Classical and Modern Standard Arabic

Putten, M. van; Lucas, C.; Manfredi, S.

Chapter 3

Classical and Modern Standard Arabic

Marijn van Putten
University of Leiden

The highly archaic Classical Arabic language and its modern iteration Modern Standard Arabic must to a large extent be seen as highly artificial archaizing registers that are the High variety of a diglossic situation. The contact phenomena found in Classical Arabic and Modern Standard Arabic are therefore often the result of imposition. Cases of borrowing are significantly rarer, and mainly found in the lexical sphere of the language.

1 Current state and historical development

Classical Arabic (CA) is the highly archaic variety of Arabic that, after its codification by the Arab Grammarians around the beginning of the ninth century, becomes the most dominant written register of Arabic. While forms of Middle Arabic, a style somewhat intermediate between CA and spoken dialects, gain some traction in the Middle Ages, CA remains the most important written register for official, religious and scientific purposes.

From the moment of CA's rise to dominance as a written language, the whole of the Arabic-speaking world can be thought of as having transitioned into a state of diglossia (Ferguson 1959; 1996), where CA takes up the High register and the spoken dialects the Low register.[1] Representation in writing of these spoken dialects is (almost) completely absent in the written record for much of the Middle Ages. Eventually, CA came to be largely replaced for administrative purposes by Ottoman Turkish, and at the beginning of the nineteenth century, it was functionally limited to religious domains (Glaß 2011: 836). During the nineteenth-century

[1]Diglossic situations are often seen as consisting of a high register (often called H) and a low register (L). These two are seen to be in complementary distribution, where each register is used in designated environments, where the H register takes up such domains like formal speeches and writing, while the L register is used in personal conversation, oral literature etc.

Marijn van Putten. 2020. Classical and Modern Standard Arabic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 57–82. Berlin: Language Science Press. DOI:10.5281/zenodo.3744505

Arabic literary revival known as the *Nahḍa*, CA goes through a rather amorphous and decentralized phase of modernization, introducing many neologisms for modern technologies and concepts, and many new syntagms became part of modern writing, often calqued upon European languages. After this period, it is customary in scholarly circles to speak of CA having transitioned into Modern Standard Arabic (MSA), despite the insistence of its authors that CA and MSA are one and the same language: *al-ʕarabiyya l-fuṣḥā* 'the most eloquent Arabic language' (Ryding 2011: 845).

2 Contact languages

Considering the significant time-depth of CA and MSA, contact languages have of course changed over time. Important sources of linguistic contact of the pre-Islamic varieties of Arabic that come to form the vocabulary for CA are Aramaic, Greek and Ethio-Semitic. While there are already some Persian loanwords in the very first sources of CA, this influence continues well into the Classical period, and ends up having a marked effect on CA and MSA alike.

2.1 Aramaic

Aramaic becomes the dominant lingua franca in much of the Achaemenid empire, and both written and spoken varieties of Aramaic continue to play an essential role all throughout Arabia, Syria and Mesopotamia right up until the dawn of Islam. As such, a not insignificant amount of vocabulary has been borrowed from Aramaic into Arabic, which shows up in CA. Moreover, Aramaic was an important language of Christianity and Judaism, and a noticeable amount of religious vocabulary from Aramaic has entered CA (§3.4.2). There may even be some structural influence on the phonology of pre-Classical Arabic that has made it into CA (§3.1).

2.2 Greek

Greek was the language of state of the Byzantine Empire, which, when not directly ruling over Arabic-speaking populations, was at least in close contact with them. This can be seen in the significant amount of Greek vocabulary that can be detected in CA. Aramaic, however, has often borrowed the same terms that we find in CA, and it is usually difficult, if not impossible, to decide whether a Greek word entered Arabic directly from Greek or through the intermediary of Aramaic (§3.4.3).

2.3 Persian

After the rise of Islam, Greek and Aramaic quickly lose the central role they once played in the region, and they do not continue to influence CA significantly in the Islamic period. Persian, however, of which a number of words can already be detected in the Qur̲an, continues to have a pronounced influence on Arabic, and many more Persian words enter CA throughout its history (§3.4.5).

2.4 Ethio-Semitic and Old South Arabian

It is widely recognized that some degree of influence from Ethio-Semitic can be identified within CA (§3.2.3; 3.4.1). Many of the Ethio-Semitic words that have entered into Qur̲anic Arabic presumably arrived there through South Arabian contact after the invasion of Yemen by Christian Ethiopia in the sixth century. Also previous South Arabian contact must probably be assumed, and the divine epithet *ar-Raḥmān* is usually thought to be a borrowing from South Arabian, where it in turn is a borrowing from Aramaic (Jeffery 2007 [1938]: 140–141).

While Ethio-Semitic contact has been fairly well-researched, research into contact with Ancient South Arabia is still in its infancy. The exact classification of the Old South Arabian languages and their relation to Modern South Arabian and Ethio-Semitic is still very much under debate. A simple understanding of this highly multilingual region seems impossible. Due to the extensive contact within South Arabia and the South Arabian languages, it is not always easy to pin down the exact vector of contact between CA and these languages of South Arabia and Ethiopia (§3.4.4).

2.5 Arabic dialects

The spoken Arabic dialects, of course, have had and continue to have a noticeable influence on CA and MSA (§2.5; 3.2.1; 3.2.2; 3.3; 3.5). It seems that from the very moment CA became canonized as an official language, it was already a highly artificial register that nobody spoke in the form in which it was canonized. Especially the Ḥiǧāzi conquerors had a noticeable effect on the language – no doubt through mediation of the Qur̲anic text. Noticeable irregularities in the treatment of the glottal stop, for example, have entered the language, and have influenced the treatment of certain morphological features (§3.2).

2.6 Ottoman Turkish

In the Ottoman period, Ottoman Turkish becomes the official language in use in the Middle East, and replaces many of the sociolinguistic functions that CA had

previously had. The imposition of this official language had a significant effect on the Arabic vernaculars throughout the Middle East (even outside the borders of the Ottoman Empire), but also had a noticeable impact on the vocabulary of CA, especially in the eighteenth and nineteenth centuries, which feeds into MSA (§3.4.6).

3 Contact-induced changes in Classical and Modern Standard Arabic

3.1 Phonology

Due to the highly conservative nature of CA, finding any obvious traces of contact in phonological change is very difficult. From the period in which Sibawayh describes the phonology of the *ʕarabiyya* until today, only minor changes have taken place in the phonology of CA. The most obvious example of this is the loss of the lateral realization of the *ḍād*, which in Sibawayh's description is still a lateral, while today it is generally pronounced as [dˤ]. Blau (1969: 162–163) convincingly attributes this development to influence from the modern dialects. In most modern Arabic dialects, the reflexes of *ḍ* [ɮˤ] and *ð̣* [ðˤ] merged to *ð̣* [ðˤ].[2] In sedentary dialects that lose the interdentals, this merged sound subsequently shifts to *ḍ* [dˤ]. As such, original *ð̣ and ḍ* are either both pronounced as an emphatic interdental fricative or both as an emphatic dental stop. As virtually all modern dialects, however, have lost the lateral realization of *ḍ*, the sedentary stop realization was repurposed for the realization of *ḍ*, to introduce the phonemic distinction between *ð̣* and *ḍ* in MSA.

As this is a case where the speakers influencing the phonology of the RL are SL-dominant, this change in pronunciation of the *ḍ* from a lateral to a stop realization can be seen as a form of imposition on the phonology of MSA. It should be noted, however, that the type of imposition we are dealing with in this case is of quite a different character than what is traditionally understood as imposition within the framework of Van Coetsem (1988; 2000). In this case, we see a conscious effort to introduce a phonemic distinction lost in the SL between original *ḍ* and *ð̣* by using two different dialectal outcomes of the merger of these two phonemes.

