*Article*

# Connecting the Lines between Old (Epigraphic) Arabic and the Modern Vernaculars

Ahmad Al-Jallad ●

Faculty of Theology, University of Groningen, 9712 CP Groningen, The Netherlands; a.m.al-jallad@rug.nl

**Abstract:** This paper investigates three linguistic features—wawation, the 1CS genitive clitic pronoun, and the relative pronoun—that are shared between the ancient epigraphic forms of Arabic and modern dialects, to the exclusion of Classical Arabic. I suggest that these features represent the earliest linguistic layer of the modern dialects.

## 1. Introduction

It has been widely recognized that the diverse forms of spoken Arabic today do not descend in a linear manner from the literary Arabic of medieval prose and poetry—conventionally termed Classical Arabic—or the language of the Quranic Consonantal Text (QCT), Old Ḥigāzī (for the most recent appraisal, see Holes 2018a, pp. 1–28; Al-Jallad 2020b, chps. 4 and 5). Indeed, when viewed through the lens of the comparative method, many modern Arabic vernaculars exhibit features that are more archaic than their Classical Arabic counterparts. Na'ama Pat-El (2017) has skillfully identified a number of such features in her 2017 article "Neo-Arabic and Comparative Semitics". Clive Holes has also done pioneering work on pre-Islamic relics in the modern vernaculars of the Gulf, especially in the realm of the lexicon (Holes 2018b, pp. 112–32). Van Putten and Benkato (2017) isolated relics of an earlier stratum of Arabic in loans in Awjila Berber that is distinct from the present-day dialects of Libya. And I have suggested that the phonology of the emphatics of pre-Hilalian Maghrebian Arabic may be connected to the pre-Islamic dialects of the Levant (Al-Jallad 2015). The existence of these features implies that an unidentified stratum of Arabic that failed to achieve written form in the early Islamic period contributed to the formation of modern vernaculars.

This essay explores the possibility that such ancestors may be attested in the pre-Islamic epigraphic record. Before approaching this question, however, it is important to recognize two things. The modern vernaculars never existed in a vacuum; they have experienced considerable contact with the literary register, which has contributed significantly to their lexicons and to their grammatical structure. In addition to this, interdialectal contact has led to an amalgamation of grammatical features in living speech, ones that originate in different times and places. An obvious example of this is the verb *šāf* "he saw", which is nearly pan-Arabic today. *šāf*, although presently widespread in the Maghreb, was likely a late introduction through inter-dialectal contact (Aguadé 2018, p. 57). It is absent in Maltese, which became isolated from the Arabic sprachraum by the 13th century, and is not used in several pre-Hilalian dialects. These only know *ṛa*. The same applies to the Levant. There, *šāf* is the primary verb used to express "to see" in Lebanon, yet Cypriot Arabic, which originates on the Levantine coast and became isolated from the Arabic-speaking world by the 13th c. CE, does not use this etymon. Instead, it employs two verbs for "to see"—*ra* (Proto-Arabic *raʾaya; Classical Arabic raʾā; Borg 2004, p. 214) and *kišˁe* (Qəltu qašaˁ; Borg 2004, p. 388). The latter is fossilized as a presentative in Damascene Arabic, *šaˁ* (Souag 2016). While it is clear that Cypriot Arabic shares a common ancestor with the

dialects of the Levant, in the intervening centuries since its isolation, a new verb for "to see" spread as a result of contact with other dialects, in this case perhaps northern Arabian ones.

Likewise, Cypriot Arabic does not know the pseudo-verb *bVdd-* "to want" and instead makes use of a verb derived from the root *rwd*, *piri* (< *birīd* "he wants"; Borg 2004, p. 256; cf. Classical Arabic *yurīdu*). This it shares in common with the Qəltu dialects, while the modern dialects of the Levantine coast employ *bVdd*. The latter may also find its source in the North Arabian dialects, where "to want" can be expressed with the prepositional phrase, (*i*)*b-widd*-PN, or simply with *widd*-PN, literally meaning "in PN's wish" and "PN's wish", respectively. If we employ an archaeological metaphor, a dialect area, such as the Levant, can be regarded as an archaeological section. The layers would reflect different chronological strata of contact-based features and local innovations. While *šāf* and *bVdd* may reflect relatively late layers, this paper is interested in identifying the very earliest linguistic strata in the modern vernaculars.

Almost all who have discussed Arabic's past begin its historical period with the Quran and the nearly contemporary oral poems, passed on traditionally from *rāwī* to *rāwī* until achieving written form in the 8th–9th centuries at the earliest. The Quran itself is far from a linguistic unity. It minimally comprises a consonantal text, *rasm*, which reflects the local dialect of the Ḥigāz, while the reading traditions imposed upon it draw on various 7th and 8th c. varieties. The combination of these two linguistic types sometimes produces features that may never have been used in spoken language (Van Putten 2021, §3.4; Al-Jallad 2020b, pp. 57–72). Likewise, the oral poems can provide us with a glimpse of the performance language of that particular tradition, but we cannot know how much the odes changed over time as they were passed from generation to generation. Finally, their linguistic unity is little more than an assumption rather than a demonstrable fact. No one has yet, as far as I know, engaged in a truly comparative examination of the poetic tradition's language on its own terms.

Another corpus suitable for comparison exists: pre-Islamic epigraphy.[1] These texts, which are carved in nearly half a dozen scripts, offer both advantages and disadvantages. To begin with the latter, the inscriptions do not belong to a living tradition. While the researcher has the work of early Islamic philologists to rely upon when approaching the Qaṣīdah odes and the Quran, the meaning of the pre-Islamic inscriptions must be reconstructed. However, with a proper comparative approach, and with due attention to archaeological and historical contexts, one can be confident about the meaning and grammar of a large part of the corpus. Nevertheless, the consonantal Semitic scripts that encode these ancient Arabic vernaculars provide us with a very limited view of their phonologies and morphology.

