

Chapter 16

Modern South Arabian languages

Simone Bettega

Università degli Studi di Torino

Fabio Gasparini

Freie Universität Berlin

In the course of this chapter we will discuss what is known about the effects that contact with Arabic has had on the Modern South Arabian languages of Oman and Yemen. Documentation concerning these languages is not abundant, and even more limited is our knowledge of the history of their interaction with Arabic. By integrating the existing bibliography with as yet unpublished fieldwork materials, we will try to provide as complete a picture of the situation as possible, also discussing the current linguistic and sociolinguistic landscape of Dhofar and eastern Yemen.

1 History of contact between Arabic and the Modern South Arabian languages

Much to the frustration of modern scholars of Semitic, the history of the Modern South Arabian languages (henceforth MSAL) remains largely unknown.¹ To this day, no written attestation of these varieties has been discovered, and it seems safe to assume that they have remained exclusively spoken languages throughout all of their history. Since European researchers became aware of their existence in the first half of the nineteenth century (Wellsted 1837), and until very recently, the MSAL were thought by many to be the descendants of the Old (epigraphic)

¹§1 was authored by Simone Bettega, while §2 was authored by Fabio Gasparini. §3 and §4 are the result of the conjoined efforts of both authors. In particular, Gasparini was responsible for analyzing most of the primary sources and raw linguistic data, while Bettega worked more extensively on the existing bibliography.



South Arabian languages (Rubin 2014: 16). This assumption has been conclusively disproven by Porkhomovsky's (1997) article, which also contributed significantly to the re-shaping of the proposed model for the Semitic family tree. This modified version of the family tree (which finds further support in the recent works of Kogan 2015 and Edzard 2017) sets the MSAL apart as an independent branch of the West Semitic subgroup, one whose origins are therefore of considerable antiquity. This brings us to the question of when it was that the MSAL (or their forebears) first came into contact with Arabic. This might have happened at any time since Arabic-speaking people started to penetrate into southern Arabia, a process that – as we know from historical records – began in the second half of the first millennium BCE (Robin 1991; Hoyland 2001: 47–48). Roughly one thousand years later, almost the whole population of central and northern Yemen was speaking Arabic, and possibly a considerable portion of the southern population as well (Beeston 1981: 184; Zammit 2011: 295). It is therefore possible that Arabic and the MSAL have been in contact for quite some time, and it seems likely that the intensity and effects of such contact grew stronger after the advent of Islam (Lonnet 2011: 247). It is also possible, as some scholars have written, that the MSAL “represent isolated forms that were never touched by Arabic influence until the modern period” (Versteegh 2014: 127). Admittedly, evidence to support either one of these hypotheses is scarce, and at present it is probably safer to say that our knowledge of the history of contact between Arabic and the MSAL before the twentieth century is fragmentary at best. This is why studies on the outcomes of such contact are of particular interest, since they could help to shed light on parts of that history. This is also why, in the course of this chapter, we will refrain from addressing the question of how contact with the MSAL affected the varieties of Arabic spoken in Oman and Yemen, and focus solely on the influence of Arabic on the MSAL. Although there is plenty of evidence that South Arabian exerted a powerful influence on the Arabic of the area (see for instance Retsö 2000 and Watson 2018),² it is often difficult to assess whether this influence is the result of contact with forms of Old South Arabian or more recent interaction with the MSAL. Such a discussion, also because of space constraints, is beyond the scope of the present article.

As far as the interaction between Arabic and the MSAL in the twentieth and twenty-first centuries is concerned, Morris (2017: 25) provides a good overview of the multilingual environment in which the MSAL were and are spoken:

²To the point that so-called mixed varieties are reported to exist, whose exact linguistic nature seems difficult to pinpoint. See Watson et al. (2006) and Watson (2011) for discussion.

Speakers of [a Modern South Arabian] language always had to deal with speakers of other MSAL, as well as with speakers of various dialects of Arabic. The Baṭāḥirah, for instance, did nearly all their trade with boats from Ṣūr and other Arabic-speaking ports; they lived and worked with the Arabic-speaking Janaba, while being in contact with speakers of Ḥarāsīs and Mahra. The Ḥarāsīs interacted with the Arabic speakers surrounding their Jiddat al-Ḥarāsīs homeland, traded in the Arabic-speaking markets of the north, and in the summer months went to work at the northern date harvest. Mehri speakers lived beside and traded with Arabic-speaking Kathīri tribesmen in the Nejd region, Ṣḥeret speakers in the mountains, and Arabic speakers in the coastal market towns of Dhofar. Ṣḥeret speakers interacted with the Mahra, some of whom settled among them, and with Arabic-speaking peoples of the coast as well as the desert interior [...] There was marriage between Arabic-speaking men of the coastal towns and MSAL-speaking women of the interior, and over time, families of Mehri and Ṣḥeret speakers settled in or near the towns, with the result that even more Arabic speakers became familiar with these languages. (Morris 2017: 25)

2 Current state of contact between Arabic and the Modern South Arabian languages

Today, six Modern South Arabian languages exist, spoken by around 200,000 people in eastern Yemen (including the island of Soqatra) and western Oman. These six languages are: Mehri, Hobyōt, Ḥarsūsi, Baṭḥari, Ṣḥeret/Jibbāli and Soqoṭri. They are all to be regarded as endangered varieties, though the individual degree of endangerment varies remarkably. No exact census concerning the number of speakers is currently available (Simeone-Senelle 2011: 1075), but we know that Mehri is the most spoken language, with an estimated 100,000 speakers. It is followed by Soqoṭri (about 50,000 speakers), Ṣḥeret (25,000), Ḥarsūsi (a few hundred), Hobyōt (a few hundred) and Baṭḥari (less than 20 speakers). The main causes of endangerment are reckoned to be shift to Arabic and the disappearance of traditional local lifestyles. In addition, the current political situation in Yemen is having effects on the linguistic landscape of the region which are difficult to document or foresee: the area is currently inaccessible to researchers, and there is no way to know how the conflict will affect the local communities.