Other cases of phonetic imposition on MSA from the modern dialects may especially be found in the realization of the *ǧīm*. While Sibawayh's description of the *ǧīm* was probably a palatal stop [ɟ], today the realization that seems to carry

[2]Not all dialects, however, see Behnstedt (2016: 16ff.).

the most prestige and is generally adhered to in Quranic recitation is [ʤ]. However, here too we often find imposition of the local pronunciation of this sound in MSA. In spoken MSA of Egyptians the *ǧīm* is regularly pronounced as [g], the realization of the *ǧīm* in Egyptian Arabic. Likewise, Levantine Arabic speakers whose reflex of the *ǧīm* is [ʒ] will often use that realization when speaking MSA.

If we shift our focus to developments that began in the pre-Classical period and continue in CA, we find that there are several phonetic developments that bear some similarity to developments of Aramaic. It has therefore, not unreasonably, been suggested that such developments are the result of contact with Aramaic.

The first of these similar phonetic developments shared between CA and Aramaic is the shift of the semivowels *w* and *y* to *ʔ* between a preceding *ā* and a following short vowel *i* or *u*. This can be seen, for example, in the similar outcomes of the active participles of hollow roots. This similarity was already remarked upon and described by Brockelmann (1908: 138–139), e.g.:

(1) a. CA *qāwimun > *qāʔimun* ‘standing’
b. Aram. *qāwim > *qāʔem* ‘standing’

However, it is clear that, at least in Nabataean Arabic, this development had not yet taken place (Diem 1980: 91–93). This is a dialect that was certainly in contact with Aramaic, as most of the writing of the Nabataeans was in a form of Aramaic. As such, we may plausibly suggest that this development took place *after* the establishment of linguistic contact between Aramaic and Arabic. It is quite difficult to decide whether this development, if we are correct to interpret it as the result of contact-induced change, is the result of imposition, borrowing or convergence. We do not have a clear enough picture of the sociolinguistic relations between Aramaic and pre-Classical Arabic to identify the type of contact situation that would have caused it. One is tempted to see it as the result of imposition simply because of the fact that phonological borrowing seems to be uncommon (Lucas 2015: 526).[3]

As proposed by Al-Jallad (this volume), another possible case of contact induced phonological change between Aramaic and pre-Classical Arabic is the shift of pausal *-at* to *ah*, found only in nouns and not in verbs. Huehnergard & Rubin (2011: 267–268) already suggested that this development, which cannot be due to a development in a shared ancestor, may have been the result of areal diffusion.

[3]We cannot discount the possibility of parallel development, however. Akkadian seems to have undergone an almost identical development (Huehnergard 1997: 196), where it is not likely to have been the result of contact.

Whether we can really interpret the development of Aramaic as similar to that of CA, however, depends somewhat on the interpretation of the Aramaic evidence. While we can indeed see a development of the original Aramaic feminine ending *-at that is written with ⟨-h⟩ in consonantal writing, which might suggest it has shifted to /ah/, one also finds that all other cases of word-final nominal *t* have been lost, while not leaving a consonantal *-h*, e.g.:

(2) a. *ṣalōt > *ṣlō* ⟨ṣlw⟩ 'prayer'
b. *zakūt > *zkū* ⟨zkw⟩ 'merit, victory'
c. *ešāt > *ʔešā* ⟨ʔšʔ⟩ 'fire'
d. *bayt > *bay* ⟨by⟩ 'house'

For this parallel loss of final *t* in all other environments, Beyer (1984: 96, fn. 4) prefers to interpret the ⟨-h⟩ as a *mater lectionis* for final /ā/ or /a/. In this interpretation, the development of Aramaic compared to Arabic is quite different, since in Arabic the ⟨-h⟩ is clearly consonantal, and the loss of final *t* does not happen after long vowels in Arabic:

(3) Aramaic
a. *kalbat > *kalbā* ⟨klb⟩ 'bitch' (*-at# > -a/ā*)
b. *ʔešāt > *ʔešā* ⟨ʔšʔ⟩ 'fire' (*-āt# > -ā*)

(4) Arabic
a. *kalbat > *kalbah*
b. *kalbāt > *kalbāt* remains unchanged

However, if one takes the ⟨-h⟩ of the feminine to originally represent *-at > *-ah*, and the loss of *t* in other word-final positions to be a different development, one could reasonably attribute the development in Arabic to the result of contact with Aramaic, as it is clear that in many varieties of pre-Islamic Arabic, the *-at > *-ah* shift had not yet taken place.[4]

3.2 Morphology

3.2.1 Imposition of the taCCiʔah stem II verbal noun for glottal-stop-final verbs

A well-known feature of Ḥiǧāzī Arabic in the early Islamic period, and a feature that is found in many of the modern dialects, is the (almost) complete loss of the

[4]For a discussion on the development of the *-at > *-ah* shift in pre-Islamic Arabic see Al-Jallad (2017: 157–158).

glottal stop (Rabin 1951: 130–131; van Putten 2018). This loss has usually caused glottal-stop-final roots to be reanalyzed as final-weak verbs, e.g. Cairene *ʔara*, *ʔarēt* 'he read, I read' (< *qaraʔa, *qaraʔtu).

A typical feature of final-weak verbal noun formations in CA is their formation of the verbal noun of stem II verbs. Sound verbs form verbal nouns using the pattern taCCīC, e.g. *taslīm* 'greeting' from *sallama* 'to greet'. Final-weak verbs, however, regularly use the pattern taCCiyah instead (Fischer 2002: 44), for example, *tasmiyah* 'naming' from *sammā* 'to name'.[5]

In CA, the *ʔ* generally functions as a regular consonant. Thus a verb like *qaraʔa/yaqraʔu* 'to read, recite' does not differ significantly in its behavior from any other triconsonantal verb such as *fataḥa/yaftaḥu* 'to open'.

However, verbs with *ʔ* as final root consonants unexpectedly frequently side with the final-weak verbs when it comes to the verbal noun of stem II verbs (Fischer 2002: 128). For example, *hannaʔa/yuhanniʔu* 'to congratulate' does not have the expected verbal noun **tahnīʔ, but instead *tahniʔah* 'congratulation'. Other examples are:

(5) a. *nabbaʔa* v.n. *tanbiʔah* (besides *tanbīʔ*) 'to inform'
b. *barraʔa* v.n. *tabriʔah* 'to acquit'
c. *hayyaʔa* v.n. *tahyiʔah* (besides *tahyīʔ*) 'to make ready'
d. *naššaʔa*, v.n. *tanšiʔah* (besides *tanšīʔ*) 'to raise (a child)'

Some other verbs with the same pattern do have the expected CA form such as *baṭṭaʔa* v.n. *tabṭīʔ* 'to delay'.

This behaviour can plausibly be attributed to the fact that in many (if not most) spoken varieties of Arabic, from early on the final-glottal-stop verbs had already merged completely with the final-weak verbs, and as such a verb like *hannaʔa* had come to be pronounced as *hannā*, and was thus reanalyzed as a final-weak verb. Like original final-weak verbs, their regular verbal noun formation would be *tahniyah*. When verbs of this type were employed in CA, the weak root consonant *y* was replaced with the etymological glottal stop *ʔ*, rather than completely converting the verbal noun to the regular pattern. This is a clear example of the imposition of a morphological pattern onto CA grammar by speakers of Arabic dialects.

[5]This is an ancient idiosyncrasy of final-weak verbs. While the taCCīC formation is not a regular formation in other Semitic languages, when it does occur, the final-weak verbs have a feminine ending, e.g. Hebrew *tarmi-ṯ* 'betrayal', *toḏå* 'praise' (< *tawdiy-ah), see Brockelmann (1908: 385–387).

3.2.2 Imposition of the ʔaCCiyāʔ broken plural pattern

A similar case of imposition, where the morphological categories of glottal-stop-final roots behave in the grammar as if they are final-weak, may be found in the broken-plural formation of CaCīʔ nouns and adjectives. The broken-plural formation most generally used for final-weak adjectives with the pattern CaCiyy (< *CaCīy) is ʔaCCiyāʔ. For example, *ɣaniyy* pl. *ʔaɣniyāʔ* 'rich', *waliyy* pl. *ʔawliyāʔ* 'close associate', *daʃiyy* pl. *ʔadʃiyāʔ* 'bastard', *sawiyy* pl. *ʔaswiyāʔ* 'correct', *ḫaliyy* pl. *ʔaḫliyāʔ* 'free'.