These materials come with advantages as well. We can be sure that their language was not filtered through later, Classicizing traditions. They reflect a register of Arabic used at the time they were produced, and since many are simple graffiti, they likely reflect something close the vernacular of their writers. The pre-Islamic inscriptions, moreover, stretch much further into the past than the pre-Islamic odes, as far back as the middle of the first millennium BCE if not earlier, and cover a wider geographic area, spanning from the Syrian desert to the Yemeni frontier.

As such, how can this corpus aid in the understanding of the linguistic history of the Arabic vernaculars? The answer is not straightforward. In some cases, we may posit a direct developmental trajectory between a phenomenon attested in the ancient sources, but in others, similarities may point towards parallel developments in the history of the language. The following pages will identify three features that the modern dialects share with the ancient epigraphy to the exclusion of normative Classical Arabic. I would suggest that these are reflective of the earliest linguistic layer of present-day vernacular Arabic.

## 2. Wawation

Proto-Arabic inherited the Proto-Semitic case system with only a few changes, including the emergence of a new declension (Huehnergard 2017; Al-Jallad and van Putten 2017;

[Al-Jallad forthcoming](#)), but the case system began to disappear in several ancient dialects of Arabic at approximately the turn of our era, mainly concentrated in the Nabataean realm ([Corriente 1976](#); [Blau 2006](#)). The first stage of this process appears to have been the loss of final short vowels and then the loss of *nunation* (tanwīn), which resulted in a new set of final vowels in triptotic nouns. While a couple of inscriptions attest a functional declensional system in this state, the majority situation generalizes the nominative ending in all syntactic positions.[2] This feature—conventionally termed *wawation*—is encountered not only in the Nabataean inscriptions, but wherever one finds triptotic Arabic names in the Aramaic inscriptions of the first millennium BCE and the first half of the first millennium CE. Perhaps the earliest attestation of this feature in the Aramaic script is found in the 5th c. BCE votive inscription of Qaynu son of Guśam king of Qaydar at Tell Maskhūṭah, Egypt ([Rabinowitz 1956](#)). Wawation is attested continuously throughout the centuries in northern Arabic dialects, appearing on the anthroponyms and tribal names in the Namārah inscription and even in 6th c. CE Arabic inscriptions from Syria and North Arabia ([Al-Jallad forthcoming](#)).

> Tell Maskhūṭah (5th c. BCE)
>
> C *zy* **qynw** *br* **gšm** *mlk qdr qrb l-hnˀlt*
>
> "That which **Qaynu** son of Guśam has offered to han-ˀIlat (the goddess)"
>
> Namārah inscription (S. Syria) (328 CE)
>
> *w-mlk ˀl-ˀšryn w-***nzrw** *w-mlwk-hm w-ḥrb* **mdḫgw**
>
> "He ruled the two Syrias and **Nizāru** and their kings and waged war upon **Maḏḥigu**"
>
> Ḥarrān inscription (S. Syria) (568 CE)
>
> *ˀnˀšrḥyl br* **ṭlmw**
>
> "I am Šaraḥīl son of **Ẓālimu**"

The distribution of ancient *wawation* is as follows: with a few exceptions, it appears on triptotic anthroponyms and on Arabic proper nouns. It does not attach to names terminating with the feminine ending *-at*, nor does it attach to diptotic names belonging to patterns such as *fuʿal*, *ˀafʿal*, and *fVʿlān* or names defined by the article. It is reasonable to assume that this distribution applied to nouns as well, although it is impossible to prove as there are so few examples of Arabic prose written in the Classical Nabataean script. JSNab 17, an Arabic inscription carved in the Nabataean script from Madāˀin Ṣāliḥ (dated 267 CE; [Fiema et al. 2015](#)), marks all triptotic nouns with wawation, including definite forms: *ˀlḥgrw* = ˀal-Ḥiǧr, the ancient name of Madāˀin Ṣāliḥ, *ˀlqbrw* = ˀal-qabru 'the grave' ([Fiema et al. 2015](#)). While wawation does not apply to anthroponyms with the definite article—for example, the name *marˀalqays* (=imruˀulqays) is always written *mrˀlqys* and never *mrˀlqysw*—its application appears to have been extended in the realm of nominal morphology, at least in some varieties.

The *u* termination is also encountered in the modern Arabic vernaculars of southwest Arabia, concentrated in the Yemeni Tihāmah, extending as far north as the dialect of Balqarn ([Behnstedt 2016](#), p. 81; [Greenman 1979](#); [Alqahtani 2015](#)). Nouns terminating in a non-etymological *u* have a distribution virtually identical to anthroponyms terminating in *waw* in the ancient inscriptions: it is restricted to triptotic nouns and does not occur on nouns with the feminine ending *-at*. The striking congruence of both of these systems motivated [Blau](#) ([2006](#)) to compare them directly. While he stops short of suggesting a genealogical relationship between the dialects of Southwest Arabia and the ancient North Arabian dialects, the particular sequences of changes required to produce a nearly identical distribution at both ends of the ancient Arabic sprachraum does suggest that the feature may share a common ancestor.

