'Middle Arabic'?

Morpho-syntactic features of clashing grammars in a 13th-century Arabian text

ALEX BELLEM & G. REX SMITH

1   Introduction

In the early years of the 13th century, a traveller from the east of the Islamic world committed to writing his experiences of his journey round the west and south of the Arabian Peninsula. Ibn al-Mujāwir's travelogue is concerned with matters of trade and commerce, agriculture, and the culture and mores of the peoples he encountered. He called it *Tārīkh al-Mustabṣir* (Ibn al-Mujāwir 1951–1954). The Arabic of the text is a mixed style, being a melange of Classical Arabic (CA), Spoken Arabic (SA) and features which are neither entirely CA nor SA. It is noteworthy that Ibn al-Mujāwir's own introduction to his text is couched in an elevated CA rhymed prose (*sajʿ*) style, a feature which in all likelihood provides evidence of his competence in CA and that the mixed style was deliberate (Smith 2008: 9–10).

This language-mixing in *Tārīkh al-Mustabṣir* is characteristic of a body of texts that go back as early as the first century BCE, and which stretch across many centuries. This language has traditionally been called 'Middle Arabic' (MA). The term was initially coined by scholars of the 19th century and persists to this day, although it is often taken to relate to chronological and historical 'middleness' rather than linguistic intermediacy, i.e. akin to 'Middle English' or 'Middle Welsh'.

We reject the term MA as being inappropriate and misleading for this style of Arabic. Our academic concern is with mixed Arabic as a literary medium, that is, as one part of the culture of literature and we therefore prefer the term Literary Mixed Arabic (LMA) in our study of *Tārīkh al-Mustabṣir* and also of the other texts which will form a part of our broader research.

An obvious feature of *Tārīkh al-Mustabṣir* is that CA appears to dominate the language style, with many non-CA features mixed into the CA base. This is typical of many LMA texts, particularly those of Muslim authors — it seems that the Judaeo-Christian LMA texts may have a greater proportion of Spoken Arabic-type features. For this reason, it is often said that

so-called MA is a continuum.[1] In many instances, the non-CA features are clearly taken wholesale from Spoken Arabic (SA), so that the language of the text is generally said to be a mix of CA and SA, and this is particularly the case with lexical items.

However, there are many non-CA features of the text of *Tārīkh al-Mustabṣir* that do not conform entirely to either CA or SA. These are the features that have previously been called 'pseudo-corrections' or 'hyper-' and 'hypo-corrections'.[2] These are, typically, morpho-syntactic features. However, their use is **not** unsystematic, contra what has been said elsewhere,[3] and since they are not necessarily unsystematic, we disagree that such features are to be analysed as errors or categorized as types of 'pseudo-corrections'.[4]

So what exactly are these features and how can we analyse them?

This paper takes some of the morpho-syntactic variables that have been argued to differentiate CA from SA[5] against which to begin to test systematically the language of Ibn al-Mujāwir's text. In this context, these variables can be seen as features according to which the norms of CA and the norms of (a given variety of) SA are highly likely not to be compatible. Thus, we take such variables to reveal a systematic divergence between two grammars, or in fact points at which they clash (since one could have either one form or the other, but not both).

This paper thus uses these variables to explore the hypothesis that the features of Ibn al-Mujāwir's text that are not entirely compatible with the norms of CA and/or SA arise from the particular strategy employed to resolve a clash between the two grammatical systems.

Finally, we should add here a caveat that the earliest MS of *Tārīkh al-Mustabṣir* dates to the late 16th century, i.e. almost 300 years after its composition. It is not possible to know the extent to which later scribes have amended the text, but it is generally consistent with what is known about LMA from other texts.