As far as Oman is concerned, the city of Salalah undoubtedly represents the major locus of contact between Arabic and the MSAL. The rapid growth the city has witnessed in recent years, and the improved possibilities of economic development that came with it, have led many Šheret speakers from the nearby mountains to settle in the city or its immediate surroundings, where they now employ Arabic on a daily basis as a consequence of mass education and media, neglecting other local languages. This has led to a split, in the speakers' perception, between "proper" Šheret, spoken in the mountains, and the "city Šheret" of Salalah, often regarded as a sort of "broken" variety of the language in which, among other things, code-switching with Arabic is extremely frequent. Unfortunately, data on this subject are virtually non-existent, given the extreme difficulty of documenting such an episodic phenomenon (aggravated by speakers' understandable reluctance to having their imperfect language proficiency evaluated and recorded).

Even outside the urban centers, however, contact with Arabic is on the rise. Even the most isolated variety, Soqotri, is apparently undergoing rapid change under the influence of Arabic: the existence of a koinéised variety of Soqotri, heavily influenced by Arabic, has been recently reported in Ḥadibo (Morris 2017: 27). This is not to say, of course, that all MSAL are being affected to the same degree: Watson (2012: 3), for instance, notes how "Mahriyōt [the eastern Yemeni variety of Mehri] [...] exhibits structures unattested in Mehreyyet [Mehri Omani variety] [...] and shows greater Arabic influence both in terms of the number of Arabic terms used, and the length and frequency of Arabic phrases within texts." However, no MSAL seems at present to be exempt from the effects of contact.

The case of Baḥari, which, as we have seen, is the most severely endangered of all the MSAL, exemplifies well the processes of morphological loss and erosion that a language undergoes in the final stages of endangerment. Morris (2017) reports how already in the 1970s Baḥari seemed to display many of the signs of a moribund language. In recent times:

[t]he younger generations showed little interest in their former language; they were eager to embrace Arabic and to feel themselves part of the wider Arabic Islamic community; and they were proud to call themselves 'ʿarab', with all that word's overtones of Bedouin ancestry and code of honour. (Morris 2017: 11)

In the following sections we will discuss several types of contact-induced changes in the MSAL. Although we will use material taken from all varieties, Baḥari will be in particular focus due to its singular status.

3 Contact-induced changes in the MSAL

As already noted, in the course of this chapter we will focus solely on the effects that contact with Arabic has had on the various MSAL. Therefore, Arabic will always be the source language of all the transfer phenomena considered in the next pages, while the recipient language will be, depending on the different examples, one or the other of the six MSAL. Obviously, this poses the question of who the agents of change are and were in the case of these particular phenomena, and what type(s) of transfer are we confronted with (cf. Van Coetsem 1988; 2000; Winford 2005). According to the overview of the MSAL's sociolinguistic status presented above, it should be clear by now that, while the two cases are extremely common of (a) mono- or multilingual MSAL speakers who acquire Arabic as an L2 and (b) bilingual MSAL–Arabic speakers, the opposite is not true (that is, monolingual Arabic speakers who come to acquire one or more MSAL as L2s later in life). In other words, all the transfer phenomena we will be considering in the next paragraphs are either instances of borrowing (brought about by speakers who are dominant in one or more MSAL) or convergence (brought about by speakers who are native speakers of Arabic and at least one MSAL; see Lucas 2015 for a definition of convergence).

3.1 Phonology

3.1.1 Phonetic adaptation of loanwords

As illustrated in §3.4, lexical borrowings from Arabic are extremely common in the MSAL. As Morris (2017: 13) remarks, such loanwords are often altered in order for them to acquire a “South Arabian flavour”, so to speak. The phenomenon is not one of simple adaptation dictated by difficulty of articulation, since the sounds that are replaced are present in the phonological inventory of the MSAL. In fact, the opposite appears to be true, these sounds normally being replaced by others which are typical of South Arabian but absent in Arabic. For Baḥari, Morris gives the example of Arabic pharyngealised dental fricative /ð̤/ (IPA [ð̤^h]) being replaced by the pharyngealised alveolar lateral fricative /š̤/ (mostly realised as IPA [ʃ̤^h], see §3.1.4), as in *rašš* ‘bruise’ (from Janaybi Arabic *raḏḏ*), or Arabic /š/ (IPA [ʃ]) being replaced by /s/, as in *men šān-k* ‘for you, for your sake’, in place of *men šān-k*, *šarray* ‘buyer’ for *šarray*, or *šāmāl* ‘inland, north’ for *šāmāl* (while Baḥari *šēmāl(i)* is normally used to refer to the left hand only).

Lexical borrowing can also be the cause of variation in the realisation of certain sounds, as is the case with the phonemes /g/ and /y/ (IPA [g] and [j] respectively), which represent different reflexes of Proto-Semitic *g in different Omani Arabic

dialects. It is possible to find traces of this variation in those MSAL that are in contact with more than one variety of Arabic, as is the case with Ḥarsusi: see for instance *fagr* and *fayr*, both meaning ‘dawn’, or the opposition between *yann* ‘madness’ and *genni* ‘jinni’, both from the same etymological root (Lonnet 2011: 299).