For sound nouns of this type, it is much more typical to use the plural formations CiCāC (*kabīr* pl. *kibār* 'big') or CuCaCāʔ (*faqīr* pl. *fuqarāʔ* 'poor'), although there are a couple of sound nouns that do use this plural, such as *qarīb* pl. *ʔaqribāʔ* 'relative' and *ṣadīq* pl. *ʔaṣdiqāʔ* 'friend' (Ratcliffe 1998: 106–107).[6]

CaCīC formations where the last root consonant is *ʔ*, however, behave in rather unexpected ways in CA, usually following the pattern of final-weak nouns, often even replacing the final *ʔ* with *y*, for example: *barīʔ* pl. *ʔabriyāʔ* 'free', *radīʔ* pl. *ʔardiyāʔ* 'bad'. These nouns have plurals that are proper not to the Classical form they have, but rather to the colloquial form without *ʔ*, i.e. *bariyy*, *radiyy*. Once again this can be seen as a clear case of imposition of the colloquial Arabic forms onto the classical language.[7]

3.2.3 Borrowing of the broken plural pattern CaCāCiCah

CA, like the modern Arabic dialects, is well-known for its broken-plural patterns. This is a feature it shares especially with Old South Arabian (Stein 2011: 1050–1051), Modern South Arabian languages (Simeone-Senelle 2011: 1085) and Ethio-Semitic (Weninger 2011a: 1132). The use of broken plurals has caused somewhat

[6]The pattern (with metathesis) is also regular for geminated CaCīC adjectives, e.g. *šadīd* pl. *ʔašiddāʔ* 'severe'.

[7]These two cases of imposition of glottal stop-less morphology onto CA are two of the more clear and systematic cases, but close observation of CA morphology reveals many more of these somewhat more isolated cases, e.g. *ḫaṭīʔah* 'sin' with a plural *ḫaṭāyā*, for which the expected singular would rather be *ḫaṭiyyah*; *bariyyah* pl. *barāyā* 'creature' which is a derivation from *baraʔa* 'to create'; *ðurriyyah*, *ðirriyyah* pl. *ðarāriyy* 'progeny, offspring', derived from *ðaraʔa* 'to sow, seed'. Another example of irregular treatment of *ʔ* that is presumably the result of impositition is found in verbal nouns of stem VI verbs, and *mafāʃil* plurals of hollow roots, which modern textbooks say should not have a *ʔ* despite having the environment that is expected to undergo the shift *āwu/i, āyi* > *āʔu/i, āʔi* as discussed in §3.1. The lexicographical tradition and Quranic reading traditions often record disagreements on the application of the *hamzah* in such cases. For example, we find both *tanāwuš* and *tanāʔuš* 'reaching one another', and *maʃāyiš* and *maʃāʔiš* 'ways of living'.

of a controversy in the subgrouping of the Semitic language family. Scholars who consider broken plurals a shared retention do not view their presence as important for grouping Arabic, Old South Arabian, the Modern South Arabian languages and Ethio-Semitic together (Huehnergard 2005: 159–160); while those who consider their presence an innovation in a subset of Semitic languages see this as a strong indication that these languages should be grouped together into a South Semitic branch (e.g. Ratcliffe 1998).

While most scholars today seem to agree that the broken-plural system is a shared retention (Weninger 2011b: 1116), it seems clear that the retention of a highly productive broken-plural system is to be considered an areal feature that clusters around South Arabia and the Horn of Africa. CA partakes in this areal feature.

A possible case of influence from Old South Arabian (and/or Ethio-Semitic) into Arabic is the introduction of the CaCāCiCah plural formation. In the South Arabian languages,[8] the equivalent plural formation CaCāCiCt is extremely productive, and numerous words with four consonants form their plural in this way. For example in Sabaic, mCCCt is the regular plural formation to mCCC nouns of location, e.g. *mḥfd* pl. *mḥfdt* 'tower' (Beeston 1962: 34). It is likewise common in Gəʕəz, e.g. *tänbäl* pl. *tänabəlt* 'ambassador' (Dillmann 2005 [1907]: 309), and occurs occasionally in Modern South Arabian, e.g. Mehri *məlēk* pl. *məlaykət* 'angel' (Rubin 2010: 68).

While this pattern exists in CA, it is much rarer than the other broken plural formations of four consonantal forms, i.e. CaCāCiC and CaCāCīC. In the Quran, *malak* pl. *malāʔikah* 'angel' is the only plural with this pattern. This noun is widely recognized as being a loanword from Gəʕəz *malʔak*, *malāʔəkt* (Jeffery 2007 [1938]: 269), in part on the basis that it shares this plural formation: the word seems to have been borrowed together with its plural formation. Considering the rarity of this pattern in Arabic and how common it is in South Arabian, it seems possible that the pattern was introduced into Arabic through South Arabian contact. However, the absence of other clearly identifiable South Arabian loanwords with this plural pattern makes it rather difficult to make a strong case for this identification.

Another possible word of South Arabian origin with this plural pattern is *tubbaʕ* pl. *tabābiʕah* 'a Yemenite king', but evidence that this word is indeed of Old South Arabian origin is missing. The word does not occur as a separate word in Old South Arabian, and instead is only the first part of several Old South Arabian theophoric names such as *tbʕkrb*, *tbʔʕl*. Such names should probably be

[8]South Arabian is used here as a purely geographical descriptive term, not one of classification.

understood as being related to the root √*tbʕ* which, like in Arabic, may have had the meaning 'following', so such names likely mean 'follower of the deity KRB' and 'follower of the deity ʔL'. Such names being associated with Yemenite kings may have led to the Arabic meaning of *tubbaʕ* as 'Yemenite king', but in Old South Arabian itself it does not seem to have carried a meaning of this kind.

All in all, the evidence for this being a pattern that is the result of South Arabian influence is rather slim, although the rarity of the pattern in CA does make it look unusual. If the interpretation of this plural pattern as being a borrowing from South Arabian is correct, it seems that some South Arabian nouns were borrowed along with their respective plural. This would be a case of morphological borrowing rather than the more common type of morphological influence through imposition.[9]

Note that this plural pattern has become the productive plural pattern for quadriconsonantal loanwords regardless of them being of South Arabian origin or elsewhere, e.g. *biṭrīq* pl. *baṭāriqah* 'patrician' (< Latin *patricius*), *ʔusquf* pl. *ʔasāqifah* 'bishop' (< Greek *epískopos*), *ʔustāð* pl. *ʔasātiðah* 'master' (< Middle Persian *ōstād*), *tilmīð* pl. *talāmiðah* 'student' (< Aramaic *talmīḏ*).

3.3 Syntax

Due to CA being the High register in a diglossic situation for centuries, we should presumably consider the majority of the written material produced in this language to be written exclusively by non-native speakers. Moreover, a large proportion of its writers all throughout its written history must have been speakers not only of Arabic vernaculars but also of entirely different languages such as Persian and Turkish. It seems highly unlikely that such a multilingual background of authors of CA would have been completely without effect on the syntax of the language; however, as it is difficult to decide from what moment onward we can speak of true diglossia, and what the syntax was like before that period, it has not yet been possible to trace such influences in detail.

There is, however, promising research being done on influence on MSA syntax from the speakers of modern Arabic dialects. Wilmsen (2010) convincingly describes one such point of influence in a paper on the treatment of object pronouns in Egyptian and Levantine newspapers.

