

## 2019 UNESCO International Year of Indigenous Languages - The Modern South Arabian Languages

By Miranda Morris

The Modern South Arabian Languages (MSAL), Mehri, Soqoṭri, Śherēt (also known as Jibbāli or Shahri), Ḥarsūsi, Hobyōt and Baḥari, are unwritten Semitic languages spoken by minority populations in south-east Yemen, southern Oman and the fringes of southern and eastern Saudi Arabia. The name 'Modern South Arabian' is somewhat confusing, as these unwritten languages are neither 'modern' nor comprehensible to an Arabic speaker. They are called 'Modern South Arabian' languages (henceforth referred to as MSAL) to differentiate them from 'Old South Arabian', which refers to the four related languages which were written in the Ancient South Arabian script and are now extinct. The MSAL belong to the South Semitic branch of the Semitic language family, which also includes Ethiopian Semitic. This is distinguished from the Central Semitic branch, which includes the more widely known Arabic, Aramaic, and Hebrew. The MSAL are believed to be the remnants of a pre-Arabic substratum that once stretched over the whole of southern Arabia, and across the Red Sea, into the highlands and littoral of East Africa.

Due to rapid economic and socio-political change in recent decades, the spread of Arabic among MSAL speakers has resulted in these languages increasingly falling into disuse. In both Oman and Yemen today, the official language is Arabic: of education, government, the media and commerce. Being in competition with another more widely-spoken and literate language is a common problem for purely oral languages. However, in the case of the MSAL, the official language in question is Arabic. As this is also the language of the Qurʾān, and one which Muslims (nearly a quarter of the world's population) work hard to learn and understand, it means that these six minority languages are in competition with an extremely high-prestige language.

The six languages of this group are in varying stages of endangerment. Note that with the exception of figures for speakers of Soqoṭri and Baḥari, the figures given below are best estimates, and speaker numbers for all MSAL are rapidly declining:

- i. Mehri is the most widespread language, spoken by people of the Mahra tribes in Oman, Yemen, and parts of southern and eastern Saudi Arabia. The Mahra are estimated to be some 180,000 people, though the actual number of those among them who speak Mehri is difficult to estimate since the language is spoken across three state boundaries, and many Mahra no longer speak Mehri;
- ii. Soqoṭri, spoken exclusively in the islands of the Soqoṭra Archipelago, a World Heritage Site, has some 60,000 speakers;
- iii. Śherēt, spoken by a variety of tribes within the Dhofar region of Oman, has some 30,000 speakers;
- iv. Ḥarsūsi, formerly spoken by members of the Ḥarsūsi tribe across the Jiddat al-Ḥarāsīs in central Oman, has under 1,000 speakers;

- v. Hobyōt, spoken by a variety of tribes on both sides of the Yemeni / Omani border, likewise has under 1,000 speakers;
- vi. Baḥari, spoken by members of the Baḥari tribe who live along the shore opposite the Al-Hallaniyah islands and in the desert plateau above, has less than 30 speakers.

However, today, the majority of speakers of MSAL also speak Arabic.

The areas in which the MSAL are still spoken are the only regions within the Arabian Peninsula to have retained the Semitic languages spoken prior to the spread of Islam and subsequent Arabisation of the Peninsula. In all other communities, Arabic appears to have superseded the original languages. As such, the documentation and description of the MSAL is of crucial importance to understanding the historical development of the Semitic language family as a whole.

These languages are noted for their retention of ancient Semitic phonological and grammatical features that have disappeared from other Semitic languages, suggesting that the MSAL are the oldest extant Semitic languages. Retained linguistic features include:

- a) In the consonantal phonology, a contrast of three plain voiceless sibilants (s-like sounds), known to have existed in Ancient South Arabian, of which one is a lateral sibilant. In Mehri, for example, *šiff* 'to fart silently' with a palato-alveolar sibilant, pronounced as in English 'sh' in 'sheep', contrasts with *šiff* 'to want' with a lateral sibilant, pronounced as in Welsh 'll' in words such as 'llid' 'song'; and *škawn* 'thorns' with a lateral sibilant contrasts with *skawn* 'they m. lived in' with an alveolar sibilant, pronounced as in English 's' in 'soul' (Watson, al-Mahri et al, in press).
- b) In the pronoun system, dual contrasts with singular and plural in all persons. In Mehri, for example, *akay* 'we two' contrasts with *hōh* 'I' and *nḥāh* 'we'. Similarly, *atay* 'you two' contrasts with *hēt* 'you singular (s.)' and *atēm* 'you masculine plural (m.pl.)' and *atēn* 'you feminine plural (f.pl.)' (Rubin 2010, 2018; Watson 2012). Within the Semitic language family, no other purely spoken language retains the dual pronoun, and the long-extinct Ugaritic is apparently the only other Semitic language to have exhibited a first person dual pronoun.

