew both Nabataean Aramaic and Arabic who established the orthographic conventions of Arabic. Another kind of evidence is the existence of specific shared linguistic features which link Nabataean Arabic with the Arabic of the Qurʾānic *rasm*, or unpointed consonantal skeleton.

We have already noted that the use of the definite article *ʾal-* was a Nabataean feature. Early Arabic dialects, including epigraphic Safaitic, used a variety of forms of the definite article, including *h-, hn-, ʾ-, and ʾl-, or they dispensed with the article altogether (e.g. Ḫismaic). Although some earlier researchers had identified the *ʾl-* form as a distinctive feature of the precursor variety to Classical Arabic, Al-Jallad (2014: 5-6, 13-15) has argued that the *ʾl-* article is not an innovation that can be used to distinguish a separate genetic grouping within the Arabic varieties. Instead, Al-Jallad argues that all the diverse forms of the Old Arabic definite article were diffused across dialects.

Nevertheless, although the *ʾl-* article may not be genetically significant as a diachronic marker, the frequency of its use is a distinguishing feature of Nabataean Arabic, where it is the only form of the definite article attested. On the other hand, in the Old Arabic dialects of the nomads the *ʾl-* article was non-standard (Al-Jallad 2014:14). So in its frequency, the *ʾl-*
article does link the language of the Qurʾān to Nabataean Arabic, even if it is not a reliable genetic marker of a diachronic development within Old Arabic.

There are other features which link the Qurʾānic rasm with the Graeco-Arabica – Arabic written in Greek letters – of Arabia Petraea, the Nabataean heartland of the Southern Levant. After they were conquered by the Romans in 106 CE the Nabataeans gradually shifted to using Greek as the language of formal written communications. There is a considerable number of subsequent Graeco-Arabica inscriptions and papyri from the region of the Nabataean settlements in the Southern Levant, and, as Al-Jallad (2017: 105) has observed, the linguistic features of these materials ‘usually agree’ with the Arabic substratum of Nabataean against the ‘Ancient North Arabian’ epigraphy, particularly in the use of the definite article αλ, reflecting /ʾal/, and the feminine ending -α, reflecting /-ah/ in contrast to /-at/. The dialects of the Southern Levant can broadly speaking be referred to as Nabataen.

The analysis in this section is made possible by Al-Jallad’s superb survey of Graeco-Arabica, of the Southern Levant. Al-Jallad (2017:99) defines this area as comprising: ‘southern Syria (areas including the Lejā, i.e. Trachonitis, Umm al-Jimāl, Bošrā, and the Ḥawrān), central and southern Jordan (including areas such as Moab, Edom, Petra, and the Ḥismā), and Israel (areas in the Negev such as Beersheba, Elusa, and Nessana).’ This region coincides with the area under Nabataean control before the Roman conquest of the Petra in 106 CE, which was renamed Arabia Petraea after conquest, and administered from Bosra, the former northern Nabataean capital. Given the political stability of Arabia Petraea after the 106 CE overthrow of the Nabataean kingdom, and the continuing influence of Bosra, which had been a major Nabataean city, it is not surprising that Al-Jallad’s 2017 study of Graeco-Arabica across this region points to a consistency and uniformity in dialect features which aligns the Arabia Petraea Graeco-Arabica corpus with Nabataean Arabic.

19. Greek was also adopted as the language of administration by the Ghassanid Arabs who had settled in the Levant (Macdonald, 2009d:24).

20. Nabataean remains have been identified at over 1,000 sites along the Nabataean trade routes connected Petra with Damascus in the North and the Hejaz in the South, and throughout the Negev stretching down to the port of Gaza (Joukowsky 2007:716).
The five features we shall consider here are word-final unstressed inflectional vowels (iʿrāb); the feminine ending -al/-ah; alif maqṣūrah; retention of the glottal stop, and the assimilation of the definite article al- to following coronals, or ‘sun letters’.

The arguments presented here rely upon two assumptions: a) that the spelling of the Qurʾānic rasm was phonemic at the time the Arabic script became fixed, and b) the rhyme patterns in the Qurʾān reflect the phonology of the dialect in which the Qurʾān was first composed and recited. We shall see that, in each of these respects, the Qurʾānic rasm agrees with what we know of the Arabia Petraea Graeco-Arabica, and where evidence is available, specifically with the Arabic or Petra. Al-Jallad has already acknowledged this, commenting that ‘The Graeco-Arabica [of the Southern Levant] generally agrees with Qurʾānic orthography’ (2017: 153). However the arguments presented here are more thorough, and also take into account evidence from Qurʾānic rhyme schemes for Qurʾānic phonology.