2   The features

---

[1] Lentin (2008: 216). He also says (p.219) that 'Middle Arabic pragmatically deal[s] with Arabic diglossia, by filling the space of the linguistic continuum between both polar varieties'. See also Zak & Schippers (2010:113) and Smith's review, forthcoming.
[2] Lentin (2008: 217), too, rightly disputes such analysis. See Blau (1970); also Hary (2007).
[3] E.g. Lentin (2008: 219): 'It should be kept in mind that the occurrence of these [Middle Arabic typical] features is never systematic'.
[4] Further evidence supporting the view that these are not 'errors' or 'corrections' is the *Kitāb al-Iʿtibār* of Usāmah Ibn Munqidh, which is typical of (non-Judaeo-Christian) Literary Mixed Arabic.
[5] Ferguson (1959); Cohen (1970); Versteegh (1984); but see Behnstedt & Woidich (2005), Watson (2011). See also the list, and discussion and exemplification, of typical features of 'Middle Arabic' in Lentin (2008).

There are a number of features explicitly identified first by Ferguson (1959) and added to by Cohen (1970) and Versteegh (1984) which are claimed to differentiate CA from SA. As discussed by Watson (2011: 860–861), these features are not universally present in all Arabic dialects — contra CA — and there is ever increasing counter-evidence of dialects which do not adhere to one or other of these features. Nevertheless, in contradistinction to CA, very many dialects do share most of these features, and they are generally seen as typical. For our purposes, since we are looking at a text which is composed in a language which is not entirely Classical and not entirely dialectal Arabic, such a list is a useful tool for investigating exactly how LMA can appear to be somehow between the two. Faced with a binary opposition (e.g. to have dual verb inflection or not to have dual verb inflection), which way does LMA swing? Or is there, rather, a third alternative which resolves this clash of two differing grammar systems, thereby creating forms unique to LMA? This latter would indicate that deviations from CA may not be entirely random in nature, but that there is something more systematic at play.

Watson (2011: 859–860) lays out these features which are likely to differentiate many dialects of Arabic from CA as a list of 34 variables. We focus here on some of the morpho-syntactic variables, leaving aside issues of phonology. For some of these features, we were able to perform a search of an electronic version of *Tārīkh al-Mustabṣir* which is now available online.[6] While this electronic version of the text contains differences from the original Löfgren edition (Ibn al-Mujāwir 1951–1954) from which it appears to have been copied, it was still possible to search electronically and then check against the hard copy of Löfgren. In this way, a number of interesting points have been uncovered and it is possible to identify trends. It was not really feasible to conduct a systematic search for some of the features (e.g. each occurrence of a geminate verb, as opposed to being able to search for all instances of ان- and ين- when investigating dual forms), although we had already flagged up some interesting examples arising from the list of features, which we include below.

From the list of differentiating features, those which we discuss in this paper are:
1. Duals
2. Geminate verbs (i.e. identical C_2 and C_3)
3. Form IV verbs

---

[6] The digitized, anonymously re-edited version was obtained from www.al-mostafa.com (accessed 22nd April 2013). We are grateful to Murshed al-Hakmani for finding and sending it to the first-named author. However, as noted above, it should be treated with caution, since it is not an accurate copy and contains many errors and omissions.

We also look at negation, since this was a salient feature of the text for which certain types did not concord with either CA or SA (to our knowledge).

3   The data

The Arabic examples and lexical items in the following sections are transliterated into roman orthography, as per convention, alongside the Arabic original. However, the transliteration of LMA is fraught with problems, since the Arabic script – as is usual – in *Tārīkh al-Mustabṣir* does not notate short vowels, which has the advantage for the composer of such a text that many word forms are ambiguous, with no case endings, etc. The dilemma is therefore whether to use conventional CA transliteration – thus potentially misleading the reader as to the original – or whether to attempt to mix CA transliteration with transliteration of obviously non-CA forms, e.g. (3), below, الرجل 'the man' as *al-rajulu* (CA) or *al-rajul* (SA/pausal form). We have therefore adopted what we consider a relatively neutral transliteration of each word with no final short vowels, i.e. each word in pausal form, thus: *al-rajul*. For consistency – and to avoid assumptions as to whether a form is more CA-like or more SA-like – we apply this even to triconsonantal clusters, e.g. we transliterate بنته 'his daughter' as *bint-h*. The data from *Tārīkh al-Mustabṣir* is in italics; where we discuss specifically CA forms (not data from the text), we use angle brackets, e.g. <hātayni> 'these two (f. obl.)'. The reader should, where possible, follow the Arabic script.