Wilmsen (2010: 104) shows that, in the case of ditransitive verbs, Egyptian and Levantine have a different natural word order. In Egyptian Arabic, the direct

[9]This can be seen as a type of "Parallel System Borrowing" similar to that which we find in Berber languages. Berber languages, like Arabic, have apophonic plurals; but Arabic nouns are simply borrowed along with their own Arabic broken plurals (Kossmann 2010).

object must precede the indirect object as in (6), while in Levantine Arabic the indirect object preceding the direct object is preferred, as shown in (7):

(6) Egyptian
rabbi-na yḫalli-**hū-l-ak**
Lord-1PL keep.IMPF.3SG.M-3SG.M-DAT-2SG.M
'Our lord keep him for you.'

(7) Levantine
aḷḷa yḫallī-l-ak iyyā
God keep.IMPF.3SG.M-DAT-2SG.M ACC.3SG.M
'God keep him for you.'

Wilmsen argues that the following two variant sentences in a Reuters news story written in MSA, the original in (8), likely written by an Egyptian, and the slightly altered version in (9), which appeared in a Lebanese newspaper, show exactly this difference of word order found in the respective spoken dialects:

(8) MSA (Egyptian)
al-ʔawrāq-i llatī **sallamat-hā** **la-hu** ʔarmalat-u
DEF-papers-OBL REL.SG.F give.PRF.3SG.F-3SG.F DAT-3SG.M widow-NOM
ʕabdi l-wahhāb
PN
'the papers, which Abdel Wahhab's widow had **given him**'

(9) MSA (Lebanese)
al-ʔawrāq-i llatī **sallamat-hu** **ʔiyyā-hā** ʔarmalat-u
DEF-papers-OBL REL.SG.F give.PRF.3SG.F-3SG.M ACC-3SG.F widow-NOM
ʕabdi l-wahhāb
PN
'the papers, which Abdel Wahhab's widow had **given him**'

Wilmsen (2010: 114–115) goes on to examine three newspapers (the London-based, largely Lebanese, *al-Ḥayāt* of the years 1996–1997; the Syrian *al-Θawra* of the year 2005 and the Egyptian *al-ʔAhrām*), and shows that with the two most common verbs in the corpus with such argument structure (*manaḥa* 'to grant' and *ʔaʕṭā* 'to give'), the trend is consistently in favour of the pattern found. The recipient–theme order is overwhelmingly favoured in the Levantine newspapers, while the theme–recipient order is clearly favoured by the Egyptian newspaper. The results are reproduced in Tables 1 and 2.

Table 1: Occurences of theme–recipient and recipient–theme order with *manaḥa* 'to grant'

Database	theme–recipient	recipient–theme
al-Ḥayāt 96	29	56
al-Ḥayāt 97	27	52
al-Θawra	27	66
al-ʔAhrām	44	8

Table 2: Occurrences of theme–recipient and recipient–theme order with *ʔaʕṭā* 'to give'

Database	theme–recipient	recipient–theme
al-Ḥayāt 96	11	23
al-Ḥayāt 97	8	22
al-Θawra	9	38
al-ʔAhrām	33	2

From this data it is clear that the dialectal background of the author of an MSA text can indeed play a role in how its syntax is constructed, despite both resulting sentences being grammatically acceptable in CA/MSA.[10]

This (and any contact phenomenon in MSA–dialect diglossia) should be seen as a case of imposition, where the dialect SL, in which the speakers/writers are dominant, has influenced the MSA RL.

It stands to reason that such syntactic research could be undertaken with CA works as well. Taking into account the biographies of authors, it might be possible to find similar imposition effects that can be connected to different dialects and languages in former times. To my knowledge, however, this work has yet to be undertaken.

3.4 Lexicon

In terms of lexicon, Jeffery's indispensable (2007 [1938]) study of the foreign vocabulary in the Quran allows us to examine some of the important sources of lexical influence on pre-Classical Arabic. Influence from Greek, Aramaic, Gəʕəz and Persian are all readily recognizable.

[10] Other works that discuss clear cases of country-specific language use of MSA include Ibrahim (2009), Parkinson (2003), Parkinson (2007) and Parkinson & Ibrahim (1999).

3.4.1 Gəʕəz

Nöldeke (1910) is still one of the most complete and important discussions of Gəʕəz loanwords in CA. Both Gəʕəz and Arabic display a significant amount of religious vocabulary that is borrowed from Aramaic. It is quite often impossible to tell whether Arabic borrowed the word from Gəʕəz or from Aramaic. Such examples are *ṭāɣūt* 'idol', Gz. *ṭaʕot*, Aram. *ṭāʕū* 'error, idol' (Nöldeke 1910: 48); *tābūt* 'ark; chest', Gz. *tabot* 'ark of Noah, ark of the covenant', Aram. *tēḇō* 'chest; ark' (Nöldeke 1910: 49).

There is religious vocabulary that is unambiguously borrowed from Gəʕəz, e.g. *ḥawāriyyūn* 'disciples' < Gz. *ḥäwarəya* 'apostle' and *muṣḥaf* 'book (esp. Qur̲an)' < Gz. *mäṣḥäf* 'scripture', but there is also religious vocabulary borrowed unambiguously from Aramaic, e.g. *zakāt* 'alms' < Aram. *zāḵū* 'merit, victory'; *sifr* 'large book' < Aram. *sp̄ar, sep̄rā*. It is therefore just as likely that Arabic would have borrowed such Aramaic loanwords via Gəʕəz as directly from Aramaic.

Some religious vocabulary from Aramaic and Hebrew can be shown to have arrived in Arabic through contact with Gəʕəz, since these words have undergone specific phonetic developments shared between CA and Gəʕəz but absent in the source language. As these often involve core religious vocabulary, and the Christian Axumite kingdom was established centuries before Islam, it seems reasonable to assume such words to be borrowings from Gəʕəz into CA, e.g. CA *ǧahannam* 'hell' < Gz. *gähännäm* (but Hebrew *gehinnom* and Syriac *gehannā*) and CA *šayṭān* 'Satan' < Gz. *śäyṭan* (but Hebrew *śåṭån* and Syriac *sāṭānā*).[11]

3.4.2 Aramaic

As already remarked upon by Retsö (2011), Aramaic loanwords in CA often have an extremely archaic character. The Aramaic variety that influenced Qur̲anic and pre-Classical Arabic had not undergone the famous *bəḡaḏkəp̄aṯ* lenition of postvocalic simple stops, nor had it lost short vowels in open syllables. This necessarily means that the form of Aramaic that influenced Qur̲anic and Classical Arabic, even the religious vocabulary, cannot be Syriac, which almost certainly underwent both shifts before becoming a dominant religious language. The *bəḡaḏkəp̄aṯ* spirantization can be dated between the first and third centuries CE, and the syncope of short vowels in open syllables takes place sometime in the middle of the third century (Gzella 2015: 41–42). However, Classical Syriac itself, as an important vehicular language of Christianity, only emerges in the fourth century CE, well after these developments had taken place (Gzella 2015: 259).

[11] Leslau (1990) often reverses the directionality of such borrowings, though without an explanation as to why he thinks a borrowing from CA into Gəʕəz is more likely.

Had *bəgāḏkəp̄aṯ* taken place, we would expect Syr. *ḡ*, *ḏ*, *ḵ*, and *ṯ* to be borrowed with their phonetic equivalents in CA: *ɣ*, *ð*, *ḫ*, and *θ* respectively.[12] This, however, is not the case; instead these consonants are consistently borrowed with the stop equivalents *ǧ*, *d*, *k*, and *t*, and without the loss of vowels in open syllables, clearly showing that these Aramaic loanwords predate the phonetic developments in Classical Syriac.

(10) a. *malakūt* 'kingdom', Syr. *malkūṯ-ā* 'kingdom' < *malakūt-ā
b. *malik* 'king', Syr. *mleḵ* 'king' < *malik[13]
c. *masǧid* 'place of worship, mosque', Syr. *masgeḏ-ā* 'place of worship' < *masgid-ā

Even the proper names of Biblical figures have a markedly un-Syriac form.