The Southwest Arabian dialects, however, attest an important difference. There are some dialects where wawation is in complementary distribution with tanwīn. The former appears in pause and the latter in context. Nöldeke was the first to hypothesize

that the Nabataean *w* had developed from *-un*, but in these Tihāmī dialects we see the process in action. The asymmetric situation is rare, isolated to a few dialects of the ʿAsīr (Behnstedt 2016, p. 81). Rather, most dialects of the area have generalized one form. Those on the Tihāmī coast have generalized *u* while most in the ʿAsīr have only the nunated ending, either *un* or *in*. Thus, as Blau (2006) suggested, the following relative chronology appears secure (Figure 1):

| | | |
|---|---|---|
| Stage 1 | *al-baytu* | *baytun* |
| Stage 2 | *al-bayt* | *baytun* |
| Stage 3 | *al-bayt* | *baytu (pause) – baytun (context)* |
| Stage 4 | *al-bayt - baytu (generalized)* | *al-bayt - baytun/in (generalized)* |

**Figure 1.** Stages in the development of wawation.

Those dialects exhibiting the *baytu/baytVn* opposition appear to be more archaic than the Nabataean situation at first glance, but this may simply be an accident of attestation. Since most of the nouns attested in Nabataean occur in an Aramaic linguistic setting, it may be the case that their attested forms are pausal. While there is no direct evidence for the preservation of nunation in Nabataean inscriptions, a clue might be found in the Nahal Hever papyri, which are first c. CE legal documents from the Dead Sea area. The Arabic noun for "contract" is attested with an otiose final *nūn*, ʿ*qdn*. Although Yardeni (2014) suggested that this could possibly be a first person pronominal suffix, it would make little sense in this context. Rather, one could carefully hypothesize that it be interpreted as the ad-hoc writing of context form, with nunation. An even earlier example of functional nunation is attested in a widely known yet unpublished inscription from the Taymāʾ area. The text—carved in an oasis North Arabian alphabet—was authored by the king of Dūmat (mod. Dawmat al-Jandal) and can be dated to the middle of the 6th c. BCE based on its reference to the Babylonian king Nabonidus. All non-pausal, non-construct, and non-diptotic nouns terminate in a *nun*.[3]

### The Bsrn inscription

ʾ*n : bsr**n** : ʿbd : nbwnʾd : mlk : bbl : nẓrt : h-ġn**m** : b-mʾt**n** : frs**n** : w-mʾt**n** : rkb : ʾbl*

'I am Bsrn servant of Nabonidus king of Babylon; I have guarded the spoils with a cavalry unit and a unit of cameleers'

The phrase *mʾt frs* "cavalry unit" is widely attested in the Safaitic inscriptions, which are about half a century later (Macdonald 2014). The appearance of *nūn*s in this inscription suggest that the two words do not form a genitive construction but rather a noun and adverb, *bi-miʾatin farasan*. The final word of the inscription, ʾ*bl*, lacks a *nūn*, perhaps suggesting that it is a pausal form.

This distribution could indicate that both the ancient northern Arabic dialects and those of southwest Arabia share a common ancestor that had undergone the changes described above. Over the passage of time, each group altered the asymmetric *pausal* vs. *context* distribution by generalizing one form. The *u* termination was eventually favored in Nabataean and the Tihāmah while the nunated form was favored elsewhere. Some varieties of Nabataean further generalized wawated forms to the definite declension as well, producing the situation we find in JSNab 17.

If the genealogical connection between these two dialect groups is correct, then it may suggest that an ancient dialect of Arabic similar to what is attested in the Bsrn inscription moved south sometime in the first millennium CE and replaced the pre-Arabic languages of the ʿAsīr and Tihāmah.[4] We should further note that Nabataean Arabic and the dialects of southwest Arabia differ in the form of the definite article, *al* and *am* respectively. Thus, it is possible that the definite article of the ancestral dialect to both was *han-*, as attested in the Tell al-Maskhūṭah inscription. This morpheme split into ʾ*al-* in the north and ʾ*am* in the south (on the chronology of the Arabic article, see Al-Jallad 2021) (Figure 2).

|  |  |  |  |  |  |
|---|---|---|---|---|---|
| **Proto-Dialect** | *han-baytu* | *baytun* |  |  |  |
| **Northern** | *(h)al-bayt* | *baytVn – baytu* | **ʿAsīrī** | *am-bayt* | *baytVn – baytu* |
| **Nabataean** | *ʾal-bayt(u)* | *baytu* | **Tihāmī** | *am-bayt* | *baytu* |

**Figure 2.** Evolution of wawation in Nabataean, ʿAsīrī, and Tihāmī Arabic.

Wawation is today not only attested in southwest Arabia. It is also found in the Levant and Mesopotamia, where it is realized as *u* or *o*, depending on the dialect. It has a much more restricted distribution: the feature is found on high frequency kinship terms, such as Levantine Arabic ʿ*ammu* "paternal uncle", *ḫālu* "maternal uncle", *sīdu* "grandfather", *ǧaddu* "idem.", and on feminine nouns, *ḫāltu* "maternal aunt", etc. In northern Mesopotamia the *u/o*-termination applies only to masculine kinship terms, while feminine nouns terminate in -*a*; in Mardin, feminine vocative nouns terminate in -*e*. This distribution speaks against viewing the suffix as a third person masculine singular clitic; there would be no reason that it should be restricted to masculine nouns. Grigore (2007, p. 203) suggested that, at least for the dialect of Mardin, the termination could have a Kurdish source, but Procházka favors a Semitic origin as its distribution extends far beyond the areas in which Persian or Kurdish influence would seem possible (Procházka 2020, pp. 95–96). If I may go further, I would suggest, given the broader Arabic context, that the *u/o*-termination is a reflex of wawation as attested in Nabataean and in the southwestern Arabic dialects. The distribution in the Mesopotamian dialects matches the situation in Nabataean—it does not apply to nouns terminating in the feminine ending. The etymology of the feminine -*a* remains unclear. Perhaps Grigore (2007, p. 203) is correct to see a connection with Kurdish. While the masculine wawated form would have had an Arabic origin, speakers could have understood it as the same morpheme as the Kurdish vocative ending in a bilingual setting. The absence of any marking on feminine kinship terms perhaps motivated the borrowing of the Kurdish feminine ending to produce an etymologically mixed paradigm nearly identical with the Kurdish vocative paradigm.