3.1   Number

Dual number is included in Ferguson's (1959: 620–621) list as a differential feature. He notes that adjectives, pronouns and verbs do not have a dual form in the modern dialects, only singular or plural, whereas CA has dual inflection. Further, the dual is invariant, having only the form which is equivalent to the CA oblique, i.e. CA <-ayni> / SA '-ayn'.

There are two issues with duals: first, in the case of an expressed dual-number-inflected noun, whether the inflectional suffix form conforms with CA grammar or not (i.e. nominative or oblique); second, whether any adjectives, verbs and so on carry dual inflectional agreement — as per CA norms — or whether they have non-CA number marking, such as plural.

A search of *Tārīkh al-Mustabṣir* reveals that the dual forms in the text are used mostly in accordance with the grammar of CA, with occasional exceptions. For instance, there are examples where a noun has the dual inflectional suffix *-ān / -ayn* in accordance with CA

4

grammar; where the suffix is -*ayn* (for CA genitive or accusative) it is consistent with the grammar of both CA and SA. However, there are some instances of an invariant SA -*ayn* where CA grammar norms predict *-ān*. There are occasional examples which are non-CA non-SA, e.g.

(1) فاذا اصبح خرج وترك نعلاه (L54.2)[7]

    fa-ʾidhā ʾaṣbaḥ kharaj wa-tarak naʿlā-h

    'When morning comes, he goes out and leaves behind his sandals'

In (1), *naʿlā-h* is the direct object and should thus be in the oblique form *naʿlay-h* according to both CA and SA grammar. The actual form adheres neither to CA grammar, nor the invariant SA form, yet the use of the CA nominative — clearly not SA — lends a flavour of something which feels stylistically more literary (than SA) simply by virtue of being a CA form, even though it does not accord with CA grammar here.

Agreement with dual nouns is generally as per CA, i.e. a dual suffix in the same case or a dual verb form. However, there are occasional examples of usage which are unorthodox by CA standards, e.g.

(2) وفيهم اثنان احدهما يسمى سيار والثانى مياس فسكنوا جدة (L43.4)

    wa-fī-him ithnān aḥad-humā yusammā sayyār wa-ʾl-thānī mayyās fa-sakanū juddah

    'There were two [men], one of them called Sayyār and the other Mayyās, and they settled in Jeddah'

The data in (2) is interesting because it shows CA nominative dual *ithnān* and a following (CA) dual pronominal suffix in *aḥad-humā*. However, the verb which follows is masculine plural *sakanū*, which is typically a dialectal agreement pattern. The sentence therefore displays specifically CA features and specifically SA features, so that overall the effect is somewhere between the two, as if an intermediate register.

---

[7] References are given to the Löfgren edition of *Tārīkh al-Mustabṣir* (Ibn al-Mujāwir 1951–1954), by page and then line number.

More interesting, however, was one counter-example to this:

(3) ضفرت شعرتها دبوقتان وتشد كل دبوقة منهما ... يمسك الرجل تلك الدبوقتين ولا يزال يمدهما الى ان يقلعهما

من الاصل (L56.10–13)

...ḍafarat shaʿrat-hā dabbūqatān wa-tashidd kull dabbūqah min-humā... yamsik al-rajul tilk al-dabbūqatayn wa-lā yazāl yamudd-humā ʾilā ʾan yaqlaʿ-humā min al-ʾaṣl

'…[the bride has] plaited her hair [into] two plaits, and let each of the two plaits down... the man grasps those plaits and keeps pulling until he pulls them out by the roots'

The *dabbūqatān* of the first line should be an accusative, by the norms of CA, i.e. *dabbūqatayn*, which would also have concurred with the norms of SA (generally, invariant *-ayn*). However, the form used is an inappropriate CA nominative case; thus it is associated with CA but not correct by CA grammar, as in example (1) above, for *naʿlā-h*. In (3), there are a number of pronominal suffixes which are dual form, as per CA grammar, and then another interesting form *tilk al-dabbūqatayn*, where the demonstrative pronoun is feminine singular (instead of CA feminine dual <taynika>). Meanwhile, the *dabbūqatayn* is oblique, as per CA norms.