Other linguistic features of interest to the MSAL include:

- c) A highly non-concatenative morphological system. By non-concatenation we mean that many inflections are expressed within the word stem rather than as explicit suffixes or prefixes. Thus, in Mehri while *tḥōm* means 'you m.s./she want(s)', *tḥaym* with a change in the vowel of the verb stem means either 'you f.s. want' or 'you m.pl. want'. In Šḥerēt, the difference between masculine and feminine in some adjectives is indicated by an 'o' vowel in the masculine, and 'i' vowel in the feminine, as in the word for 'green', which is *šəṣror* in the masculine, *šəṣrir* in the feminine, and the word for 'cold', which is *kaṣmun*, in the masculine, and *kaṣmin* in the feminine (cf. Rubin 2014).
- d) A great deal of syncretism, particularly in the verb system (Eades 2014). By 'syncretism', we mean that one particular form can indicate two or more morphological categories. Thus in all the MSAL, the basic form of the perfect verb indicates both third person m.s. and third person f.pl., as in Mehri: *šini*

‘he/they f. saw’ and *assōfar* ‘he/they f. travelled’. In Šherēt this syncretism goes even further, with the basic form of the perfect verb indicating third m.s. and third m.pl. and third f.pl. Within the imperfect verb, certain verbal forms show syncretism for many more morphological categories. Thus, in the indicative of the verb *abōšar* ‘to go around twilight’ in Mehri the form *tābašran* stands for all second persons – you m.s., you f.s., you m.pl., you f.pl., ‘you go around twilight’ – and for third f.s. and f.pl., thus ‘she goes around twilight’ and ‘they f. go around twilight’ (Watson, al-Mahri et al, in press).

- e) The development of consonants not attested in other Semitic languages. These include the emphatic counterpart of š (pronounced as ‘sh’), ṣ̌, in all the MSAL. Thus in Mehri, *fišš* ‘to release air’ with a plain final ‘š’ contrasts with *fišṣ̌* ‘to press hard’ with an emphatic ‘ṣ̌’. Šherēt has developed more consonants than the other MSAL, and contrasts *four* plain voiceless sibilants: in addition to *s*, *ṣ* and *š*, it has a sibilant produced with lip rounding, which is commonly transcribed as *ṣ̣̌*. Thus, *šum* ‘they m.’ contrasts with *šuhum* ‘with them m.’ Šherēt also has a voiced lateral sibilant, which we and others in the field transcribe as *ž*, as in *iž irun* [owners goats] ‘goat owners’.
- f) Finally, when using a possessive pronoun with a noun, the MSAL, in contrast to other Semitic languages, requires the definite article. Thus in Mehri *wōz* ‘goat’ becomes *hōz* ‘the goat’ and ‘my goat’ then takes the pronoun suffix *-i* to give *hōzi* [the-goat-mine] ‘my goat’.

In addition to their importance to Semitic linguistics, the documentation of the MSAL is crucial to producing a permanent record of the fast-disappearing cultural traditions and socio-economic practices of the speakers. These once self-sufficient communities of MSAL speakers lived close to the natural environment for centuries, and of necessity had to rely directly on it, not only for their survival, but also for their cultural and spiritual needs. Language documentation provides an insight into the culture and way of life of speakers, and preserves traditional knowledge built up over generations by language communities, such as methods of animal husbandry, land-use management, sustainable harvesting of the sea, use of the local vegetation, and many other areas of traditional expertise. This information is of interest not only to linguists, but also to historians, anthropologists, sociologists, and specialists in other scientific fields. Such documentation is also of great interest to the language communities themselves: they are keen to maintain a record of their heritage and traditions in a world which is undergoing rapid change. The Baṭāḥira, for instance, have an extensive lexicon not only for the fish they catch, but also for the differing physical conditions they display at different times of the year; and the lexis for butchering and preparing a turtle as food differs totally from that for butchering and preparing domesticated livestock as food in the MSAL. The passing seasons are very specific and marine-orientated for the Baṭāḥira: as is common to all the MSAL, each season is demarcated by the rising and setting of certain ‘stars’, but the Baṭāḥira extend this to include details of winds, the state of the sea and the seasonal availability of shellfish, shark and turtle, as well as fish. On Soqoṭra, the annual development of the flowers, fruit, leaflets, wood and blood-red resin of the commercially important ‘dragon’s blood tree, *Dracaena cinnabari*, one of the 300 or so endemic plants on the island, is described in minute lexical detail, almost matching the complex vocabulary surrounding the date-palm, so central to the life of the islanders.