The Levantine dialects appear to have extended the domain of wawation through analogy, appending the suffix to the female counterparts of male kinship terms; a similar extension of nunation occurred in Classical Arabic as Van Putten (2017) convincingly reconstructs the feminine ending as diptotic in Proto-Arabic.

The Levantine situation may, therefore, reflect a continuation of ancient Nabataean-type wawation, which survived marginally while the rest of the nominal system shifted—either through contact or through internal development—to favor the non-wawated paradigm. The early 6th century CE Arabic inscription from Jebel Usays[5] already demonstrates that the local Levantine dialects of Arabic had dispensed with wawation on personal names and nouns; thus, it is already possible at this point that the feature was restricted to kinship terms. It is not surprising that kinship terms would preserve older layers of morphology, and so this solution, if correct, would provide a unified analysis of wawation across Arabic.

To conclude, the linguistic stratum of wawation in the Levantine and northern Mesopotamian dialects, the ancient dialects of the southern Levant, and the modern Tihāmī and ʿAsīrī dialects would appear to share a non-Classical Arabic common ancestor with this distinct declensional profile.

## 3. 1CS Genitive Clitic Pronoun

The next feature I would like to consider is the 1CS genitive clitic pronoun. In all forms of Arabic, the shape of this pronoun is dependent upon the termination of the noun to which it attaches, as in other Semitic languages, but its distribution can vary from dialect to dialect. The pronoun has two allomorphs: -*ī* and -*ya*.

\*ī

    Classical Arabic: conditioned—following short vowels or consonants: *kitāb-ī*
    Ugaritic: conditioned—ø = /ī/ on nominative singular + fem. pl. nouns
    Phoenician: conditioned—ø = /ī/, nominative + accusative

\*ya

    Classical Arabic: conditioned—following long vowels and diphthongs: *ʿalay-ya*
    Gəʿəz: unconditioned—*hagaré-ya*
    Ugaritic: conditioned—*y* = /ya/, gen + acc singular, and other nouns; on prepositions
    Phoenician: conditioned—*y* = \*/ya/, genitive nouns

Some contemporary Arabic dialects, most notably those spoken in North Africa, employ the \**ya* allomorph following certain prepositions: Maghrebian *liya* "to, for me"; *biya* "in/by me", in contrast to normative Classical Arabic *lī* and *bī*, respectively. This distribution may in fact not be innovative. Various Quranic reading traditions produce such forms, but perhaps more importantly, the *rasm* itself demonstrates that this allomorph was in existence and had a much wider distribution.

**Quran**

69:19

<div dir="rtl">فَأَمَّاْ مَنْ أُوتِىَ كِتَبَهُ بِيَمِينِهِ فَيَقُولُ هَآؤُمُ ٱقْرَءُوا كِتَـٰيَهْ</div>

*faʾammā man ʾūtiya kitāba-hū bi-yamīni-hī fa-yaqūlu hāʾumu qraʾū **kitāb-iyah***

"and whosoever has received his record in his right hand will exclaim—Behold! Read aloud **my record**"

69:20

<div dir="rtl">إِنِّ ظَنَنتُ أَنِّ مُلَـٰق حِسَابِيَهْ</div>

*ʾinnī ḏanantu ʾannī mulāqin **ḥisāb-iyah***

"I had thought that I would surely face **my doom**"

69:28

<div dir="rtl">مَآ أَغْنَىٰ عَنِّ مَالِيَهْ</div>

*mā ʾaġnā ʿannī **māl-iyah***

"**My wealth** has not availed me"

In Sūrat al-Ḥāqqah, the termination *iyah*, where the final *h* should be understood as *hāʾu s-sakt*, i.e., a pausal *h* following a short vowel, is used on nouns that are syntactically nominative (*māliyah*) and accusative (*kitābiyah* and *ḥisābiyah*). The employment of the *ya* allomorph in these contexts is certainly motivated by rhyme, but there are other places in the Quran that demonstrate that its conditioning environment was slightly different from normative Classical Arabic. The vocative expression in Quran 12:84, ياسفى, is read by Ḥafṣ as *yā ʾasafā* and by al-Kisāʾī as *yā ʾasafē*, translated as "woe to me" (lit. O my woe). Q 5:31 attests a similar construction, يويلى, Ḥafṣ *yā waylatā*, al-Kisāʾī *yā waylatē*. The *alif maqṣūrah*, read by Ḥafṣ as *ā* and al-Kisāʾī as *ē*, reflects the outcome of an original triphthong, \**yā ʾasafa-ya > yā ʾasafē* (Old Ḥigāzī; al-Kisāʾī) and *yā ʾasafā* (Ḥafṣ) (Al-Jallad and van Putten 2017, pp. 113–14). Thus, these expressions preserve a situation where Arabic deployed the *ya* suffix following a short /a/, the accusative. Finally, in agreement with the modern North African varieties, the first person clitic following the preposition *li-* is sometimes realized as *ya*, depending on the reading tradition. Ḥafṣ reads لى as *liya*, for example, in Q 36:22.