3.1.2 Affrication of /k/ > [tʃ]

It can also be the case that some phonetic processes regularly taking place in the local Arabic varieties but otherwise unknown to MSAL phonology are transferred to original MSAL vocabulary. This is what happens in Baḥari, where some speakers may show an affricate realisation of the voiceless occlusive [k] > [tʃ], which resembles the Janaybi Arabic realisation of the phoneme /k/ (whose complementary distribution with the voiceless plosive realisation [k] is still unclear). For example, some speakers regularly produce /yənkaʕ/ ‘come.3SG.M.SBJV’ as [jənʔaʕ] instead of [jənʔkaʕ].

3.1.3 Stress

The structural similarity of Arabic and the MSAL can sometimes cause stress patterns which are typical of the former to be applied to the latter, as is the case with ‘she began’: Soqotri *bédʔəh*, (local) Arabic *bədəʔat*, Soqotri with an Arabic stress *bədəʔəh* (Lonnet 2011: 299).

3.1.4 Realisation of emphatics

This is a topic that has attracted the attention of several scholars since the publication of Johnstone’s (1975) article on the subject, because of the realisation of the so-called Semitic “emphatics” as glottalised consonants. Glottalisation is a secondary articulatory process in which narrowing (creaky voice) or closure (ejective realisation) of the glottis takes place: the action of the larynx compresses the air in the vocal tract which, once released, produces a greater amplitude in the stop burst (Ladefoged & Maddieson 1996: 78).

Lonnet (2011: 299) notes a tendency for speakers of various MSAL to replace the ejective articulation of certain consonants (especially fricatives, see Ridouane & Gendrot 2017) with a pharyngealised realisation, typical of Arabic emphatics. Pharyngealisation is a kind of secondary articulation involving a constriction of the pharynx usually realised through tongue-root retraction, resulting in a backed realisation (Ladefoged & Maddieson 1996: 365). This process is well-documented across Semitic languages. Naumkin & Porkhomovsky (1981) note for

Soqotri an ongoing process of transition from a glottalised to a pharyngealised realisation of emphatics, with only stops being realised as fully glottalised items. Work by Watson & Bellem (2010; 2011) and Watson & Heselwood (2016) shows the co-occurrence of pharyngealisation and glottalisation in relation to pre-pausal phenomena in Şanʿāni Arabic, Mahriyōt and Mehreyyet (respectively the westernmost Yemeni and Omani varieties of Mehri). Dufour (2016: 22) states that “le caractère éjectif des phonèmes emphatiques ne fait aucun doute, en jibbali comme en mehri” (“the nature of the emphatic phonemes is undoubtedly ejective, in Jibbali as much as in Mehri”).³ Finally, in Baḥari only /k/ is realised as a fully ejective consonant [kʰ]. /t/ and the fricative emphatics, on the other hand, are described as mainly pharyngealised (and partially voiced, in the case of fricatives; Gasparini 2017).

Unfortunately, since there is no thorough phonetic description of any MSAL that predates the 1970s, it is impossible to ascertain whether these realisations (which, again, range from fully glottalised to fully pharyngealised) are the result of the influence of Arabic, or have arisen as the consequence of internal and typologically predictable developments. It is likely, though, that bilingualism and constant contact with Arabic have at least favoured this phonetic change. Evidence in support of this view may come from the fact that speakers who are poorly proficient in Arabic and live in rural and more isolated areas are more likely to preserve a glottalised realisation of the emphatics (as emerges from direct fieldwork observations).

3.2 Morphology

3.2.1 Nominal morphology

Morphological patterns which are typical of Arabic can enter a language through borrowing, as is the case with the passive participle pattern for simple verbs, which is *mVCCūC* in Arabic and *mVCCīC* in MSAL. Soqotri *maḥlāk*, for instance, is clearly derived from Arabic *maḥlūq* ‘human being’ (lit. ‘created’), while this is not the case for Ḥarsusi *mḥəlīk* (Lonnet 2011: 299). Also, in the realm of verbal derivational morphology, certain phenomena can be introduced into the recipient language through lexical borrowing: this is the case with gemination and prefixation of *t-* in Ḥarsusi, as in the participle *məṯḥaffi* ‘barefoot’ (from Omani Arabic *mithaffi*; Lonnet 2011).

In general, Arabic loanwords are normally well integrated in MSAL morphology, probably because of the high degree of structural similarity that exists be-

³Authors’ translation.

tween these languages. One example, reported by Lonnet (2011), is that of *bəkerēt* ‘cow’, a fully integrated loan from Arabic used in Ḥarsusi and Western Yemeni Mehri, which possesses its own plural and diminutive form (*bəkār* and *bəkərənōt*, respectively).

Arabic loans in several MSAL stand out because of their characteristic feminine ending in *-V(h)* instead of *-(V)t*, as in Šehri *saṣah* ‘watch’ and *ṭorəh* ‘revolution’ (but consider the more adapted *rišt* ‘trigger’ from Omani Arabic *rīšah*; Lonnet 2011).