Thus, the noun phrase (NP) *tilk al-dabbūqatayn* carries a grammatical feature which is SA and one which is CA. In the second of these instances the grammar of CA clashes with the grammar of SA, so the resolution is to use one feature of each, and this thereby mixes the two grammatical systems. The very interesting example is the first, in which *dabbūqatān* is presumably associated with CA grammar, yet used in the wrong context, which avoids a form which is both CA or SA; since there are also definite CA forms *-humā*, the unexpected nominative flags up this phrase as non-SA, but also non-CA since the nominative is not grammatical. What is interesting is that this is a case where the two grammars overlap — they do not clash here. A feature is therefore changed to create a LMA intermediate form which mimics a clash resolution.

3.2   Form IV verbs

The list of features which are often thought to vary between CA and SA includes form IV verbs, i.e. of the pattern <ʾafʿala>. It is noticeable in the text that there are many occurrences

of form IV verbs, which is strikingly CA rather than (typically) SA. For instance, ʾanfadẖ is attested many times, with the meaning 'he sent':

(4) (L45.9) أنفذ صاحب مكة الى شيخ التجار

ʾanfadẖ ṣāḥib makkah ʾilā shaykẖ al-tujjār

'The lord of Mecca sent a message to the shaykh of the merchants'

It is thus very interesting to observe a form which is ungrammatical by CA norms but not an obvious SA form:

(5) (L70.11) فقال ما فعل الله بزبا؟ فقال بيد, اى هلك فسمى البلد زب بيد

fa-qāl: mā faʿal Allāh bi-zabā? fa-qāl: bīd, ayy halak, fa-summī ʾl-balad zab bīd

'[Someone] asked: "What became of [that ruler] Zabā?" He replied: "bīd," i.e. he was annihilated, so the area was called "Zab bīd" (for Zabīd)'

In (5), bīd is potentially puzzling, since it is a form I intransitive verb in CA, which therefore cannot be passivized. The internal passive is not typically SA, therefore bīd appears neither CA nor SA and it is not clear what it means, according to the 'rules' of either. However, with the view that there are expressions in the text which appear to be both CA and SA, i.e. to carry a feature of each simultaneously, the word makes more sense. There are many attestations in LMA of form I verbs used in place of form IV verbs (which, according to the perception of what is typically SA rather than CA, as exemplified by our 'list'). This is noted by Hary (2007: 277), who notes the use of bād 'he destroyed' (for *ʾabād) in a Christian LMA text. Used transitively, this verb can then undergo passivization with the CA feature of the internal passive, and the meaning 'he was annihilated / destroyed' is now clear. This means that the form is in fact a neat amalgam of a CA feature and a SA feature, therefore mixed, rather than 'neither CA nor SA'.[8]

An analogous example of an unexpected form IV verb which ends up bearing both CA and SA features (and thus truly mixed) is as follows:

(6) (L23.15) الا ترى أنا قد أضرينا بهذا الرجل

ʾa-lā tarā ʾannā qad ʾaḍarraynā bi-hādẖā ʾl-rajul

'Do you not realise that we have done this man harm?'

---

[8] Such forms have been categorized in the literature as 'hypocorrections', i.e. 'half corrected' to CA, but the 'correction' is not taken far enough. This view is less popular than it once was (see e.g. Lentin 2008; compare Blau 1970); it should be clear from our discussion that we disagree in the strongest terms with this analysis.

The verb ʾaḍarraynā is clearly a form IV verb, thus associated with CA rather than SA; however, it also exhibits non-CA treatment of geminate radicals, with the colloquial -ay- following the geminate -rr-. This is analogous with colloquial forms of verbs such as marr 'to pass', which has the CA 1st person singular perfect form <marartu> 'I passed', but the colloquial form marrayt.[9] In this way, the verb explicitly signals itself as an intermediate form, i.e. definitely LMA because it carries both CA and SA features.