(11) a. *zakariyā*, *zakariyāʔ*, Syr. *Zḵaryā* < *zakaryā
b. *mīkāʔīl, mīkāʔil*,[14] Syr. *mīḵāʔel* < *mīkāʔēl

In other words, far from Syriac being "undoubtedly the most copious source of Qurʾānic borrowings" (Jeffery 2007 [1938]: 19), the Aramaic vocabulary in the Quran seems to not be Syriac at all.[15] Any isogloss that would allow us to identify it as such is conspicuously absent. This has important historical implications, as the presence of supposed Syriac religious vocabulary in the Quran is viewed as an important indication that Syriac Christian thought had a pronounced influence on early Islam (e.g. Mingana 1927: 82–90; Jeffery 2007 [1938]: 19–22).[16] While

[12]Retsö (2011) suggests that *b̲* could also be borrowed as *w*. This might be true, but at least the phonetic match in this case is not perfect.

[13]This word is not recognized as an Aramaic loanword by Jeffery (2007: 270), but it likely is. All the Semitic cognates of this noun are derived from a form *malk, which should have been reflected in CA as *malk*. However, we find it with an extra vowel between the last two root consonants. This can be best understood as the epenthetic vowel insertion as it is attested in Aramaic which was then subsequently borrowed with this epenthesis into Arabic. I thank Ahmad Al-Jallad for pointing this out to me.

[14]Most readers of the Quran read either *mīkāʔīl* or *mīkāʔil*, only the most dominant tradition today, that of Ḥafṣ, reads it in the highly unusual form *mīkāl* (Ibn Muǧāhid no date: 166).

[15]Note that Jeffery (2007 [1938]: 19) explicitly states that by Syriac he means any form of Christian Aramaic, so, besides Syriac, most notably also Christian Palestinian Aramaic. However, this caveat hardly solves the chronological problem, as the latter rises to prominence even later.

[16]Even if we were to accept the possibility that the dating of the lenition and syncope is somehow off by several centuries, the suggestion that "it is possible that certain of the Syriac words we find in the Qurʾān were introduced by Muḥammad himself" (Jeffery 2007 [1938]: 22) must certainly be rejected. In the grammatical works of Jacob of Edessa (640–708 CE) we have an unambiguous description of the lenition of the consonants (Holger Gzella p.c.). It seems highly unlikely that a wholesale lenition took place in only a few decades between the composition of the Quran and the time of his writings.

this is of course still a possibility, this has to be reconciled with the fact that the majority of clearly monotheistic religious vocabulary was already borrowed from a form of Aramaic before the rise of Syriac as a major religious language.

This does not mean that CA is completely devoid of Aramaic loanwords that have undergone the lenition of the consonants, and several post-Quranic loanwords have been borrowed from a variety which, like Syriac, had lenited its stops, e.g.:

(12) a. *tilmīð* 'student' < Syr. *talmīḏā* (Fraenkel 1886: 254)
b. *tūθ, tūt* 'mulberry' < Syr. *tūṯā* (Fraenkel 1886: 140)
c. *ḥiltīθ, ḥiltīt* '*asa foetida*' < Syr. *ḥeltīṯā* (Fraenkel 1886: 140)
d. *kāmaḥ, kāmiḥ* 'vinegar sauce' < Syr. *kāmḵā* (Fraenkel 1886: 288)
e. *karrāθ, kurrāθ* 'leek' < Syr. *karrāṯā* (Fraenkel 1886: 144)

It is interesting to note that Aramaic loanwords in Gəʕəz reflect a similar archaicity, in those cases where this is detectable. The expected lenited *ḵ* is not represented with Gəʕəz *ḫ* but with *k*, and short vowels in open syllables are retained. This might suggest that, when looking for religious influences on Islam, we should rather shift our focus to the south, where during the centuries before Islam both Judaism and Christianity were introduced, presumably through the vector of Gəʕəz. Some examples of such similarly archaic Aramaic loanwords in Gəʕəz are cited by Nöldeke (1910: 31–46), e.g.:

(13) Gəʕəz
a. *mälʔäk* 'angel', cf. CA *malak*, Syr. *malʔaḵ-ā* < *malʔak-ā
b. *mäläkot* 'kingdom', cf. CA *malakūt*, Syr. *malkūṯ-ā* < *malakūt-ā
c. *ḥämelät* 'mantle, headcloth', Syr. *ḥmīlṯ-ā* < *ḥamīlat-ā
d. *näbīy* 'prophet', cf. CA *nabiyy*, *nabīʔ*, Syr. *nḇīyyā* < *nabīʔ-ā
e. *mäsīḥ* 'Messiah', cf. CA *al-masīḥ*, Syr. *mšīḥ-ā* < *masīḥ-ā
f. *siʔol* 'hell', cf. Syr. *siwūl* < *siʔūl (cf. Hebr. *səʔol*)
g. *ʔärämi, ʔärämāwi, ʔärämay* 'heathen', cf. Syr. *ʔarmāy-ā* < *ʔaramāy-ā
h. *mänarät, mänarat* 'candlestick', cf. CA *manārah*, Syr. *mnārṯ-ā* < *manārat-ā

As of yet, there is not a clear historical scenario that helps us better understand how both CA and Gəʕəz, and, from the scanty information that we currently have, also Old South Arabian, ended up with similarly archaic forms of

Aramaic. This seems to suggest an as yet unattested, very archaic form of Aramaic in South Arabia. Alternatively, the syncope and lenition so well-known in Syriac may have had a much less broad distribution across the written Aramaic dialects than previously thought.

3.4.3 Greek (and Latin)

Besides this noticeable cluster of Aramaic and Gəʕəz words, there are of course also Greek loanwords in CA, generally in the semantic fields of economy and administration. Very often Aramaic likewise has these words, and it is usually not possible to decide whether Arabic borrowed the word from Aramaic or directly from Greek. The former direction is presumably more likely considering the broad presence of Aramaic as a lingua franca. Some examples are e.g. *dīnār* 'dinar', Aram. *dēnār*, Gk. *dēnárion*, Lat. *denarius*; *zawǧ* 'spouse, pair', Aram. *zōḡ* 'id.', Gk. *zeûgos* 'yoke'; *ṣirāṭ* 'way', Aram. *ʔeṣṭrāṭ* 'street', Gk. *stráta*, Lat. *(via) strata*; *qirṭās* 'parchment, papyrus', Aram. *qarṭīs*, Gk. *kʰártēs*; *qaṣr* 'castle', Aram. *qaṣrā*, Gk. *kástron*, Lat. *castrum*; *qalam* 'reed-pen', Gk. *kálamos* 'reed-pen'.[17]

A new influx of mostly philosophical and scientific Greek vocabulary entered CA during the early Abbasid period (mid 8th–10th centuries), at the time of the Graeco-Arabic translation movement (Gutas 1998). Once again, these words seem to have entered the language through Syriac (Gutas 2011). From this translation movement, we have words such as *ǧins* 'genus' < Syr. *gensā* < Gk. *génos*; *faylasūf* 'philosopher' < Syr. *pīlōsōp̄ā* < Gk. *pʰilósopʰos*; *kīmyāʔ* 'alchemy' < Syr. *kīmīyā* < Gk. *kʰēmeía*; and *ʔistāðiyā* 'stadium'[18] < Syr. *estaḏyā* < Gk. *stádion*.

3.4.4 Old South Arabian

It is often difficult to establish from which of the South Arabian languages a certain word originates. As Old South Arabian retained all the Proto-Semitic consonants, a borrowing from Old South Arabian or an inheritance from Proto-Semitic is often difficult to distinguish in CA. While Jeffery (2007 [1938]: 305) identifies a fair number of possible words of South Arabian origin, hardly ever

[17] Nöldeke (1910: 50) argues that the CA *qalam* must come from Greek through Gz. *qäläm*. While this is possible, there is nothing about this word that requires us to assume this directionality, nor is it particularly unlikely that CA and Gəʕəz independently borrowed this word without its Greek ending *-os*.