It is also worth noting that the Arabic ending *-V(h)* is replaced by its MSAL equivalent when the noun is in the construct state, that is, final *-t* reappears. This would also happen in Arabic, but the alteration in the quality of the vowel is a clear signal that the suffix is to be considered an MSAL morpheme. Consider the following example from Morris’ Baṭḥari recordings:⁴

- (1) Baṭḥari
 mṣayš-it-həm bəss mṣayš-it-həm ḥawla ʕār ḥāmis
 sustenance-F.CS-3PL.M only sustenance-F.CS-3PL.M once only turtle
 w-ṣayd śālā mṣayš-ah ḥawil
 and-fish nothing sustenance-F once
 ‘Their sustenance, only that! Their sustenance was once only turtle and
 fish, there was nothing to eat in the past.’

The word *mṣayšah* ‘sustenance, food’ is a loanword from Arabic (as the *-ah* ending suggests). When suffixed with the possessive 3PL.M pronoun *-həm*, however, Baṭḥari *-it* replaces *-ah/-at* (note also, in the example, the use of the restrictive adverbial particle *bəss* ‘only’, which is a well-integrated loan from dialectal Arabic and occurs in alternation with Baṭḥari *ʕār*).

Finally, in Baṭḥari the Arabic definite article (*a*)-*l-* is occasionally used instead of the MSAL definite article *a-*: *bə-l-ḥarifēt* ‘during the rainy season’.

3.2.2 Pronouns

The influence of Arabic can be observed, to an extent, even in the pronominal system, especially in those MSAL that are more exposed to contact due to the limited size of their speech communities. Lonnet (2011), for instance, reports how,

⁴Audio file 20130929_B_B02andB04_storyofcatchingturtle recorded, transcribed and kindly shared with Fabio Gasparini by Miranda Morris. The recording was produced in the context of Morris’ and Watson’s “Documentation and Ethnolinguistic Analysis of the Modern South Arabian languages” project, funded by the Leverhulme Trust. More recordings are accessible at the ELAR archive of SOAS University of London. The transcription has been adapted.

despite the fact that in the MSAL the first person suffix pronoun is normally an invariable *-i*, in Ḥarsusi this can be replaced by *-ni* after verbs and prepositions (as is the case with Arabic; see also §3.3 for another interesting example concerning the marking of pronominal direct objects).

In addition, Baḥari relative pronouns (SG: *l-*, *lī* PL: *əlli*) are close to their equivalent in Janaybi Arabic (and diverge from the rest of MSAL, where a *ḏ-* element can be found). Baḥari has also borrowed the reflexive pronoun *ʕamr-* ‘oneself’ from the Arabic dialect of the Janaba, despite the existence of an original Baḥari term with the same meaning, *ḥanef-* (note that both terms must always be followed by a suffix personal pronoun). *ʕamr-* has also been given a plural form in Baḥari, based on MSAL derivational patterns, *ḥaʕmār-* (Morris 2017: 14).

3.2.3 Baḥari verbal plural marker *-uw*

Baḥari differs from the rest of the MSAL in that all 2/3_{PL.M} verbal forms are marked by an *-uw* suffix, while in the other languages of the group these persons are marked by a *-Vm* ending and/or by internal vowel change (e.g. Mehri and Ḥarsusi *-kə(u)m* for the 2_{PL.M} and *-ə(u)m/umlaut* for the 3_{PL.M} of the perfective conjugation; *t...-ə(u)m* and *y/i...-ə(u)m* respectively for 2 and 3_{PL.M} of the imperfective conjugation; Simeone-Senelle 2011: 1093–1094).

The origin of this suffix is uncertain. Its presence might well be connected to contact with Arabic (neighboring dialects have an *-u* or *-ūn* suffix in the 3_{PL.M} person of the verb in both the perfective and imperfective conjugation) or to otherwise unattested stages of development internal to the MSAL verbal system. In this regard, Rubin (2017: 5) suggests for Mehreyyēt the presence of a subjacent *-ə-* in 2nd/3rd plural masculine verbal suffixes which could therefore be somehow related to the Baḥari *-uw* marker. However, the optional simultaneous presence of apophony within the stem of 3_{PL.M} verbal forms (similarly to what happens elsewhere in the MSAL), together with scarcity of data, prevents any conclusive assessment of the topic.

3.3 Syntax

At present, the syntax of the various MSAL has not been made the object of detailed investigation. The only scientific work dealing with this topic is Watson’s (2012) in-depth analysis of Mehri syntax. However, Watson’s thorough description provides only sporadic insights into the issue of language contact (as for instance the use in Mahriyōt, the eastern Yemeni variety of Mehri, of a *swē ~ amma... yā* construction to express polycoordination, probably to be regarded as

the result of Arabic influence; Watson 2012: 298). In general, though, the topic is left unaddressed in the literature, and more research is needed.

Gasparini's data on Baḥari offer an interesting example of Arabic influence on MSAL syntax. In Baḥari, as in the other MSAL, pronominal direct and indirect objects may require a particle *t-* to be inserted between them and the verb, depending on the morphological form of the verb itself. Masculine singular imperatives, for instance, require the presence of the marker, as shown in the following example:

- (2) Baḥari (Gasparini, unpublished data):

zum t-ī t-ih
 give.IMP ACC-1SG ACC-3SG.M
 'Give it to me.'

Example (3), in contrast, shows that the pronominal indirect object *-(ə)nī* is suffixed directly to the verb as it would be in Arabic (see §3.2.2).

- (3) Baḥari (Gasparini 2018: 66):

zəm-ənī θrɛh
 give.IMP-1SG two.M
 'Give me both (of them).'