During a recent Leverhulme-funded project, *Documentation and ethnolinguistic analysis of the Modern South Arabian languages*: RPG-2012-599, which ran from 2013-2016, a large corpus of annotated, transcribed and translated oral data from five of the Modern South Arabian languages, Baḥari, Hobyot, Ḥarsusi, Šherēt and Mehri, was collected by Janet Watson, Miranda Morris and Domenyk Eades. This has been archived with the Endangered Languages Archive (ELAR) at SOAS in London, <https://elar.soas.ac.uk/> (Eades & Morris 2016; Morris 2016a, 2016b; Watson & Morris 2016a, 2016b). To this collection was added the 1970s recordings of MSAL speakers made by Miranda Morris when living and working in Dhofar, and later, further recordings made in the 1980s under the sponsorship of the Oman Diwan of Royal Court. The many hours of recordings of Soqotri made between 1990 and 2010 by Morris when working on Soqotri have not yet been archived, but are planned to be archived at ELAR in the near future.

As is the case in many oral cultures, these languages have an extremely rich poetic tradition. As well as poems and songs, this oral tradition includes prayers, lullabies, work chants, community wisdom enshrined in poetic couplets, messages in code, riddles, and stories centred on a short poem or exchange of poems. This valuable and irreplaceable repository of linguistic and ethnographic material is likely to be eroded even more rapidly than the language itself. A high percentage of the material recorded, much of it already archived at ELAR, is of this nature. For Mehri, the reader should consult Liebhaber's excellent digital exhibit, *When Melodies Gather* (Liebhaber 2018), which contains sound files, video files, transcripts and translations of different poetic types in Mehri.

On the website of the *Documentation and ethnolinguistic analysis of the Modern South Arabian languages* project, <http://ahc.leeds.ac.uk/modern-south-arabian-languages> can be found a map of the areas in which the MSAL are spoken; details of the Arabic-based script devised for the project; a photographic library; videos of speakers using Šherēt and Mehri taken by Janet Watson; sample sound-files; other relevant websites; theses and dissertations on the MSAL by Omani and Yemeni scholars; three children's books in Mehri by Janet Watson and Abdullah al-Mahri, and much else. A full and regularly updated bibliography of the MSA languages is also available here.

For those interested, a sample transcribed and translated text in Mehri is available here:

<https://elar.soas.ac.uk/Collection/MPI976775>

And one in Šherēt here: <https://elar.soas.ac.uk/Collection/MPI972274>

An online article that includes photos and a video and translated sound file is available here: <https://www.al-fanarmedia.org/2019/02/ancient-cousins-of-arabic-survive-in-oman-but-for-how-long/>

As regards recently published work on the MSAL, grammars of Mehri by Janet Watson (Watson 2012) and Aaron Rubin (Rubin 2010, 2018), and of Šherēt by Aaron Rubin (Rubin 2014), have appeared recently, while a teaching grammar of Mehri by Janet Watson, Abdullah al-Mahri and others is in preparation. In 2014 and 2018, volumes I and II of *Corpus of Soqotri Oral Literature* by Naumkin, Kogan, Isa al-Da'rhi, Ahmed al-Da'rhi,

Maysoon al-Da'ri, Cherkashin, Bulakh and Vizirova appeared (Naumkin, Kogan et al 2014, 2018), offering a wide variety of annotated texts. These are presented in both English and Arabic, and are also transcribed phonetically and in an Arabic-based orthography. Collections of transcribed and translated texts in Baḥari and Hobyōt are being prepared for publication by Miranda Morris (Morris, in prep. a, in prep. b), while *The Oral Art of Soqatra*, by Miranda Morris and Tanuf Nuh II-Kishin, a collection of over 2,000 poetical texts, is being finalised for publication (Morris & II-Kishin, in prep.).