3.3   Negatives

While negatives are not included specifically in the list, it is an obvious feature of dialectal Arabic that negatives differ from CA. It will suffice here to outline and exemplify briefly some of the significant aspects of negation attested in Tārīkh al-Mustabṣir. It should be noted that Lentin (2008: 221) discusses negation in LMA broadly. He notes that a typical feature of LMA is that verbal negation is predominantly with mā; lam is marked and is often used with perfect verbs, or to negate nominals. For Lentin this exemplifies a common LMA process, which is to borrow 'a linguistic tool' from CA, use it in a pseudo-genuine way (e.g. lam as negator) but without conforming to CA syntactic norms. We agree with this, but would take it a step further to say that this is an example of clash resolution which marks LMA as a distinct linguistic variety and not just the 'no language's land'.

Several features stand out in Tārīkh al-Mustabṣir. Firstly, there is not a single occurrence of lan, the CA future negator. Further, lam is used as a verbal negator with past, present and future reference. It is invariably followed by an imperfect verb form (we did not find an example of lam + perfect, which is reportedly not uncommon in LMA, cf. Lentin 2008: 221); although it is often not clear from the MS whether this imperfect is in the apocopate form, where a 'weak' root follows lam, this is sometimes apocopate (as per CA) and sometimes not apocopate (not CA, but not SA since lam is not a negator generally used in most dialectal Arabic, and is thus marked). The use of mā as a verbal negator is frequent, but almost all instances seem to be followed by ʾillā, or sometimes siwā, thus with a limiting or exceptive meaning ('only'). mā is followed by both perfect and imperfect verb forms, but its function is predominantly in the mā… ʾillā / siwā structure, and it was very hard to find exceptions — the one exemplified below is in reported speech and to negate the main verb of the apodosis of a conditional. The other negating particle is lā, which is found in the text as both a verbal and a

---

[9] In e.g. Iraqi Arabic. Cf. the well-known poem of Muẓaffar al-Nawwāb which starts with the line marraynā bī-kum Ḥamad 'We passed you by, Hamad'.

nominal negator. As a verbal negator, it negates more often imperfect verb forms, although there are a significant number of occurrences of *lā* + perfect; a notably common occurrence of *lā* is to negate *zāl* to give the meaning 'no longer, still', which appears unusual since one would probably expect CA <mā zāla / lam yazal> for past, <lā yazālu> for non-past. (Note that there are a few instances of *lam yazal* with the same meaning, and two instances of *mā zāl*.) The other use of *lā* is nominally, to negate the genus (e.g. *lā shakk* 'there is no doubt'[10]). Finally, *laysa*[11] is used as per CA to mean 'not to be' (e.g. with a following noun or prepositional phrase), and is additionally used on a significant number of occasions with a following verb in the imperfect, again a feature of CA.

These negative structures can be exemplified as follows.

(i) *lam*
a. past reference

(7) (L51.2) يقال انه كان فى قديم العهد لم يكن هذا بحر

yuqāl ʾinna-h kān fī qadīm al-ʿahd lam yakun hādhā baḥr

'It is said that in ancient times this [area] was not sea [but land]'

Noteworthy in (7) is that while the sentence is generally grammatical by CA norms (note especially the apocopate after *lam* and the past time reference), the nominal predicate of *yakun* is not marked for accusative case. Thus the negated verb phrase has both CA and SA features.

b. present reference

(8) (L25.4) ولم يورث احدهم بنته الدراهم

wa-lam yūrith ʾaḥad-hum bint-h al-darāhim

'but none of them bequeaths his money to his daughter'

The negation in (8) shows forms that would be expected as per CA grammatical norms, but the reference is clearly present-time, which is a context in which *lam* would be disallowed in CA. This is a case where the LMA in use in this text has developed a linguistic feature which is