The pre-Islamic Arabic inscriptions also attest a different distribution of the *-ya* allomorph. The Safaitic inscription BES15 799 attests a construction that is identical to the Quranic use of the *-ya* allomorph in the vocative.[6]

**BES15 799**

*wgd sfr bny f ṯql ʿl-bny w ql ḫbl**y***

"he found the inscription of Bonayy and was weighed down (by grief) on account of Bonayy and said: woe to me (*ḫabla-ya* lit. O my woe)"

The use of the *-ya* allomorph following the short high vowel /i/ is also attested in the pre-Islamic corpus. A Thamudic D inscription from the northern Ḥigāz attests this allomorph following the preposition *bi*.[7]

**UdhThamD 1 = JSTham 213**

*rbt śq* **by** *ʾ{l} kn ʾmt śkrn*

'There is much longing **in me** (*biya*) for Kn the maidservant of śkrn.'

Finally, the Dumaitic inscription WDum 3 = WTI 23 attests the *-ya* allomorph on a noun which is syntactically in the genitive case. Its presence implies that the genitive ending was still productive in this stage of the language.[8]

**WDum 3; WTI 23**

*h rḍw w nhy w ʿtrsm sʿd-n ʿl-wdd-**y***

'O Ruṣaw and Nuhay and ʿAttarsamē, help me in the matter of my wish (*widādiya*)'

The combination of these facts indicates that the Proto-Arabic distribution of the *ī* and *ya* allomorphs of the 1CS genitive pronoun was different from normative Classical Arabic. Rather, its appearance following the accusative in vocatives /a/, and short /i/, following prepositions like *li* and *bi*, and the genitive in Dumaitic, indicates a distribution similar to Ugaritic. Thus, we can reconstruct the Proto-Arabic situation as such:

**Nouns**
Nom: \**gamal-ī*
Gen: \**gamali-ya* (Attested: Dumaitic; relics: QCT)
Acc: \**gamala-ya* (Relics: vocative in QCT and Safaitic)
**i-vowel prepositions:**
\**li-ya*
\**bi-ya*
**Long vowels + diphthongs**
\**ʿalay-ya*
\**yadā-ya*

In this light, modern vernaculars that exhibit forms such as *biya* and *liya* continue the ancient situation, while Classical Arabic is innovative in its generalizing of the *-ī* ending to these propositions. As one reviewer of this paper pointed out to me, the quality of the vowel of the preposition in the Maghrebian varieties suggests that its immediate ancestor was long, *liya < \*līya*. Maghrebian Arabic generally loses etymologically short vowels, except in unstressed function words, where they are reanalyzed as long, e.g., the third masculine plural pronoun *hūma < hum*. Thus, an original \**liya* would have plausibly yielded *līya*; the same applies to the form *biya*.

The vocative form may also be attested in some modern dialects. In some Levantine dialects, the expression *yābāye* is used in situations of distress. It translates literally as "O my father." If the expression goes back to \**yā ʾabā-yah*, with *hāʾu s-sakt*, then it would parallel similar constructions in the Quran and Safaitic.

*Hāʾu s-sakt* must be reconstructed for the ancestor of the forms *liya* and *biya* as well. The presence of a final *a* in these cases is anomalous, as final-short vowels, including *a*, have generally been lost in the modern vernaculars (Figure 3).

| Proto-West Semitic | Classical Arabic | Egyptian |
|---|---|---|
| \**kʾatala* | *qatala* | *ʾatal* |
| \**taḥta* | *taḥta* | *taḥt* |
| \**sāmiʿīna* | *sāmiʿīna* | *sam ʿīn* |

**Figure 3.** Loss of final *a* in Modern Egyptian.

Thus, the survival of the vowel suggests the presence of a final *h*, protecting it from apocope. In other words, the antecedent of dialectal *biya* was not \**biya* but rather \**biyah*, as attested in Sūrat al-Ḥāqqah.

To conclude, both the distribution and form of the 1CS genitive clitic pronoun in the modern dialects speaks against a Classical Arabic origin, but should rather be connected with phenomena attested marginally in the QCT and in the ancient inscriptions.

## 4. Relative Pronoun

The relative pronoun ʾ*allaḏī* is restricted to southwest Arabia today (Behnstedt 2016, p. 74), but in former times it was much more widely distributed (Holes 2018a, p. 13). It is the primary form attested in Middle Arabic texts, even those that are quite close to the vernacular. It is attested in the Damascus Psalm Fragment as ελλεδι (8th–early 9th c.; Al-Jallad 2020b, p. 26). If this form was common in medieval vernaculars, it has today given way to the virtually pan-Arabic relative pronoun *ʾalli* (Stokes 2018). Yet *allaḏī* seems to have spread at the expense of an earlier relative pronoun *ḏV:*. To the Arabic Grammarians, *ḏV:* was characteristic of the dialects of southwest Arabia, where it can still be heard today, and the Najdi dialect of Ṭayyiʾ (Rabin 1951, chps. 3 and 14). In the modern dialects, *ḏ*-base relatives are common in Southwest Arabia (Behnstedt 2016, p. 74) and in the Maghreb (Aguadé 2018, p. 54). The genitive particles *ḏīl* and *ḏēl* (lit. "that which is for") in the Qəltu dialects and marginally in the Levant also suggest that at one point the relative pronoun of those dialects was a simple *ḏ*-base form (Procházka 2018, p. 280; Lentin 2018, p 195).