[18] Note here the apparent application of the Syriac lenition being borrowed as such in Arabic, unlike earlier loans. But it may also be possible that the lenition is part of the Greek lenition of the *delta* instead, as we see it today in Modern Greek.

does this seem the only possibility. Another issue with identifying South Arabian loanwords is that we have very scanty knowledge of its vocabulary or its linguistic developments. As a result, Old South Arabian identifications can be quite difficult to substantiate.

In recent years several lexical studies have tried to draw connections between Old South Arabian and Arabic vocabulary, but this is often based on certain semantic extensions or uses of words as described in CA dictionaries. While these observations may eventually be proven correct, it is somewhat difficult to evaluate whether we are truly dealing with borrowings in these cases, and the extremely limited knowledge that we have of the vowel system of the different Old South Arabian languages makes it difficult to evaluate this in detail. Several interesting suggestions are given by Weninger (2009), Hayajneh (2011) and Elmaz (2014; 2016).

To illustrate the difficulties we run into when trying to identify Old South Arabian borrowings in Arabic, let us examine the word *tārīḫ* pl. *tawārīḫ* 'date'. From the perspective of CA morphology, *tārīḫ* could only be a hypocorrect form of *taʔrīḫ* – which is indeed an attested biform of *tārīḫ*. The existence of the plural *tawārīḫ* rather than *taʔārīḫ*, however, seems to suggest that *taʔrīḫ* is rather a hypercorrect insertion of *hamzah* from an original form *tārīḫ*, which certainly looks foreign in its formation.

Both Hebbo (1984: 27) and Weninger (2009: 399) have suggested that this word is to be connected with the the widespread Semitic root *√wrḫ*, related to 'month' or 'moon' (cf. Hebrew *yɛraḥ* < *warḫ 'month'), which exists in Old South Arabian but not in CA.[19] The verb *ʔarraḫa* 'to date' would then reasonably be taken as a backformation from *tārīḫ*.

However, this explanation still leaves us with many problems. There is perhaps some reason to suppose that in Old South Arabian *aw would have collapsed to an unknown monophthong (Early Sabaic *ywm* 'day'; Late Sabaic *ym*). This might explain why the word is *tārīḫ* and not **tawrīḫ, but *tārīḫ* is not actually attested in Old South Arabian. So while the suggestion is certainly possible, it seems that another of the many non-Arabic Ancient northern Arabian epigraphic languages could likewise have been an origin. Barring further discoveries, many such proposed etymologies remain highly speculative, and drastically simplify the rather complex multilingual situation of pre-Islamic Arabia, where many other sources besides Old South Arabian remain possible (Al-Jallad 2018).

[19]Note, however, that the root *√wrḫ* 'month' is attested unambiguously in the singular (*wrḫ*), dual (*wrḫn*) and plural (*ʾrḫ*) in the Old Arabic corpus of Safaitic inscriptions (Al-Jallad 2015: 353).

3.4.5 Persian

Whereas with the advent of Islam the influence of Aramaic, Greek and Gəʕəz on CA quickly diminished and disappeared, the influence of Persian actually increased. While the Quran already contains a sizeable number of Persian borrowings, this only increases in the following centuries.

Some clear Persian borrowings in the Quran include: *ʔistabraq* 'silk brocade', cf. New Persian *istabra* (Eilers 1962: 204); *numruq* 'cushion' < Middle Persian *namrag*; *kanz* 'treasure' < Middle Persian *ganz/ganǧ* 'treasury' (Eilers 1962: 206). Outside of the Quran many other Persian words may be found in Arabic, e.g. *dīwān* 'archive, collected writings' < Early New Persian *dīwān* (Eilers 1962: 223), *banafsaǧ*, *manafsaǧ* 'violet' < Middle Persian *banafš* (Eilers 1971: 596); *barnāmaǧ* 'program' < Middle Persian *bārnāmag* (Eilers 1962: 217-218); *wazīr* 'minister' < Middle Persian *wizīr* (Eilers 1962: 207).[20]

3.4.6 Ottoman Turkish

The influence of Ottoman Turkish on MSA is significantly less than on the modern Arabic dialects, largely due to linguistic purism (Procházka 2011). Words that have entered MSA are words related to administration, technology and food, but also several other origins are found. For example: *damɣa* 'stamp' < *damga*; *ǧumruk* 'customs' < *gümrük* (ultimately from Latin *commercium*); *bāšā* 'pasha' < *paşa*; *bābūr* < *vapur* 'steam ship' (ultimately from French [*bateau à*] *vapeur*); *quṣāǧ* 'pliers' < *kıskaç*; *balṭa* 'axe' < *balta*; *šāwurma, šāwirma* 'lamb, etc., roasted on a spit' < *çevirme*; *qāwurma, qāwirma* 'fried meat' < *kavurma*; *kufta* 'meatballs' < *köfte*.

Of some interest is the *-ci* suffix that denotes professions and characterizations in Turkish. This suffix has developed some amount of productivity in modern dialects (especially in Iraq, Syria and Egypt), where it may even be suffixed to nouns of non-Turkish origin. In MSA the suffix is attested not infrequently, although it would probably go too far to say that it is productive. Some examples are *nawbatǧī* 'on duty; command of the guard' < *nawba* 'shift, rotatation' + *-ci*; *qahwaǧī* 'coffeehouse owner' < *qahwa* 'coffee' + *-ci*; *xurdaǧī* 'dealer in miscellaneous smallwares' < *hordaci* 'id.'; *balṭaǧī* 'sapper, pioneer' < *baltaci* 'sapper'; *būyāǧī* 'painter, bootblack' < *boyaci* 'painter'.

3.4.7 Influence of Standard Average European

A rather different, but nevertheless important factor of language contact for MSA, especially in the journalistic style, was described by Blau (1969). Blau argues

[20] I thank Chams Bernard for updating the transcription of the Middle Persian forms.

that, under the influence of what he dubs "Standard Average European" (SAE; cf. Whorf 1956), MSA (as well as Modern Hebrew) has taken on a large amount of vocabulary,[21] phraseology, and syntax similar to the journalistic language use of European languages, though the actual languages of influence could be quite different in different countries (e.g. Russian and Yiddish for Modern Hebrew; English for Egyptian MSA, French for Lebanese, Moroccan, Tunisian and Algerian MSA).[22] Examples of such influence take up over a hundred pages in Blau's pioneering work.

Blau identifies examples of lexical expansion of existing words to include lexical associations present in SAE, e.g. *saṭḥī* 'flat' is extended in meaning towards 'superficial' due to influence of, e.g. French *superficiel* and German *oberflächlich* (Blau 1969: 65); *ǧaww* 'air, atmosphere' comes to be used in a metaphorical sense in the same way English uses 'atmosphere', e.g. *ǧawwu s-siyāsati mukahrabun* 'the political atmosphere is electrified' (Blau 1969: 69).

Even whole phrases may show up as loan translations, such as MSA *ʔanqaða l-mawqifa* 'to save the situation', cf. French *sauver la situation*, German *die Situation retten*; MSA *qatala l-waqta* 'to kill time', cf. French *tuer le temps,* German *die Zeit totschlagen* (Blau 1969: 76). Even such highly specific metaphorical expressions as 'to miss the train', in the meaning of missing an opportunity, appears in MSA *ʔasriʕ wa-ʔillā fātaka l-qiṭāru* 'hurry, otherwise you will miss the train' (Blau 1969: 101).

Such linguistic influence, of course, does not lend itself particularly well to be classified within the framework of Van Coetsem (1988; 2000), as the writers of MSA in these cases are dominant in neither the source language(s) nor the recipient language, a situation which is a rather unique result of the Arabic diglossia in combination with the influence of foreign journalistic styles that have transformed the way in which MSA is written.