---

[10] This is also in this text used as a pun meaning 'poking' (Smith 2008: 284).
[11] Contra the rationale outlined in §3 above for the 'neutral' transliteration of the data in *Tārīkh al-Mustabṣir*, we notate *laysa* with the final vowel that would be expected of CA, since it is unequivocally CA and in this lexical item it would seem actually wrong to use the form that our transliteration system would predict, i.e. *lays*.

neither CA nor SA, but it appears to be fairly systematic. That is, there appears to be a grammatical system here which is in this respect partially independent of both CA and SA.

c. future reference

(9) ولم تمت الى يوم القيمة (L101.10)

wa-lam tamut ʾilā yawm al-qiyāmah

'and she will not die until the day of resurrection'

The negation in (9) is similar to that in (8), in that the phrase is as per CA norms, with the exception that the time reference is clearly future, which is not normal CA usage.

(ii) *mā*

a. *mā* + imperfect (exceptive)

(10) ما أظن السيف اصله الا من الصاعقة التى ضربها يافث بن نوح (L28.6)

wa-mā ʾaẓunn al-sayf ʾaṣl-h ʾillā min al-ṣāʿiqah allatī ḍarab-hā Yāfiṯ bn Nūḥ

'I think the origin of this sword can only be the thunderbolt which Japheth, son of Noah, fashioned'

b. *mā* + perfect (exceptive)

(11) ما بقى فى الوهط من الشجر سوى شجرة توت (L28.6)

mā baqī[12] *fī ʾl-wahṭ min al-shajar siwā shajarat tūt*

'Only one mulberry tree has remained in al-Wahṭ'

c. *mā* + perfect

(12) ولو قتلتنى وأخذت الغنم ما نجوت (L19.4)

wa-law qatalt-nī wa-ʾaxadht al-ġanam mā najawt

'If you were to kill me and take the sheep, you would not get away with it'

---

[12] Hypothetically, this may be read as *baqā* for some varieties of SA. However, here we use *baqī* for two reasons. Firstly, it is consistent with our transliteration scheme, which is maximally 'neutral', and this is anyway the CA form; secondly, the SA forms in this text are south Arabian, in which this form would have a final *ī*, as in contemporary Sana'ani Yemeni Arabic *bigi* 'stay, remain' (Qafisheh 1999: 44; the phonological final short *i* is not relevant to our treatment above).

The use of *mā* in the MS is almost always in a structure with *ʾillā* or *siwā*, with very few exceptions, such as (12). There were also two instances in the whole MS of *mā zāl*, with this verb otherwise co-occurring with *lam* or *lā*.

(iii) *lā*

a. *lā* + imperfect

(13)   مكة لا تنصرف لانها مؤنثة (L2.12)

makkah lā tanṣarif li-ʾanna-hā muʾannathah

'The word "Mecca" cannot be fully inflected because it is feminine'

The negator *lā* functions in (13) as per CA norms.

b. *lā* + perfect

(14)   ولا زال الرجلان يعملان فى النقر والحفر (L108.7)

wa-lā zāl al-rajulān yaʿmalān fī ʾl-naqr wa-ʾl-ḥafr

'The men remained at work hewing and excavating'

Perfects occurred only rarely with *lā*, with the exception of *lā zāl* (as in 14), which occurred frequently.

(iv) *laysa*

a. *laysa* + non-verbal predicate

(15)   وليس هذا الفن عندهم عار (L7.17)

wa-laysa hādhā ʾl-fann ʿind-hum ʿār

'They do not regard this practice as a shameful act'

b. *laysa* + verb

(16)   ليس يحكم عليهم سلطان (L26.11)

laysa yaḥkum ʿalay-him sulṭān

'they are under the authority of no ruler'

The use of *laysa* + an imperfect verb is a CA form used as a 'strong' negative (for emphasis), and there are a number of instances in this text.

PRE-PUBLICATION VERSION WITH MINOR ERRORS

To conclude the section on negatives, the text of *Tārīkh al-Mustabṣir* shows negation which diverges from CA usage, yet is not SA. There is a noteworthy degree of systematicity, indicating that there is a linguistic system here which has its own unique features.