The relative *ḏV:* is attested across the pre-Islamic Arabic Sprachraum (Figure 4)—indeed, the form ʾ*allaḏī* has not yet appeared in the pre-Islamic epigraphic record, although its feminine counterpart ʾ*allatī* has been attested once in the Ḥigāz.

| Script | Location | Form |
|---|---|---|
| Hismaic | S. Jordan/N.W. Saudi Arabia | *ḏ* (ms); *ḏ ʾt* (fs) |
| Safaitic | Syro-Arabian *ḥarrah* | *ḏ* (ms); *ḏ ʾt*, *ḏt* (fs); *ḏw* (mpl) |
| Nabataean | Southern Levant–North Arabia | *dw* (generalized?) |
| Dadanitic | Al-ʿUlā/NW Arabia | *ʾlt* (fs) |
| Thamudic D | Medina | *ḏ* (ms) |
| Ancient South Arabian musnad | Qaryat al-Fāw | *ḏw* (mpl) |

**Figure 4.** Distribution of relative pronouns in the epigraphic record; data from Al-Jallad (2018).

In at least Safaitic and Hismaic it seems to inflect for case, gender, and number, with the plural form appearing as *ḏw* /ḏawū/ or /ḏawī/). Even as far south as Qaryat al-Faw, in the linguistically mixed inscription from the site, the Rbbl bin Hfʿm grave inscription, the plural form is attested as *ḏw* (Beeston 1979; Al-Jallad 2014). In Safaitic the relative may rarely agree in definiteness with its antecedent, producing *hḏ* /haḏḏī/.

The presence of the *ḏ*-base relative pronoun in all other branches of Semitic permits its secure reconstruction to Proto-Arabic, although there is not enough information to determine the details of its inflectional paradigm (Huehnergard 2017, pp. 16–17). This in turn indicates that the *ʾallaḏī* and later *ʾalli* forms are innovative, and spread at a later period, similar to *šāf* and *bVdd* discussed in the introduction.

Since *ḏV:* is an archaism it cannot be used to argue for a shared genealogical relationship between the dialects that preserve traces of it. It does, however, demonstrate that these dialects do not descend linearly from Classical Arabic, which had replaced this form with the *allaḏī*-type relative. Moreover, its presence throughout pre-Islamic Arabic prevents us from assuming that the *ḏ*- base relative pronoun in the modern vernaculars is a result of "South Arabian" influence, as has been previously suggested (Corriente 2007). The relative was not bound to a single geographic area in pre-Islamic times, but was in use from Yemen to Syria. Rather, it was the *allaḏī*-type relative that appears to have had a specific geographic distribution, restricted to the Ḥigāz. Today's dialect geography reflects a reversal of the pre-Islamic situation. The *allaḏī*-type relative, including *alli*, has spread at the expense of the older *ḏ*-type, which is today restricted to the periphery of the Arabic *sprachraum*.

## 5. Concluding Remarks

The features discussed here are but a small sample of possible Old Arabic relics strewn throughout modern Arabic vernaculars. They nevertheless motivate one to think in terms of a three-dimensional dialect continuum, extending not only geographically but also chronologically. Interdialectal contact, substrate contributions from the pre-Arabic languages of all regions to which Arabic spread, and the heavy superstrate influence of Classical Arabic prevent us from regarding any dialect as a monogenetic descendent of a pre-Islamic variety. Yet there can be no doubt that pre-Islamic phonological and morphological features absent in Classical Arabic contributed to the formation of the modern vernaculars.

**Data Availability Statement:** Not applicable.

## Appendix A



**Figure A1.** Safaitic Inscription BES15 799 (courtesy OCIANA).

BS 821: *l mġyr bn msk bn ʿmd bn mlk w wgd sfr bny f ṯql ʿl bny w ql ḫbly w ḏkr rgl f ʾdn {ʿ}l rgl*

"By Moġayyer son of Māsek son of ʿamīd son of Mālek and he found the writing of Bonayy and was weighed down by grief for Bonayy and said "O my woe" and he remembered Rāgel and was debased (by grief) for Rāgel"

Commentary:

This text was discovered in the Jordanian Ḥarrah at 32.43341; 37.270460, during the 2015 campaign of the Badia Epigraphic Survey project. The author produced three other Safaitic inscriptions KRS 38, 1885, and 1886, in the same general region.

*wgd sfr*: "he found the writing", a common inscriptional genre produced upon the finding of the inscription's of distant or deceased loved ones.

*ṯql*: "he was weighed down", cf. Classical Arabic *ṯaqula*. The verb is only attested in grieving contexts and so should be construed as a metaphor for worry and grief.

*ql ḫbly*: "he said: woe to me!" The meaning of this line was discussed in section three of this paper. A similar expression is attested in KRS 941: *w ql ḫbl-h trḥ* "sorrow afflicted him".

*w ḏkr rgl w ʾdn*: "he remembered Rgl (likely vocalized as Rāgel) and was debased. ʾ*dn*, the causative of *danna* "to make lowly" should be construed as a passive here with an unexpressed agent, namely, grief.

## Notes

[1]　For a summary and linguistic classification of these texts, see Al-Jallad (2018) and Macdonald (2004).

[2]　The inscriptions that continue to exhibit a living case system are the En Avdat inscription (see Macdonald's contribution to Fiema et al. 2015); the inflection of Nabataean theophoric names such as ʿ*bdʾlbʿly* /ʿabdu-ʾal-baʿli/; and the Jebel Ramm Hismaic inscription (Macdonald 2018a, 2018b; Al-Jallad 2020a).