The Bibliography on the Resources page of the *Documentation and ethnolinguistic analysis of the Modern South Arabian languages* website can be consulted for the increasing number of publications that are appearing on the MSA languages and the areas where they are spoken.

The continued survival of these languages depends on their ability to hold their own against the Arabic that dominates their societies, and to do so with little or no official support or recognition. And this depends firstly, on the enthusiasm of speakers to use their language amongst themselves and in the home, and secondly, on there being enough people from a language community who are interested both in speaking and in studying their language. School, the media, the internet and the drift towards towns, where Arabic is the lingua franca, have already compromised the former; and if the latter is to be productive, it has to get underway quickly, while there are enough people fluent in their language still living.

The likelihood of introducing the study of the MSAL and basic phonetics into the school curricula of Oman and Yemen might seem extremely slim at present. But we might perhaps hope that in the not-too-distant future ways of transcribing these languages in a modified Arabic script will be seriously discussed in both countries, and a standardised system devised. Because it is certain that if speakers don't start to write their languages, they will certainly lose them altogether.

## References

- Eades, Domenyk 2014. Syncretism in the verbal morphology of the Modern South Arabian Languages. In: Orhan Elmaz & Janet Watson (eds), *Languages of South Arabia (Supplement to the Proceedings of the Seminar for Arabian Studies Volume 44)*. Oxford: Archaeopress. 19–28.
- Eades, Domenyk & Miranda J. Morris 2016. *Documentation of Modern South Arabian: Harsusi*. ID: Harsusi (0314). London: SOAS. Endangered Languages Archive, ELAR.
- Liebhaber, Sam 2018. *When Melodies Gather: Oral art of the Mahra*. Stanford: Stanford University Press.
- Morris, Miranda 2016a. *The Documentation of Modern South Arabian: Bathari*. ID: Bathari (0364). London: SOAS, Endangered Languages Archive, ELAR.
- Morris, Miranda 2016b. *The Documentation of Modern South Arabian: Hobyot*. ID: Hobyot (0309). London: SOAS, Endangered Languages Archive, ELAR.
- Morris, Miranda (in prep. a) (provisional title) Baḥari Texts. Wiesbaden: Harrassowitz.
- Morris, Miranda (in prep. b) (provisional title) Hobyōt Texts. Wiesbaden: Harrassowitz.
- Morris, Miranda & Tanuf Nuh II-Kishin (in prep.) *The Oral Art of Soqatra*. Leiden: Brill.
- Naumkin, Vitaly, Leonid Kogan, Dmitry Cherkashin, Maria Bulakh, Ekaterina Vizirova, 'Īsa Gum'ān al-Daarhi & Aḥmad 'Īsa al-Daarhi 2014. *Corpus of Soqotri Oral Literature*. Volume 1. Leiden: Brill.
- Naumkin, Vitaly, Leonid Kogan, Dmitry Cherkashin, Maria Bulakh, Ekaterina Vizirova, 'Īsa Gum'ān al-Daarhi & Aḥmad 'Īsa al-Daarhi 2018. *Corpus of Soqotri Oral Literature*. Volume 2. Leiden: Brill.
- Rubin, Aaron 2010. *The Mehri Language of Oman*. Leiden: Brill.

Rubin, Aaron 2014. *The Jibbali (Shahri) Language of Oman*. Leiden: Brill.  
Rubin, Aaron 2018. *Omani Mehri: A New Grammar with Texts*. Leiden: Brill.  
Watson, Janet C.E. 2012. *The Structure of Mehri*. Wiesbaden: Harrassowitz.  
Watson, Janet C.E. & Miranda Morris 2016a. *Documentation of the Modern South Arabian Languages: Mehri*.  
ID: Mehri (0307). London: SOAS. Endangered Languages Archive, ELAR.  
Watson, Janet C.E. & Miranda Morris 2016b. *Documentation of the Modern South Arabian Languages: Shehret*.  
ID: Shehret (0308). London: SOAS. Endangered Languages Archive, ELAR.  
Watson, Janet C.E., Abdullah al-Mahri, Ali al-Mahri & Bxayta Musallam Kḥōr al-Mahri (in press) *Taghamk Āfyat: A Course in the Mehri of Dhofar*. Wiesbaden: Harrassowitz.