In other words, the introduction of the Arabic form of the object pronoun has caused the Baḥari object marker to disappear. Note that informants judged the alternative construction *zum t-ī θrɛh*, (with the use of the object marker *t-* and the 1SG object pronoun marker *-ī*) to be acceptable, but this form was not produced spontaneously.

A peculiarity of the MSAL spoken in Oman is the use of circumstantial qualifiers, a type of clausal subordination well attested in Gulf Arabic (Persson 2009). Baḥari regularly introduces predictive and factual conditional clauses asyndetically by using the structure [SBJ.PRO w-SBJ.PRO]. Consider (4):

- (4) Baḥari (Gasparini, unpublished data)

hēt w-hēt aṣbaḥ-k aḥayr saḥīr-e t-ōk
 2SG.M and-2SG.M wake_up.PRF-2SG.M better brand.PTCP-PL.M ACC-2SG.M
 lā w ham aṣbaḥ-k aḥass hāmā-k? w-maraḏ zēd
 NEG and if wake_up.PRF-2SG.M worse hear.PRF-2SG.M and-illness huge
 l-ōk nḥā saḥīr-e t-ōk škīl-e t-ōk
 to-2SG.M 1PL brand.PTCP-PL.M ACC-2SG.M scar.PTCP-PL.M ACC-2SG.M

mən a-gab
because_of DEF-infection.

‘In case you wake up feeling better / (we) do not brand you but in case you wake up feeling worse / do you understand? And you are seriously ill / we brand you and scar you because of the infection.’

The first clause *hēt w-hēt ašbaḥ-k aḥayr* is an asyndetical circumstantial qualifier functioning as a predictive conditional clause. It contrasts with *w ham ašbaḥ-k aḥass*, in which the conjunction *ham* introduces a counterfactual conditional clause.

In Omani Mehri conditional clauses are commonly introduced through conjunction of pronouns (Watson et al. forthcoming: 211). This structure is unattested in Yemeni Mehri:

- (5) Mehri (Watson et al. forthcoming: 211)
 sēh wa-sēh t-ḥam-ah lā ḥib-sa yi-ḳal-am
 3SG.F and-3SG.F 3SG.F-want.IMPF-3SG.M NEG parents-3SG.F 3M-let.IMPF-PL
 t-ēs ta-ghōm š-ih lā
 ACC-3SG.F 3SG.F-go.SBJV with-3SG.M NEG
 ‘If she doesn’t want him, her parents won’t let her go with him.’

These uses closely resemble those of Gulf Arabic, where circumstantial qualifiers are widely attested to codify predictive and factual conditional and consecutive clauses.

3.4 Lexicon

In the case of the MSAL, it can often be difficult to clearly set apart the effects of Arabisation from those of modernisation and lifestyle changes (which is not surprising, since the two phenomena are interrelated). According to what the speakers themselves report,

it was only since the introduction of formal education, and the awareness of [Modern Standard Arabic] via the media, that Arabic became the second language for many of the MSAL speakers in Dhofar, and, in the case of younger speakers, often to the detriment of their proficiency in their MSAL variety (Davey 2016: 11).

As a consequence, phenomena of borrowing (such as code-switching and loan-words) are particularly common, especially in those varieties (and in the idiolects

of individuals) that are more exposed to Arabic. The following is a good example of code-switching in Baḥari (note that the speaker in question tended to employ Janaybi forms more than other informants):

- (6) Baḥari (Gasparini, unpublished data)
- mēt məssəlīm nə-šāhəd l-ök
 die.PRF.3SG.M muslim 1PL-say_šahada.IMPf for-2SG.M
 w-y-sabbah-uw w-y-kabbər-uw w-y-hälul-uw
 and-3M-pray.IMPf-PL and-3M-pray-PL and-3M-praise_allah-PL
 ‘(If) a Muslim dies / we say the *šahada* for you / and they pray and say
 ‘*allāhu ʔakbar*’ and praise Allah.’

In (6) the speaker makes use of several Arabic verbs related to the semantic field of religious practices, which are not lexically encoded in Baḥari. This might indirectly show the introduction of new ritual practices at a certain point of the history of the tribe. Note that C₂-geminate stems such as *ysabbahuw* and *ykabbəruw* represent verbal patterns not attested in MSAL morphology, and are therefore easily identifiable as loans.

Morris (2017: 15) makes the important remark that lexical erosion is directly connected with the loss of importance of a language in the eyes of its speakers. She gives the example of the Baḥari word for ‘home, living quarters’, for which speakers nowadays frequently resort to some version of Arabic *bayt*, while the many possible original synonyms are falling into disuse. Many of these (*kəḍöt*, *mōhen*, *maššar*, *mōḥayf*, and *ḥader*) are connected to traditional ways of living which have all but disappeared in the course of the last 40–50 years, so that speakers probably judge them inadequate to refer to modern built houses.

3.4.1 Numerals

Watson (2012: 3) reports that “[w]hile Mehri cardinal numbers are typically used for both lower and higher cardinals in Mehreyet, Mahriyöt speakers, in common with speakers of Western Yemeni Mehri, almost invariably use Arabic numbers for cardinals above 10.” This type of lexical substitution connected to numerals higher than ten is also mentioned by Lonnet (2011) and Simeone-Senelle (2011: 1088), who states that “[n]owadays the MSAL number system above 10 is only known and used by elderly Bedouin speakers.” Watson & Al-Mahri (2017: 90) note that it is mostly younger generations (especially in urban settings) who have lost the ability to count beyond ten. Interestingly, they point out that telephone numbers are given exclusively in Arabic, “possibly due to the lack of a single-word MSAL equivalent to Arabic *ṣufr* ‘zero’.”