[3]　The text was published on Twitter by its discoverer, mr. Aqla al-Rabiah: https://twitter.com/aqlaalrbeah/status/1293867413197520896 (accessed on 18 October 2021). A preliminary edition of the text can be found here: https://safaitic.blogspot.com/2021/06/king-of-ancient-dumat-addumatu.html (accessed on 18 October 2021).

[4]　The movement from north to south is assured by the chronology of the inscriptions. Wawation of this sort appears to be in place as early as the 5th c. BCE in the north, while at the same time Southwestern Arabia was dominated by the Ancient South Arabian languages. The pre-Arabic situation in the ʿAsīr is so far unclear, but the existence of a number of texts from the region that defy interpretation indicates considerable linguistic diversity before Arabic dominated the region; see https://safaitic.blogspot.com/2021/08/more-pre-arabic-texts-from-asir.html (accessed on 18 October 2021). It is unclear when the process of Arabicization began, but the first appearance of Arabic-like features in the inscriptional record dates to the turn of the era, the same time when groups called ˣ*rb* appear in the inscriptions.

[5]　This inscription begins with the author's name ʾ*nʾ rqym bn mʿrf ʾl-ʾwsy* "I am Ruqaym son of Muʿarrif the Aws-ite". Wawation would be expected on both Rqym and Mʿrf according to its normal distribution in the 6th c. Arabic inscriptions and in Nabataean. On this text, see Macdonald's contribution to Fiema et al. (2015).

[6]　This inscription had not been previously edited. See Appendix A for the edition. Note that *y* only has a consonantal value in Safaitic and cannot indicate word-final *ī*.

[7]　This is my interpretation of the text, based on parallels in other Thamudic D inscriptions. The *editio princeps* differs from my reading. See Macdonald (2018a, 2018b).

[8]　Note that *matres lectiones* are not used in the orthography of the Oasis North Arabian scripts to indicate final long vowels, as shown with the verb *sʿdn* /sāʿidū-nī/ "help me". The final *y* of the inscription must therefore be consonantal.

## References

Aguadé, Jordi. 2018. The Maghrebi dialects of Arabic. In *Arabic Historical Dialectology*. Edited by Clive Holes. Oxford: Oxford University Press, pp. 29–63.

Al-Jallad, Ahmad, and Marijn van Putten. 2017. The Case for Proto-Semitic and Proto-Arabic Case: A reply to Jonathan Owens. *Romano-Arabica* 17: 87–117.

Al-Jallad, Ahmad. 2014. On the genetic background of the Rbbl bn Hfʿm grave inscription at Qaryat al- Fāw. *Bulletin of the School of Oriental and Asian Studies* 77: 445–65. [CrossRef]

Al-Jallad, Ahmad. 2015. On the voiceless reflex of *ṣand *ṯin pre-Hilalian Maghrebian Arabic. *Zeitschrift für Arabische Linguistik* 62: 88–95.

Al-Jallad, Ahmad. 2018. What is Ancient North Arabian? In *Re-Engaging Comparative Semitic and Arabic Studies*. Edited by Daniel Birnstiel and Na'ama Pat-El. Wiesbaden: Harrassowitz, pp. 1–45.

Al-Jallad, Ahmad. 2020a. Notes on the language of the Hismaic Inscriptions and a re-reading of Line 1 of the Madaba Hismaic Inscription. *Journal of the Royal Asiatic Society* 30: 561–69. [CrossRef]

Al-Jallad, Ahmad. 2020b. *The Damascus Psalm Fragment: Middle Arabic and the Legacy of Old Ḥigāzī*. With a contribution by Ronny Vollandt. Chicago: Oriental Institute.

Al-Jallad, Ahmad. 2021. The History of the am-Definite Article—South Arabian or Arabic? *Zeitschrift für Arabische Linguistik* 73: 53–70. [CrossRef]

Al-Jallad, Ahmad. forthcoming a. "Night" in Proto-Semitic and the emergence of a new nominal declension in Arabic. *Studia Orientalia Electronica*.

Al-Jallad, Ahmad. forthcoming b. One wāw to rule them all: The origin and fate of wawation in Arabic. In *Scripts and Scripture*. Edited by Fred Donner and Rebecca Hasselbach. Chicago: Oriental Institute.

Alqahtani, Khairieh. 2015. A Sociolinguistic Study of the Tihami Qahtani Dialect in Asir, Southern Arabia. Ph.D. thesis, University of Essex, Essex, UK.

Beeston, Alfred F. L. 1979. Namarah and Fāw. *Bulletin of the School of Oriental and African Studies* 42 1: 1–6.

Behnstedt, Peter. 2016. *Dialect Atlas of North Yemen and Adjacent Areas*. Leiden: Brill.

Blau, Joshua. 2006. Problems of Noun Inflection in Arabic: Reflections on the Diptote Declension. In *Biblical Hebrew in Its Northwest Semitic Setting: Typological and Historical Perspectives*. Edited by Steven E. Fassberg and Avi M. Hurvitz. Jerusalem: Magnes Press, Winona Lake: Eisenbrauns, pp. 27–32.

Borg, Alexander. 2004. *A Comparative Glossary of Cypriot Maronite Arabic (Arabic-English): With an Introductory Essay*. Leiden: Brill.

Corriente, Frederico. 1976. From Old Arabic to Classical Arabic through the Pre-Islamic Koiné: Some Notes on the Native Grammarians' Sources, Attitudes and Goals. *Journal of Semitic Studies* 21: 62–98. [CrossRef]

Corriente, Frederico. 2007. On the prehistory of the Arabic language. *Aula Orientalis* 25: 141–53.