3.5 Influence of the early Islamic vernaculars

While, as a general rule, CA retains its archaic features, such as the retention of glottal stop in all positions and the lack of vowel harmony and syncope, we occasionally find single lexical items which optionally allow innovative forms which presumably stem from spoken vernaculars before the standardization of the classical language. This tends to be visible especially for words that have lost

[21] For further discussion of the development of Modern Standard Arabic technical vocabulary see Dichy (2011) and Jacquart (1994).

[22] The influence of French in terms of borrowings and adaptations is especially salient in literary Arabic as used in the Maghreb. Kropftisch (1977) is an excellent study on this topic.

the glottal stop, a feature usually attributed to the Ḥiǧāzī variety of the early Islamic period. For example, CA has *nabiyy* 'prophet', *nubuwwah* 'prophethood' from the root *√nbʔ*;[23] likewise *bariyyah* 'creature' from the root *√brʔ*.[24]

The likely loss of postconsonantal *ʔ* in Ḥiǧāzī Arabic has influenced the way the verb *raʔā* 'to see' (*√rʔy*) is conjugated. Its imperfect irregularly loses the *ʔ*: *yarā* 'he sees'. Similarly the verb *saʔala* 'to ask' (*√sʔl*) has two different imperatives, either the regular *isʔal* or the Ḥiǧāzī *sal* (< *sʔal). The imperative *ʔalik* 'send!' must be the imperative of an otherwise unattested verb *ʔalʔaka 'to send', which has likewise irregularly lost its postconsonantal *ʔ*. Besides verbs, we may also see the irregular lack of representation of post-consonantal *ʔ* in other nouns, e.g. *malak* 'angel', which, considering its plural *malāʔikah* and etymological origin, was presumably originally *malʔak.

The pseudo-verbs *niʕma* 'what a wonderful …' and *biʔsa* 'what an evil …', are presumably originally from *naʕima and *baʔisa, with vowel harmony and syncope. These original forms have disappeared from the classical language in their pseudo-verbal use, only retaining their verbal meaning: *naʕima* 'to be happy, glad' and *baʔisa* 'to be miserable, wretched'. However, other pseudo-verbs retain both unharmonized and unsyncopated forms as optional variants even in their pseudo-verbal use: *ḥasuna, ḥusna, ḥasna* 'how beautiful, magnificent', and *ʕaḏuma, ʕuḏma, ʕaḏma* 'how powerful, mighty'. Such syncopated and harmonized forms are claimed by the Arab grammarians themselves to be part of the eastern dialects, and absent in the Ḥiǧāzī dialects (Rabin 1951: 97), but surprisingly are retained for such pseudo-verbs.

Syncopated forms, while reported for regular verbs as well by the Arab grammarians (e.g. *šihda* or *šahda* for *šahida*), never occur in the Classical language. For some CaCiC nouns, syncopated forms are reported by lexicographers (e.g. *katf* and *kitf* besides *katif*), but it is not clear whether these syncopated forms are used in CA outside of these lexicons.

These kinds of dialectal forms that appear to have been incorporated into CA are indicative of the artificial amalgam that makes up the language, and require a much more in-depth discussion than the present chapter allows. It seems clear that the vast amount of dialectal variation that is described by the Arab grammarians, judiciously collected by Rabin (1951), does not end up in CA, but some amount of variants are either allowed, or are the only possible form present in the standard. The exact parameters that determine how and why such dialectal forms were incorporated into the language are currently unclear.

[23] In several Quranic reading traditions these are still read *nabīʔ* and *nubūʔah*, as expected (Ibn Muǧāhid no date: 106–107).

[24] Read as *barīʔah* in several Quranic reading traditions (Ibn Muǧāhid no date: 693).

4 Conclusion

Due to CA and MSA being almost exclusively High literary registers, with no true native speakers, the type of language contact that we see in the Islamic period is rather different from what we may see in more natural language contact situations. We mostly see imposition of certain dialectal forms onto the Classical ideal. An interesting exception to this is the calquing of MSA words and phraseology upon "Standard Average European", where the speakers are dominant in neither the recipient nor the source language.

Borrowing can be detected in phonology, morphology and vocabulary from Greek, Aramaic and Ethio-Semitic from the pre-Islamic period, which were then inherited by CA. In the Islamic period, it is mostly vocabulary that is borrowed, with a significant number of loans coming from Greek, Persian and Ottoman Turkish into CA.

Examining these pre-Islamic borrowings, it has become clear that the Aramaic that has primarily influenced CA, contrary to what is popularly believed, was not a form of Syriac, but rather a more archaic variety. The historical implications of this have not yet been well-integrated into our understanding of pre-Islamic linguistic diversity in Arabia and neighbouring regions.

While some studies have looked at syntactic imposition of the spoken dialects onto MSA with promising results, this has not yet been applied to medieval texts written in CA. Nevertheless, considering the clear ethnic and geographic diversity of writers of CA, it seems likely that future work should be able to detect such influences even in the medieval period.

Further reading

- Jeffery (2007) [1938] is still one of the most comprehensive books on loanwords in Quranic Arabic.
- Hebbo (1984) is an in-depth study of foreign words as they appear in the Sīrah of Ibn Hišām.
- Fraenkel (1886) is an in-depth discussion of Aramaic loanwords in Arabic, but in some respects outdated.
- Nöldeke (1910) contains an important section on loanwords both from Arabic to the Ethio-Semitic languages and the other way around.
- Blau (1969) is a pioneering work researching the interaction between European literary languages and the effects they have on the literary style of Modern Standard Arabic and Modern Hebrew.

➢ The chapters on language contact in the *Encyclopaedia of Arabic Language and Linguistics* are also highly useful and informative, and contain many up to date references for contact with Greek (Gutas 2011), Persian (Asbaghi 2011), Aramaic Retsö (2011), and Turkish (Procházka 2011).

Acknowledgements

I thank Stefan Procházka, Christopher Lucas, Maarten Kossmann and Ahmad Al-Jallad for providing me with important references, comments and suggestions.

Abbreviations

*	reconstructed form	M	masculine
**	unattested form	MSA	Modern Standard Arabic
1, 2, 3	1st, 2nd, 3rd person	NOM	nominative
ACC	accusative	OBL	oblique
Aram.	Aramaic	PL	plural
CA	Classical Arabic	PN	personal name
CE	Common Era	PRF	perfect (suffix conjugation)
DAT	dative	REL	relative pronoun
F	feminine	RL	recipient language
Gk.	Greek	SG	singular
Gz.	Gəʕəz	SL	source language
IMPF	imperfect (prefix conjugation)	Syr.	Syriac
Lat.	Latin	v.n.	verbal noun

References

Al-Jallad, Ahmad. 2015. *An outline of the grammar of Safaitic inscriptions*. Leiden: Brill.

Al-Jallad, Ahmad. 2017. Graeco-Arabica I: The southern Levant. In Ahmad Al-Jallad (ed.), *Arabic in context: Celebrating 400 years of Arabic at Leiden University*, 99–186. Leiden: Brill.

Al-Jallad, Ahmad. 2018. What is Ancient North Arabian? In Daniel Birnstiel (ed.), *Re-engaging comparative Semitic and Arabic studies*, 1–44. Wiesbaden: Harrassowitz.

Asbaghi, Asya. 2011. Persian loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

Beeston, Alfred F. L. 1962. *A descriptive grammar of Epigraphic South Arabian*. London: Luzac.

Behnstedt, Peter. 2016. *Dialect atlas of North Yemen and adjacent areas*. Leiden: Brill. (Translated by Gwendolin Goldbloom).

Beyer, Klaus. 1984. *Die aramäischen Texte vom Toten Meer samt den Inschriften aus Palästina, dem Testament Levis aus der Kairoer Genisa, der Fastenrolle und den alten talmudischen Zitaten*. Vol. 1. Göttingen: Vandenhoeck & Ruprecht.

Blau, Joshua. 1969. *The renaissance of Modern Hebrew and Modern Standard Arabic: Parallels and differences in the revival of two Semitic languages*. Berkeley: University of California Press.

Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen. Band I: Laut- und Formenlehre*. Berlin: Reuther & Reichard.

Dichy, Joseph. 2011. Terminology. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

Diem, Werner. 1980. Untersuchungen zur frühen Geschichte der arabischen Orthographie, II: Die Schreibung der Konsonanten. *Orientalia* 49. 67–106.

Dillmann, August. 2005. *Ethiopic grammar*. 2nd edn, revised by Carl Bezold and translated by James A. Crichton. Eugene, Oregon: Wipf & Stock. Originally published in 1907 by Williams & Norgate.

Eilers, Wilhelm. 1962. Iranisches Lehngut im arabischen Lexikon: Über einige Berufsnamen und Titel. *Indo-Iranian Journal* 5(3). 203–232.

Eilers, Wilhelm. 1971. Iranisches Lehngut im Arabischen. In *Actas, IV Congresso de Estudos Árabes e Islâmicos, Coimbra-Lisboa 1 a 8 de Septembro de 1968*, 581–657. Leiden: Brill.

Elmaz, Orhan. 2014. Investigating South Arabian words in al-Khalīl's Kitāb al-ʿAyn. *Supplement to the proceedings of the Seminar of Arabian Studies* 44. 29–42.

Elmaz, Orhan. 2016. *ʿarim* – A Sabaic word in the Qurʾān? In Roswitha G. Stiegner (ed.), *Süd-Arabien / South Arabia: A great "lost corridor" of mankind*, vol. 1, 213–244. Münster: Ugarit-Verlag.

Ferguson, Charles A. 1959. Diglossia. *Word* 15. 325–40.

Ferguson, Charles A. 1996. Epilogue: Diglossia revisited. In Alaa Elgibali (ed.), *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi*, 49–67. Cairo: American University in Cairo Press.

Fischer, Wolfdietrich. 2002. *A grammar of Classical Arabic*. 3rd, revised edn. Translated by Jonathan Rogers. New Haven: Yale University Press. Originally published in 1972 by Harrassowitz.

Fraenkel, Siegmund. 1886. *Die aramäischen Fremdwörter im Arabischen*. Leiden: Brill.

Glaß, Dagmar. 2011. Creating a modern standard language from medieval tradition: The Nahḍa and the Arabic academies. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 835–44. Berlin: De Gruyter Mouton.

Gutas, Dimitri. 1998. *Greek thought, Arabic culture: The Graeco-Arabic translation movement in Baghdād and early ʿAbbāsid society*. London: Routledge.

Gutas, Dimitri. 2011. Greek loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

Gzella, Holger. 2015. *A cultural history of Aramaic: From the beginnings to the advent of Islam*. Leiden: Brill.

Hayajneh, Hani. 2011. The usage of Ancient South Arabian and other Arabian languages as an etymological source for Qurʾānic vocabulary. In Gabriel Said Reynolds (ed.), *New perspectives on the Qurʾān: The Qurʾān in its historical context 2*, 117–146. London: Routledge.

Hebbo, Ahmed. 1984. *Die Fremdwörter in der arabischen Prophetenbiographie des Ibn Hischām (gest. 218/834)*. Frankfut am Main: Peter Lang.

Huehnergard, John. 1997. *A grammar of Akkadian*. Atlanta: Scholars Press.

Huehnergard, John. 2005. Features of Central Semitic. In Agustinus Gianto (ed.), *Biblical and Oriental essays in memory of William L. Moran*, 155–203. Rome: Pontificio Istituto Biblico.

Huehnergard, John & Aaron D. Rubin. 2011. Modern Standard Arabic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 259–278. Berlin: De Gruyter Mouton.

Ibn Muǧāhid, Abū Bakr. no date. *Kitāb al-sabʕah fī al-qirāʔāt*. Šawqī Ḍayf (ed.). Cairo: Dār al-Maʕārif.

Ibrahim, Zeinab. 2009. *Beyond lexical variation in Modern Standard Arabic: Egypt, Lebanon and Morocco*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Jacquart, Danielle (ed.). 1994. *La formation du vocabulaire scientifique et intellectuel dans le monde arabe*. Turnhout: Brepols Publishers.

Jeffery, Arthur. 2007. *The foreign vocabulary of the Qurʾān*. Leiden: Brill. Originally published in 1938 by Benoytosh Bhattacharyya.

Kossmann, Maarten. 2010. Parallel system borrowing: Parallel morphological systems due to the borrowing of paradigms. *Diachronica* 27(3). 459–487.

Kropftisch, Lorenz. 1977. Der französische Einfluß auf die arabische Schriftsprache im Maghrib. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 128. 39–64.

Leslau, Wolf. 1990. *Arabic loanwords in Ethiopian Semitic*. Wiesbaden: Harrassowitz.

Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowern & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.

Mingana, Alphonse. 1927. Syriac influence on the style of the Ḳurʾān. *John Rylands Library Bulletin* 11. 77–98.

Nöldeke, Theodor. 1910. *Neue Beiträge zur semitischen Sprachwissenschaft*. Strasbourg: Karl J. Trübner.

Parkinson, Dilworth B. 2003. Future variability: A corpus study of Arabic future particles. In Dilworth B. Parkinson & Samira Farwaneh (eds.), *Perspectives on Arabic linguistics XV: Papers from the fifteenth Annual Symposium on Arabic linguistics, Salt Lake City 2001*, 191–211. Amsterdam: John Benjamins.

Parkinson, Dilworth B. 2007. Sentence subject agreement variation in Arabic. In Zeinab Ibrahim & Sanaa A. M. Makhlouf (eds.), *Linguistics in an age of globalization: Perspectives on Arabic language and teaching*, 67–90. Cairo: American University in Cairo Press.

Parkinson, Dilworth B. & Zeinab Ibrahim. 1999. Testing lexical differences in regional standard Arabic. In Elabbas Benmamoun (ed.), *Perspectives on Arabic linguistics XII: Papers from the twelfth Annual Symposium on Arabic linguistics*, 183–202. Amsterdam: John Benjamins.

Procházka, Stephan. 2011. Turkish loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

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

Ratcliffe, Robert R. 1998. *The broken plural problem in Arabic and comparative Semitic: Allomorphy and analogy in non-concatenative morphology*. Amsterdam: John Benjamins.

Retsö, Jan. 2011. Aramaic/Syriac loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

Rubin, Aaron D. 2010. *The Mehri language of Oman*. Leiden: Brill.

Ryding, Karin C. 2011. Modern Standard Arabic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 844–50. Berlin: De Gruyter Mouton.

Simeone-Senelle, Marie-Claude. 2011. Modern South Arabian. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1073–1113. Berlin: De Gruyter Mouton.

Stein, Peter. 2011. Ancient South Arabian. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1042–1073. Berlin: De Gruyter Mouton.

Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

van Putten, Marijn. 2018. Hamzah in the Quranic consonantal text. *Orientalia* 87(1). 93–120.

Weninger, Stefan. 2009. Der Jemen als lexikalisches Ausstrahlungszentrum in der Antike. In Werner Arnold, Michael Jursa, Walter W. Müller & Stephan Procházka (eds.), *Philologisches und Historisches zwischen Anatolien und Sokotra: Analecta semitica in memoriam Alexander Sima*, 395–410. Wiesbaden: Harrassowitz.

Weninger, Stefan. 2011a. Old Ethiopic. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1124–1142. Berlin: De Gruyter Mouton.

Weninger, Stefan. 2011b. Ethio-Semitic in general. In Stefan Weninger, Geoffrey Khan, Michael P. Streck & Janet C. E. Watson (eds.), *The Semitic languages: An international handbook*, 1114–1123. Berlin: De Gruyter Mouton.

Whorf, Benjamin Lee. 1956. *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. John B. Carroll (ed.). Cambridge, Massachusetts: Technology Press of MIT.

Wilmsen, David. 2010. Dialects of written Arabic: Syntactic differences in the treatment of object pronouns in Egyptian and Levantine newspapers. *Arabica* 57. 99–128.