3.4.2 Spatial reference terms

According to Watson & Al-Mahri (2017: 91) the MSAL employ topographically variable absolute spatial reference terms. In other words, these terms can differ depending on the language employed, the moment of the utterance and the position of the speaker in relation to absolute points of reference. For instance, in and around the city of Salalah in Dhofar, both in the mountains and on the coastal plain, the equivalents of the words for ‘sea’ and ‘desert’ are used to indicate south and north, respectively, in both Mehri (*rawram* and *nagd*) and Ṣ̌ehri (*ramnam* and *fagir*). This is because the sea lies to the south and the desert lies to the north (beyond the mountains). In other parts of the coastal plain, however, the word for ‘mountains’ (*ṣ̌her*) is used to indicate north instead. Another common way to describe south and north is to refer to the direction in which the water flows, with the result that the same word that means ‘south’ on the sea-side of the mountains can be used to indicate ‘north’ on the desert-side. However, all these rather complex sets of terms are being rapidly replaced, particularly in the speech of the younger generations and among urban populations, with the Arabic words for south and north (*ġanūb* and *šimāl* respectively).

3.4.3 Colour terms

The MSAL lexically encode four basic colour terms: white, black, red and green (Bulakh 2017: 261–262). For example, in Ṣ̌ehri one can find *lūn* for ‘white’, *ħor* for ‘black’, *ʕf̣ar* for ‘red’ (and warm colours in general, including brown) and *ṣ̌əẓṛor* for ‘green’ (and everything from green to blue). A fifth colour term, *ṣ̌of̣ṛor* ‘yellow’ (Mehri *ṣ̌āf̣ar*), is most probably an adapted borrowing from Arabic already present at the common MSAL level (Bulakh 2017: 271).

A preliminary field inquiry on the subject was conducted by Gasparini in 2017, with 6 young speakers from the city of Salalah and its immediate surroundings, all between 20 and 35 years old and all bilingual in Ṣ̌ehri and Arabic. The results of the tests showed a remarkable degree of idiolectal variation in the colour labeling systems employed by the informants, with different levels of interference from Arabic. Remarkably, when asked to label colours in Ṣ̌ehri from a printed basic colour wheel, which was shown to them during interviews, all the speakers used the Arabic word for ‘blue’, *azraq*, which seems to have replaced *ṣ̌əẓṛor* (traditionally used for both blue and green, but now confined to the latter). Two speakers also used *aħdar* for ‘green’, claiming that they could not recall the Ṣ̌ehri term. In addition, only one speaker used *ʕf̣ar* for ‘brown’, Arabic *bunnī* being preferred by the other interviewees. The three basic colours ‘white’, ‘black’

and ‘red’, however, were regularly referred to using the Ş̣ehri forms by all speakers. Summing this up, it would seem that the Ş̣ehri colour system (at least in urban environments, but see below) is undergoing a radical process of restructuring. The three typologically fundamental colour terms are retained in most contexts, and a distinction between blue and green is being introduced through reduction of the original semantic spectrum of *śəzrər*, adoption of the Arabic word for blue, and subsequent replacement of *śəzrər* with *aḥḍar* (which indicates only green in Arabic). Further distinctions are either being replaced with the corresponding Arabic terms, or introduced if not part of the original semantic inventory of the language.

On this matter, Watson & Al-Mahri (2017: 90) argue that colour terms (together with numbers) are often among the first lexical items to be lost in contexts of linguistic endangerment, and that this is precisely the case with the MSAL. They write that even children in rural communities are now employing Arabic terms to refer to the different breeds of cattle (which traditionally used to be referred to by use of the three basic colour terms ‘white’, ‘black’ and ‘red’). This is probably a result of the fact that even in villages younger generations are no longer involved in cattle herding. Examples include *aḥmar* ‘bay’ in place of Mehri *ōfar* or Ş̣ehri *fofer*, *aswad* ‘black’ in place of Mehri *ḥōwar*, and *abyaḏ* ‘white’ in place of Mehri *ūbōn*.

3.4.4 Other word classes

Watson & Al-Mahri (2017: 90) note that, since the introduction of a public school system in Arabic in the 1970s, a number of common lexical items and expressions in Mehri and Ş̣ehri have been replaced by the corresponding Arabic ones. Lonnet (2011) also remarks that borrowings from Arabic are particularly common among particles and function words. Examples include *nafs aš-ši* ‘the same thing’ for Ş̣ehri *gens*, Mehri *gans*; *lākin* ‘but’ in place of Ş̣ehri *du^hn* and *min du^hn*, Mehri *lahinnah*; *yaʕnī* ‘that is to say’ and *ʕabārah* in place of Ş̣ehri *yaḥīn*, Mehri *(y)ahah*; *tamām* ‘fine’ in place of Ş̣ehri *ḥayšōf* and Mehri *hīs taww* ~ *histaww*; Mehri and Ḥarsusi vocative *yā* ‘oh’ in place of MSAL *ʔā-*; Ş̣ehri *bdan*, Mehri *ʔabdan* ‘never, not at all’, against Mehri and Ḥarsusi *bəhawʔ*, Ş̣ehri *bhoʔ*. Consider also the case of Arabic *bəss* ‘only’, already mentioned in §3.2.1. In Mehri as in Baḥari, this particle appears now to be interchangeable with its equivalent *ār*, as example (7) shows:

- (7) Mehri (Sima 2009: 328, cited in Watson 2012: 371; transcription adapted)
 bass ta-ʔʕam-h kād aḥah ār ʔʕām ḏ-maḥḥ
 only 2SG.M-taste.IMPF-3SG.M INT fine only taste of-clarified_butter
 ‘Just taste it, like it is just the taste of clarified butter.’