3.4   Ongoing investigation

(i) mood distinctions

Ferguson (1959: 622) notes that mood distinctions tend not to be encoded explicitly in dialectal verb forms. This is very relevant to the text in hand because there is much variation in 3rd person masculine plural imperfect verb forms. In some cases the oblique form is used where CA would have an indicative form, akin to the dialectal form. In many cases, a CA indicative form is used, in concordance with CA grammar. There are, additionally, quite a number of instances of the *lam yaktubūn* ('they didn't write') type, where *lam* in CA is followed by an apocopate verb form. An analysis of these is in preparation.

We have also conducted a preliminary search for imperative forms of verbs with a weak middle radical (such as *qām* 'get up', for instance). These forms in CA would be in the apocopate form (e.g. *qum* 'get up!'), while dialectal forms tend to retain the long vowel. So far, we have found one instance of *nām* 'sleep!' (for *nam*, L36.15).

(ii) numbers

Also on the list are number forms (Ferguson 1959: 624). Agreement in LMA is noted to be inconsistent, and we have already observed some unorthodox use of numbers.

(iii) word order (SVO / VSO)

Word order will be especially interesting (and time-consuming!) to investigate. The text is replete with instances of *kān* in the imperfect (*yakūn*, etc.), in contexts which are ungrammatical in CA. A number of these instances may be a way of creating verb-initial phrases where the nominal subject is otherwise topicalized and thereby fronted. Part of this investigation will also involve looking at the use of *kān* as an auxiliary, and the use of serial/ asyndetic verbs.

4   Conclusions

- The mixing exhibited in *Tārīkh al-Mustabṣir* is not entirely arbitrary. Although there is some variation (it is possible that later scribes may have contributed to this, or not),

there is a considerable degree of systematicity in some aspects related to this mixing. Other aspects seem less systematic. What does seem systematic, however, is that while the text is generally towards the CA end of the spectrum, enough non-CA features are mixed into each sentence for the mixed nature of the language to dominate. Therefore, it is overall perhaps slightly 'High' LMA.

- LMA may be mixed to varying degrees, thus being a continuum, as noted by others. It can be highly variant as a whole (all the various texts which have been categorised as some form of so-called 'Middle Arabic'). There has been work on building a typology of typical LMA features (see e.g. Lentin 2008).

- The mixing of CA and SA to create an intermediate 'mixed' variety is not like mixing e.g. Arabic and French, or Japanese and English, that is, two obviously very different systems. There is of course enough overlap between CA and the various other varieties of Arabic that the grammar is often perfectly compatible.

- This paper has attempted to give a snapshot of how we are approaching the question 'What is LMA?' from the flip side of the coin: primarily, by looking at what are considered to be typical indicators of SA in contradistinction to CA and to check these systematically against a given text (rather than observing 'interesting' features as one goes through a text). This is because these points are likely to be where CA grammar typically (although not always, for all features) clashes with the grammar of many or most modern dialects, so it is informative to see what the author uses to resolve the clash. Since there are two possibilities (the CA norm or the SA norm), one may expect that an author composing in LMA would choose either option depending on how much of a colloquializing or classicizing effect s/he was aiming for. In fact, what is interesting is that LMA often seems to 'invent' a third way, by using a grammatical feature of each (CA and SA) side-by-side, thus flagging up the particular structure as being neither CA nor SA but somehow not foreign to either because it is truly mixed.

- Since these forms are used frequently, and in some cases with a good degree of systematicity, but with not enough invariance to assume that they are simply learned and fossilized forms, we conclude that these forms are a conscious feature of LMA: thus does Ibn al-Mujāwir choose to convey his message to his audience. The idea of a 'Middle Arabic' stage in the historical development of Arabic is misleading, at best, if not simply fallacious;[13] thus we must move on from the vocabulary of error and

---

[13] As Owens (2006: 47) puts it, 'Middle Arabic … is primarily a style, not a historical stage.'

correction and we must move on from the notion of historical middleness encoded in the term 'Middle'. The term was never appropriate in the Arabic context and the more one analyses these texts the more one becomes convinced that they are not historically middle. As we have shown, this is a specific literary form of Mixed Arabic, hence the more appropriate designation Literary Mixed Arabic.