Fiema, Zbigniew T., Ahmad Al-Jallad, Michael C. A. Macdonald, and Laïla Nehmé. 2015. Provincia Arabia: Nabataea, the Emergence of Arabic as a Written Language, and Graeco-Arabica. In *Arabs and Empires before Islam*. Edited by Greg Fisher. Oxford: Oxford University Press, pp. 373–433.

Greenman, Joseph. 1979. A Sketch of the Arabic Dialect of the Central Yamani Tihamah. *Zeitschrift für Arabische Linguistik* 3: 289–304.

Grigore, George. 2007. *L'arabe parlé à Mardin: Monographie d'un parler arabe "périphérique"*. Bucharest: Editura Universităţii din Bucureşti.

Holes, Clive. 2018a. Introduction. In *Arabic Historical Dialectology*. Edited by Clive Holes. Oxford: Oxford University Press, pp. 1–28.

Holes, Clive. 2018b. The Arabic Dialects of the Gulf: Aspects of their historical and socitic development. In *Arabic Historical Dialectology*. Edited by Clive Holes. Oxford: Oxford University Press, pp. 112–48.

Huehnergard, John. 2017. Arabic in its Semitic Context. In *Arabic in Context*. Edited by Ahmad Al-Jallad. Leiden: Brill, pp. 3–34.

Lentin, Jérôme. 2018. The Levant. In *Arabic Historical Dialectology*. Edited by Clive Holes. Oxford: Oxford University Press, pp. 170–205.

Macdonald, Michael C. A. 2004. Ancient North Arabian. In *The Cambridge Encyclopedia of the World's Ancient Languages*. Edited by Roger D. Woodard. Cambridge: Cambride University Press, pp. 488–533.

Macdonald, Michael C. A. 2014. Romans go Home? Rome and other 'outsiders' as viewed from the Syro-Arabian Desert. In *Inside and Out. Interactions between Rome and the Peoples on the Arabian and Egyptian Frontiers in Late Antiquity*. Edited by Jitse H. F. Dijkstra and Greg Fisher. Leuven: Peeters, pp. 145–64.

Macdonald, Michael C. A. 2018a. Clues to How a Nabataean May Have Spoken from a Hismaic Inscription. In *Near Eastern and Arabian Essays Studies in Honour of John F. Healey*. Edited by George J. Brooke, Adrian H. W. Curtis, Muntasir al-Hamad and G. Rex Smith. Oxford: Oxford University Press, pp. 231–39.

Macdonald, Michael C. A. 2018b. The Ancient North Arabian and Ancient South Arabian Inscriptions. In *The Darb al-Bakrah*. Edited by Laïla Nehmé. Riyadh: Saudi Comission for Tourism and National Heritage, pp. 227–84.

Pat-El, Na'ama. 2017. Digging up Archaic Features in "Neo-Arabic" and Comparative Semitic in the Quest for Proto Arabic. In *Arabic in Context*. Edited by Ahmad Al-Jallad. Leiden: Brill, pp. 441–75.

Procházka, Stephan. 2018. The Northern Fertile Crescent. In *Arabic Historical Dialectology*. Edited by Clive Holes. Oxford: Oxford University Press, pp. 257–92.

Procházka, Stephan. 2020. Arabic in Iraq, Syria, and southern Turkey. In *Arabic and Contact-Induced Change*. Edited by Christopher Lucas and Stefano Manfredi. Berlin: Language Science Press, pp. 83–114.

Rabin, Chaim. 1951. *Ancient West-Arabian*. London: Taylor's Foreign Press.

Rabinowitz, Isaac. 1956. Aramaic Inscriptions of the Fifth Century B.C.E. from a North-Arab Shrine in Egypt. *Journal of Near Eastern Studies* 15: 1–9. [CrossRef]

Souag, Lameen. 2016. Book Review: Werner Diem. *Negation in Arabic: A Study in Linguistic History*. Wiesbaden: Harrassowitz, 2014. 119 pp. ISBN 978-3-447-10117-2; David Wilmsen. *ArabicIndefinites, Interrogatives, and Negators: A Linguistic History of Western Dialects*. Oxford Studies in Diachronic and Historical Linguistics. Oxford: Oxford University Press, 2014, 245 pp. ISBN 978-0-19-871812-3. *Linguistics* 54: 223–29.

Stokes, Phillip. 2018. The Plural Demonstrative Based on *'Vl in Arabic and the Origin of Dialectal illī. In *Re-Engaging Comparative Semitic and Arabic Studies*. Edited by Daniel Birnstiel and Na'ama Pat-El. Wiesbaden: Harrassowtz, pp. 131–54.

Van Putten, Marijn. 2017. The Feminine Ending -*at* as a Diptote in the Qurʾānic Consonantal Text and Its Implications for Proto-Arabic and Proto-Semitic. *Arabica* 64: 695–705. [CrossRef]

Van Putten, Marijn. 2021. *Quranic Arabic: From Its Hijazi Origins to Its Classical Reading Traditions*. Leiden: Brill.

Van Putten, Marijn, and Adam Benkato. 2017. The Arabic Strata in Awjila Berber. In *Arabic in Context*. Edited by Ahmad Al-Jallad. Leiden: Brill, pp. 476–502.

Yardeni, Ada. 2014. A List of the Arabic Words Appearing in Nabataean and Aramaic Legal Documents from the Judaean Desert. In *Scripta Classica Israelica: Rome, Judaea, and Its Neighbours: Special Issue in Honor of Hannah M. Cotton*. Edited by Werner Eck. Jerusalem: The Israeli Society for the Promotion of Classics, pp. 301–24.