Loss of case endings (iʿrāb)

In the Qurʾānic script, unstressed inflectional short final vowels and ‘nunation’ are marked with diacritics: they are not represented in the rasm. These endings are known as iʿrāb because they were characteristic of Bedouin (aʿrāb) dialects, in contrast to settled urban dialects (Fleisch 1986, 1250). Although classical pre-Qurʾānic Arabic poetry incorporated iʿrāb endings in its rhyme schemes (Zwettler 1978, 147), the Qurʾān does not. This can be illustrated by Q1, the first surah, which has a rhyme scheme in -īn/-īm, but the inflectional endings added after this syllable, and marked by diacritics, vary between -u, -i and -a.

The indifference of Qurʾānic rhyme schemes to case endings shows that the endings had been lost to the dialect in which the Qurʾān was first recited. This agrees with what we know of Nabataean Arabic, which had lost case endings by the end of the first century BCE (Diem 1973), as had Southern Levantine Arabic by the 6th century CE (Al-Jallad 2017a, 165).21 Al-Jallad (2017a, 165) summarizes the dialect evidence as: ‘by the 6th century there can be no doubt as to the loss of case inflection, at least in Palaestina Tertia.’

---

21. Al-Jallad (2017:159) also reports that ‘there is no evidence for case inflection in the transcribed Arabic phrases in the non-literary papyri from Petra and Nessana’, and even though evidence of vocalic case endings survived in Graeco-Arabica, in the medial /o/ of genitive name compounds after it was lost in word-final position, by the 6th century, the Petra Papyri had lost case marking even in compounds (Al-Jallad 2017:165).
The -ah ending and tāʾ marbūṭah

The Proto-Arabic (and Proto-Semitic) *-at feminine ending is rendered as -h in the Qur’ānic rasm. In Nabataean Arabic the feminine ending *-at had changed to -ah in pausal (non-construct) position by the 2nd century BCE at the latest (Al-Jallad 2017a, 158–59), following a parallel change which had already taken place in Hebrew.22 Some Arabic loans into Nabataean pre-date this change while others post-date it (Al-Jallad 2017:157). Evidence for a terminum ante quem for this sound change is the Nabataean name Aretas attested in 2 Maccabees 5:8 (~ 124 BCE), which reflects Arabic ḥārethah. A late fifth century CE example is cited by Avner, Nehmé and Robin (2013:243) of the name ʿlḥb, found in a transitional Nabataean-Arabic inscription.

The change *-at > -āh became a regional feature of Southern Levantine Arabic. Al-Jallad concludes, on the evidence of Graeco-Roman inscriptions, that by the 4th Century CE this innovation had spread throughout the sedentary settlements of the Southern Levant. In contrast, ‘Ancient North Arabian; desert inscriptions do not show this change, retaining -at in all positions (Macdonald 2004, 498). Thus Al-Jallad (2017a, 158) describes the change to -ah as an isogloss which separates sedentary Arabic varieties from the dialects written in rock inscriptions by the desert nomads of the Ḥarrah.

In the rasm of the Qurʾān, the change *-at > -ah is regularized to all positions, including non-construct contexts. However the pronunciation of -ah was restored in non-pausal positions to –at by the addition of two superimposed dots – the pointing for t over the h – to give ّ (tāʾ marbūṭah). This implies that the change *-at > -ah had become regularized to all positions in the dialect which originally determined the orthography of the Qurʾānic rasm.

That this was not merely an orthographic convention, but the way the Qurʾān was originally recited, is apparent from Qurʾānic rhyme patterns, in which the feminine singular tāʾ marbūṭah endings rhyme indiscriminately with regular -h. Examples are Q80:12–13, where dhakara-hu ‘remember it’ (rasm: dh-k-r-h) rhymes with mukarramat-in ‘honored’ (rasm: m-k-r-m-h), and Q69:18–19, where khāfiat-un ‘secret’ (rasm: kh-alif-f-y-h) rhymes with kitābūah ‘my book’ (rasm: k-t-b-y-h). Assuming that case endings and final unstressed vowels had already been lost, the first rhyming pair would have originally been recited as dhakarah

22. A change of the feminine singular ending *-at > -āh is characteristic of southern Hebrew, but not of other Canaanite varieties. The Hebrew change is already reflected in the text of the Tanakh. This change may have diffused from Hebrew to Nabataean Arabic, since they were spoken in adjacent regions.