As is predictable, also in this field Baḥari is the language most affected by Arabic: besides those already cited, we might add the expressions *zēn* ‘well’, (*a*)*barr* ‘outside’ (also in Mehri, as opposed to Soqotri *ter*), *ḥalaṣ* ‘and this is it’ (used to end a narrative). Finally, Watson (2012: 3) remarks how “Mahriyōt also exhibits structures unattested in Mehreyyet such as ‘What X!’ phrases reminiscent of Arabic, e.g. *maṭwalk* ‘How tall you (SG.M) are!’”

4 Conclusions

Throughout this chapter we have repeatedly pointed out how research on the MSAL, and in particular on the effects that contact with Arabic has had on their evolution, is still far from reaching its mature stage. Much remains to be done, in particular, in terms of sheer documentation, especially in the case of the most endangered varieties (Hobyōt, Ḥarsūsi, Baḥari). In addition to this, and although Watson’s (2012) work has greatly contributed to expanding our knowledge in this area, MSAL syntax remains a strongly neglected field of inquiry. Finally, our knowledge of the history of the MSAL prior to the twentieth century (and therefore the history of their contact with Arabic) is extremely poor.

It must also be remarked that, although the most widely spoken among the MSAL are undoubtedly better documented, very little is known about the effects that urbanisation has had on their speech communities in recent years. In particular, anecdotal evidence suggests that the varieties of Šehri and Soqotri spoken in Salalah and Ḥadibo are undergoing rapid change under the influence of Arabic (both the standard variety of the language, which children learn in school, and the dialects). Fieldwork conducted in the two abovementioned urban centres could provide extremely valuable information concerning the effects of contact between Arabic and Modern South Arabian.

Despite the far-from-complete state of research in this field, what we currently know is sufficient to say that contact has had a strong impact on the MSAL. Though this is more evident in the area of lexicon, where borrowings are legion, phonetics and phonology have also been affected (though to a different extent from one language to another). Morphology and syntax, on the contrary, appear to be more resistant to contact-induced change, though in the most endangered varieties one can notice a partial disruption of the original pronominal system and verbal paradigm, and though the seemingly high degree of resistance to external influence shown by MSAL syntax could actually be due to our limited knowledge of the subject.

One last note is due concerning another heavily neglected topic, namely the effects that contact with the MSAL have had on spoken Arabic. Though we have

not addressed the question in the course of this paper, evidence drawn from the existing literature (Simeone-Senelle 2002) suggests that this influence, too, is not completely absent, and that further research in this direction could produce interesting results.

Further reading

- Morris (2017) can be thought of as a general introduction to contact between MSAL and Arabic.
- Watson & Al-Mahri (2017) offer an intriguing account of how language change, contact with Arabic and changes to the traditional environment are all deeply interrelated.
- Lonnet (2011) – although limited in scope and extension due to its nature as an encyclopedic entry – offers interesting highlights on the effects of contact on the MSAL.

Abbreviations

1, 2, 3	1st, 2nd, 3rd person	NEG	negative
ACC	accusative	MSAL	Modern South Arabian languages
BCE	before Common Era		
CS	construct state	PTCP	participle
F	feminine	PRO	pronoun
IMP	imperative	PRF	perfect (suffix conjugation)
IMPF	imperfect (prefix conjugation)	PL	plural
		SG	singular
INT	intensifier	VOC	vocative
M	masculine		

References

- Beeston, Alfred F. L. 1981. Languages of pre-Islamic Arabia. *Arabica* 28(2/3). 178–186.
- Bulakh, Maria. 2017. Color terms of the Modern South Arabian languages: A diachronic approach. In Leonid Kogan, Natalia Koslova, Sergey Loesov & Sergey Tishchenko (eds.), *Babel und Bibel 1: Annual of Ancient Near Eastern, Old Testament and Semitic studies*, 269–282. Winona Lake, IN: Eisenbrauns.
- Davey, Richard J. 2016. *Coastal Dhofari Arabic: A sketch grammar*. Leiden: Brill.