- A final interesting question to ask is whether this intermediate code, LMA, is a distinct variety or just a style (and of what)? We have shown here for *Tārīkh al-Mustabṣir* that there are many forms and structures which are not grammatical in either CA or SA (being truly mixed of features of both, simultaneously). The mixing is in this way not entirely arbitrary. This therefore begs the question, to what extent is LMA a language variety with its own linguistic system? If we think of LMA as a style, are we claiming that it does not have its own grammar? Does LMA have its own grammar?

References

Behnstedt, P. & Woidich, M. 2005. *Arabische Dialektgeographie*. Leiden: Brill.

Blau, J. 1970. *On Pseudo-Corrections in Some Semitic Languages*. Jerusalem: Israel Academy of Sciences and Humanities.

Cohen, D. 1970. Koinè, langues communes et dialectes arabes. Pages 105–125 in D. Cohen (ed.), *Études de linguistique sémitique et arabe*. The Hague and Paris: Mouton.

Ferguson, C.A. 1959. The Arabic Koine. *Language* 35.4: 616–630.

Hary, B. 2007. Hypercorrection. Pages 275–279 in K. Versteegh et al. (eds), *Encyclopedia of Arabic Language and Linguistics*, vol. II *Eg–Lan*. Leiden: Brill.

Ibn al-Mujāwir/ed. O. Löfgren. 1951–54. *Ibn al-Muǧāwir, Descriptio Arabiae Meridionalis, Ṣifat bilād al-Yaman wa-Makkah wa-baʿḍ al-Ḥiǧāz al-musammā Taʾrīkh al-Mustabṣir*. Leiden: Brill.

Lentin, J. 2008. Middle Arabic. Pages 215–224 in K. Versteegh et al. (eds), *Encyclopedia of Arabic Language and Linguistics*, vol. III *Lat–Pu*. Leiden / Boston: Brill.

Owens, J. 2006. *A Linguistic History of Arabic*. Oxford: Oxford University Press.

Qafisheh, H.A. 1999. *NTC's Yemeni Arabic–English Dictionary*. Chicago: NTC Publishing Group.

Smith, G.R. 1996. The language of Ibn al-Mujāwir's 7th/13th century guide to Arabia, *Tārīkh al-Mustabṣir*. Pages 327–351 in J.R. Smart (ed.), *Tradition and Modernity in Arabic Language and Literature*. Richmond: Curzon.

---   2008. *A Traveller in Thirteenth-Century Arabia. Ibn al-Mujāwir's* Tārīkh al-Mustabṣir. London: The Hakluyt Society, Series III, Volume 19.

---   in press. Review of Zak & Schippers 2012. *Journal of Semitic Studies.*

Usāmah Ibn Munqidh. 1930. Ph.K. Hitti (ed.), *Usāmah's Memoirs Entitled Kitāb al-Iʿtibār.* Princeton: Princeton University Press.

Versteegh, K. 1984. *Pidginization and Creolization: The Case for Arabic*. Amsterdam: Benjamin.

Watson, J.C.E. 2011. Arabic dialects (general article). Pages 851–896 in S. Weninger, G. Khan, M.P. Streck & J.C.E. Watson (eds), *The Semitic Languages: An international handbook*. Berlin: De Gruyter Mouton.

Zack, L. & Schippers, A. 2012. *Middle Arabic and Mixed Arabic: Diachrony and Synchrony.* Leiden/Boston: Brill.

Authors' addresses

Dr Alex Bellem, Modern Languages & Cultures, University of Durham, DH1 3JT, UK
e-mail alex.bellem@dur.ac.uk
Professor G. Rex Smith, Middle Eastern Studies, University of Manchester, M13 9PL, UK
e-mail grexsmith@yahoo.co.uk