On the other hand, tā’ marbūtah never rhymes with regular tā’. Q88 illustrates the contrast, keeping tā’ marbūtah separate from tā’: Q88:8–16 rhymes consistently with tā’ marbūtah and then Q88:17–20 switches to a consistent rhyme with regular tā’. The two rhyme scheme are kept separate, which indicates that regular tā’ was not the same phoneme as tā’ marbūtah.

The evidence of the rhymes indicates that at the time the Qur’ān was first recited so-called tā’ marbūtah was actually pronounced as /h/ in all positions, which is how it is actually written in the rasm. Clearly the dialect in which the Qur’ān was originally recited had regularised this Southern Levantine feature throughout the whole paradigm, including in non-pausal positions. This regularisation represents an advanced stage in the replacement of *-at with –ah, and as such it most likely to have taken place in the Southern Levant, where the sound change was first attested in Nabataean Arabic.

The change -at -ah distinguishes Nabataean Arabic, and the Arabic of the settlements which came under the Nabataean sphere of influence, from the Arabic inscriptions of the Ḥarrah (basalt desert) nomads, which retained -at, and also from the Bedouin dialects, which later determined the standardized recitation of the Qur’ān, in which the orthographic h of the rasm was corrected back to t, and distinguished orthographically by pointing as tā’ marbūta.\(^{23}\)

Alif maqsūrah (word final *-ay)

Alif maqsūrah is the use in Classical Arabic of word-final dotless yāʾ (ṣ) to represent /ā/. For example orthographic ‘ay ‘upon’ is pronounced as /ˈaː/ This reflects the Proto-Arabic *-ay. Al-Jallad (2017: 154-155) has argued that a Graeco-Nabataean inscription supports an /æ:/ pronunciation for reflexes of *-ay in Nabataean Arabic. For example, in the inscription the Nabataean deity לודיא (Dushares */du-sʰaray/) is written דועסארה, the dative of דועסארה and not of **Δουσαρας (Littmann, Magie, and Stuart, 1907-21: 234).\(^{24}\)

Eighty years ago Bergstässer and Pretzl (1938:37) had already observed that recited /ā/

\(\) 23. The correction applied in Qur’ānic recitation even to Tawrāḥ (pronounced /tawrāt/) in which the -h was not a feminine ending, but original, being borrowed from Hebrew tōrāh ‘law’.

written as *alif maqṣūrah* does not rhyme with /ā/ written with *alif* in the Qurʿān, and they came to the same conclusion as Al-Jallad, based on the Graeco-Arabica evidence, namely that *alif maqṣūrah* reflects a distinct vowel, which they transcribed as ā. For example, compare Q91 with a consistent *alif* rhyme, and Q92 with a consistent *alif maqṣūrah* rhyme, and Q79, which has alternating but distinct passages of *alif* rhyme (Q79:27–32, 42–46) and *alif maqṣūrah* rhyme (Q79:15–26, 34–41), without any intermingling of the two rhymes.²⁵

Al-Jallad (2017a, 153) observed that Graeco-Arabica ‘generally agrees with Qurʿānic orthography’ in its treatment of *-ay*. Here again, the evidence shows that the pronunciation of Qurʿānic Arabic at the time of its initial recitation is consistent with Nabataean Arabic.

---

25. A small handful of words whose *rasm* ends in *y*-alif rhyme with *alif maqṣūrah*, e.g. *dunyā* (*rasm*: d-n-y-alif, cf. Q53:29; Q79:38; Q87:16) and *ahyā* (*rasm*: h-y-alif, cf. Q53:44), which suggests that /y/ caused the raising of a following /ā/. There is also *shayʿan* (*rasm*: sh-y-alif, cf. Q53:28), but this may be an imperfect match because there are also rhymes for *shayʿan* with -īyan forms, also spelled *y*-alif, in Q19:9, 42). Concerning *ahyā* see also the discussion of *-aya* in III-y verb forms of Al-Jallad (2017a, 156): Graeco-Arabica evidence suggests that a pronunciation /-ē/ was characteristic of of *-aya* in III-y verb forms in the Southern Levant.