- Dufour, Julien. 2016. *Recherches sur le verbe sudarabique moderne*. Paris: Ecole pratique des hautes études. (Habilitation dissertation).
- Edzard, Lutz. 2017. On the role of Modern South Arabian within a comparative Semitic lexicographical project. *Brill's Journal of Afroasiatic Languages and Linguistics* 9(1–2). 119–138.
- Gasparini, Fabio. 2017. Phonetics of emphatics in Bathari. In Simone Bettega & Fabio Gasparini (eds.), *Linguistic studies in the Arabian Gulf*, 69–85. Turin: Dipartimento di Lingue e Letterature Straniere e Culture Moderne, Università di Torino.
- Gasparini, Fabio. 2018. *The Baḥari language of Oman: Towards a descriptive grammar*. Naples: Università degli studi di Napoli “L’Orientale”. (Doctoral dissertation).
- Hoyland, Robert G. 2001. *Arabia and the Arabs: From the Bronze Age to the coming of Islam*. London: Routledge.
- Johnstone, Thomas M. 1975. The Modern South Arabian languages. *Afro-Asiatic Linguistics* 1. 93–121.
- Kogan, Leonid. 2015. *Genealogical classification of Semitic: The lexical isoglosses*. Berlin: Walter de Gruyter.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Lonnet, Antoine. 2011. Modern South Arabian. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.
- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bowerman & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Morris, Miranda J. 2017. Some thoughts on studying the endangered Modern South Arabian languages. *Brill's Journal of Afroasiatic Languages and Linguistics* 9(1–2). 9–32.
- Naumkin, Vitaly & Victor Porkhomovsky. 1981. *Očerki po etnolingvistike sokotry*. Moscow: Nauka.
- Persson, Maria. 2009. Circumstantial qualifiers in Gulf Arabic dialects. In Bo Isaksson, Heléne Kammensjö & Maria Persson (eds.), *Circumstantial qualifiers in Semitic: The case of Arabic and Hebrew*, 206–286. Wiesbaden: Harrassowitz.
- Porkhomovsky, Victor. 1997. Modern South Arabian languages from a Semitic and Hamito-Semitic perspective. In Janet Starkey (ed.), *Proceedings of the Seminar for Arabian Studies vol. 27*, 219–223. Oxford: Archaeopress.
- Retsö, Jan. 2000. *Kaškaša, t-passives and the ancient dialects in Arabia*. *Oriente Moderno* 19(1). 111–118.

- Ridouane, Rachid & Cédric Gendrot. 2017. On ejective fricatives in Omani Mehri. *Brill's Journal of Afroasiatic Languages and Linguistics* 9(1–2). 139–159.
- Robin, Christian-Julien. 1991. La pénétration des arabes nomades au Yémen. *Revue des mondes musulmans et de la Méditerranée* 61(1). 71–88.
- Rubin, Aaron D. 2014. *The ġibbali (Shahri) language of Oman: Grammar and texts*. Leiden: Brill.
- Rubin, Aaron D. 2017. A re-analysis of Mehri grammar. Paper presented at the Cinquièmes journées d'études sur les langues sudarabiques modernes, September 2017, Paris.
- Sima, Alexander. 2009. *Mehri-Texte aus der jemenitischen Sharqīyah: Transkribiert unter Mitwirkung von Askari Hugayran Saad*. Wiesbaden: Harrassowitz.
- Simeone-Senelle, Marie-Claude. 2002. L'arabe parlé dans le Mahra (Yémen). In Otto Jastrow, Werner Arnold & Hartmut Bobzin (eds.), *"Sprich doch mit deinen Knechten aramäisch, wir verstehen es!": 60 Beiträge zur Semitistik: Festschrift für Otto Jastrow zum 60. Geburtstag*, 669–684. Wiesbaden: Harrassowitz.
- 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.
- 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.
- Versteegh, Kees. 2014. *The Arabic language*. 2nd edn. Edinburgh: Edinburgh University Press.
- Watson, Janet C. E. 2011. South Arabian and Yemeni dialects. *Salford Working Papers in Linguistics and Applied Linguistics* 1. 27–40.
- Watson, Janet C. E. 2012. *The structure of Mehri*. Wiesbaden: Harrassowitz.
- Watson, Janet C. E. 2018. South Arabian and Arabic dialects. In Clive Holes (ed.), *Arabic historical dialectology: Linguistic and sociolinguistic approaches*, 316–334. Oxford: Oxford University Press.
- Watson, Janet C. E. & Abdullah Musallam Al-Mahri. 2017. Language and nature in Dhofar. In Simone Bettega & Fabio Gasparini (eds.), *Linguistic studies in the Arabian Gulf*, 87–103. Turin: Dipartimento di Lingue e Letterature Straniere e Culture Moderne, Università di Torino.
- Watson, Janet C. E., Abdullah Musallam Al-Mahri & Ali Al-Mahri. Forthcoming. *Taghamk afyat: A course in Mehri*. Wiesbaden: Harrassowitz.

- Watson, Janet C. E. & Alex Bellem. 2010. A detective story: Emphatics in Mehri. In Janet Starkey (ed.), *Proceedings of the Seminar for Arabian Studies vol. 40*, 345–355. Oxford: Archaeopress.
- Watson, Janet C. E. & Alex Bellem. 2011. Glottalisation and neutralisation in Yemeni Arabic and Mehri: An acoustic study. In Majeed Hassan Zeki & Barry Heselwood (eds.), *Instrumental studies in Arabic phonetics*, 235–256. Amsterdam: John Benjamins.
- Watson, Janet C. E. & Barry Heselwood. 2016. Phonation and glottal states in Modern South Arabian and San’ani Arabic. In Youssef A. Haddad & Eric Potsdam (eds.), *Perspectives on Arabic linguistics XXVIII: Papers from the Annual Symposium on Arabic Linguistics, Gainesville, Florida, 2014*, 3–36. Amsterdam: John Benjamins.
- Watson, Janet C. E., Bonnie Glover Stalls, Khlaïd Al-Razihi & Shelagh Weir. 2006. The language of Jabal Rāziḥ: Arabic or something else? In Rob Carter & St John Simpson (eds.), *Proceedings of the Seminar for Arabian Studies vol. 36*, 35–41. Oxford: Archaeopress.
- Wellsted, James R. 1837. Narrative of a journey into the interior of Omán, in 1835. *Journal of the Royal Geographical Society of London* 7. 102–113.
- Winford, Donald. 2005. Contact-induced change: Classification and processes. *Diachronica* 22(2). 373–427.
- Zammit, Martin R. 2011. South Arabian loanwords. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