26. Al-Jallad (2017: 168) argues that the the glottal stop was originally pronounced, at least in Safaitic, based on a Safaitic-Nabataean bilingual inscription in which the Nabataean name *ʾazʿalatmaʿ* is transcribed as ‘mt lʾz: the omission of the final glottal stop in Safaitic suggests that its inclusion in other contexts was contrastive.
In this respect also, the Qurʾānic rasm aligns with Nabataean Arabic.

Unassimilated articles

Our final piece of evidence concerns the assimilation of the definite article *al-* to a following coronal consonant, this being indicated in Arabic orthography by the *shaddah* diacritic. Here again, the lack of assimilation in the Qurʾānic rasm agrees with the evidence of Graeco-Arabica and Nabataean inscriptions, and also with Saffaitic and Ḥismaic inscriptions (Al-Jallad 2017a, 166-167; Macdonald 2000, 51). There is papyri evidence from Petra as late as the 6th Century CE of unassimilated *al-* (Al-Jallad 2017:169). For example ἐλδαργαθ, reflecting /el-dargāt/, is found in a manuscript dated to 579-580 CE. This is our fifth piece of evidence linking the Arabic of the Qurʾān to the Nabataean dialect.

There is a possible distinction to be made between the Arabic dialect in which the orthographic conventions of the rasm became fixed, and the dialect in which the Qurʾān was first recited. The two are not necessarily the same. It would hardly be surprising if the early spelling conventions of Arabic were based on Nabataean Arabic, since it is well established that Arabic orthography developed from the Nabataean Aramaic script. Indeed, in relation to each of the five features we have considered, the orthography of the rasm is consistent with what we known of Nabataean Arabic. In addition, the first three features, being reflected in Qurʾānic rhyme schemes, also point to Nabataean Arabic being the variety in which the Qurʾān was first recited.

For each of the linguistic features we have considered, the rasm of the Qurʾān aligns with Nabataean Arabic. This linguistic evidence confirms the circumstantial evidence of the ‘puzzle from above’ and the ‘puzzle from below’, that the dialect in which the Qurʾān was composed was Nabataean Arabic.


29. Arjava, Buchholz and Gagos, 2007: 30, 48. However, by the first Islamic century the –*al* assimilates in the Graeco-Arabica to a following coronal consonant (Al-Jallad 2017b, 428).
**A Pure Bedouin Tongue?**

There is a potential tension in idea of Arab identity and the origin of Qur'ānic Arabic proposed here. On the one hand, we have argued that the Nabataean form of Arabic, which was associated with trade, agrarian settlement and courtly culture, supplied the Arabic standard for the emerging Arab empire and its religion, including the 'arabīya (Q12:2) in which the Qurʾān was written and spoken. On the other hand, the Bedouin nomads or 'aʿrāb were held in high esteem as the prototype of Arab identity.

In a sense this tension is apparent in the Qurʾān. On the one hand the Qurʾān takes pride in its language being ‘clear Arabic’ (Q16:103; Q26:195), while on the other hand it expresses a derogatory attitude to those who it calls 'aʿrāb (Q9:90;97-98;101; Q48:11; Q48:16; Q49:14). These are the desert-dwelling Bedouin (Q33:2), in contrast to the settled people of the city (Q9:120), who constitute the Qurʾānic messenger’s primary audience. The sympathies of the messenger are with the city dwellers, and it seems consistent that the Arabic used in the Qurʾānic rasm was a settled, even urbanized standard variety typical of Nabataean towns and villages, and not a Bedouin dialect. It also is also understandable that later Muslim scholars – many of them not Arabs (Hoyland, 2001:247) – were romanced by the idea of a pure Bedouin Arab stereotype, reflected in the saying afsahu l-ʿarabi abarruhum ‘the best speakers of Arabic are those deepest in the desert’ (Rabin, 1951:18; c.f. also Hoyland, 2001:245-246). On this assumption, medieval Muslim scholars attempted to de-urbanise the origins of Qur’anic Arabic by sourcing it somewhere in the Arabian desert. This conflated two ideas: the doctrinal necessity that the Qurʾān’s language was pure, and the stereotype that the language (and culture) of the Bedouins was pure. These joined into a single proposition: that the Qurʾān was composed in a pure Bedouin tongue. This led Muslim scholars to devote considerable efforts to determining which Bedouin dialect (or dialects) could have provided the linguistic variety of the Qurʾān. It apparently did not occur to them that this variety could be found among the settlements of Southern Levant, which by the period of the Arab conquests had converted to Christianity.
Implications for Qur’ānic History

This account of the origin and emergence of Arabic has implications for the history of the Qur‘ān. When the Qur‘ān states that Allāh only ever sends a messenger to speak in the language of his own people (Q14:4), this implies – given that the consonantal skeleton is in Nabataean Arabic – that the Qur’ānic messenger’s ‘own people’ were native speakers of the Nabataean koiné. and the title ‘Mother of Towns’ of Q42:7 must have applied to an urban settlement in the Nabataean sphere of influence. It is hard to see how this could have been Mecca. On the other hand, if Muḥammad had been debating with Meccan tribespeople in their own mother tongue, this would have been a Bedouin dialect, and not a supra-regional koiné. In the light of the evidence linking Nabataean Arabic the Qur‘ān, it seems scarcely credible that by the late 600’s CE Meccan Bedouin tribespeople would have spoken the ‘clear Arabic language’ (Q16:103; Q26:195) of the Qur‘ān as their mother-tongue. Indeed, as we have seen, it was found by the medieval Muslim grammarians that the Meccan dialect diverged from the Qur’ānic standard.

Conclusion

The evidence put forward here suggests that the Arabic in which the Qur‘ān was recited and written was Nabataean. The Nabataean variety, it is suggested, would have been widely understood throughout the Arabic speaking region because of Nabataean trade networks. It was not considered to be a different language from local Bedouin dialects, but a ‘clear’ form of Arabic. A tradition of court poetry, probably based on the Nabataean variety, had already developed in the pre-Islamic period, no doubt influenced by the former Nabataean city of Bosra, which became the capital of Arabia Petrae.30 This explains why Muslim grammarians identified the poetic koiné as the same variety as Qur’anic Arabic. Later, in the process of standardising the recitation of the Qur‘ān, other dialect features were overlaid upon the rasm of the Qur‘ān and marked by diacritics, including case endings, and the consonantal

30. One important difference between Nabataean Arabic and the poetic variety was that the latter used the i‘rāb, using this as part of its rhyme schemes. This presumably was the result of a desire to bedouinise the language of the poetry.
assimilation of the definitive article, which were not features of the dialect in which the Qurʾān was originally written.\(^{31}\)

This model of the source of Qurʾānic Arabic accounts for the evidence of pre-Islamic epigraphy, as well the testimony of the Muslim philologists who standardized classical Arabic. It resolves the puzzle from below, as well as the puzzle from above. It also accords with the linguistic evidence, that Nabataean Arabic aligns with the linguistic variety attested in the *rasm* of the Qurʾān.

The solution to the puzzle from below is that epigraphic evidence of a precursor to classical Arabic appears so rarely in inscriptions because the Nabataeans who spoke the precursor variety preferred to write in Aramaic (and for a time, in Greek). The solution to the puzzle from above is it proved impossible for early Muslim philologists to locate the source dialect for Qurʾānic Arabic among the various Bedouin tribes where they sought it, because Nabatean Arabic was not a Bedouin variety. In any case, Petra was no more, Nabataean identity had dissipated, and Nabataean Arabic was no longer the distinctive variety of any particular tribe, but had become the linguistic inheritance of peasants in the Southern Levant, and, in the poetic koiné, the common property of the Arabs.\(^{32}\)

**BIBLIOGRAPHY**


---

\(^{31}\) Rabin (1955, 25) reports evidence that the Qurʾān was commonly being recited without *iʿrāb* as late as the 2nd century AH.

\(^{32}\) The account offered here cannot be reconciled with the widely-held view of the medieval Muslim philologists that the Qurʾān was composed in the dialect of the Quraysh (Rabin, 1951:22). However Muslim philologists themselves had already supplied the evidence to disprove this view.


———, (2009c) *Literacy and Identity in Pre-Islamic Arabia.* Surrey: Farnham: Ashgate.


