

Emprunts intra-sémitiques: l'exemple des emprunts arabes en sudarabique moderne¹

Antoine Lonnet – C.N.R.S., Laboratoire d'Études Sémitiques Anciennes, Collège de France, Paris

[The Modern South Arabian (MSA) languages are the *stratum* of Arabic in southern Arabia, as well as that of Epigraphic South Arabian. The MSA family was shaped by two events: the west-east split—a response to the settling of the ancient Kingdoms—and the migration of a group (belonging to the eastern branch) to Socotra Island. *Multilingualism* is, nowadays, standard practice among the MSA speakers, who nearly all speak also Arabic and, except in Socotra, one or two other MSA languages as well. The important issue of *intra-Semitic borrowings* can be illustrated with such an example. 'Classical Arabic' is the source of the Arabic borrowings in MSA when Islamic culture is concerned, but more abundant a source are the Yemenite and Omani Arabic *dialects*. Recent loans originate in 'International standard Arabic'. The oldest ones can be situated *before* the arrival of the Sabæans, when the Peninsula was shared between North Arabian and South Arabian dialects. Since the end of the South Arabian kingdoms, MSA has been in contact only with Arabic on dry land. However, the nautical traffic of men and of their material and cultural goods—and among them languages, vocabulary—is very heavy in the north-east of the Indian Ocean, and both Arabic and MSA have been trading on that 'market' for ages. Hence, it is not easy to determine the *source* and the circuitous *route* of a given loanword. The *kinship proximity* of MSA with Arabic is an important characteristic: it appears in their extensive common lexical/radical stock, and their analogous morphology and partially identical phonology, which highly facilitates the integration of loanwords. The languages can easily be graded according to that 'similarity' to Arabic: Harsusi, Mehri, Bathari, Hobyot, Jibbali, Socotri (in the order of decreasing similarity). The present paper aims at showing how the result of such a seemingly disordered blending can be properly analysed.]

Les langues chamito-sémitiques occupent depuis au moins cinq millénaires une vaste aire afro-asiatique (le nord de l'Afrique et l'est de l'Asie). Là, elles ont certes été en concurrence avec des langues appartenant à d'autres familles, mais, plus généralement et plus remarquablement, elles ont été affectées par les interférences avec d'autres langues de la famille.

Une typologie (théorique) de ces interférences pourrait prendre en considération la mesure de la proximité linguistique (structurale, phonétique, lexicale, typologique même) des parlers en interaction. Elle devrait aussi tenir compte de la nature socio-culturelle des contacts, de leur durée.

1. Une présentation en anglais de ces résultats se trouve dans l'*Encyclopedia of Arabic Language and Linguistics* (EALL), Leiden, Brill, parution prévue pour 2006. Abréviations: SAM= sudarabique moderne, M= mehri, S= socotri, J= jibbali, H= hobyot, Ḥ= harsusi, ar.= arabe.

Les langues « empruntent » du matériel lexical lorsque les circonstances l'imposent ou lorsque la société le demande. Ce passage de frontière se fait, schématiquement, à l'occasion d'interactions sociales entre des groupes de langue différente, et concerne en premier lieu (mais pas exclusivement) des locuteurs bilingues, et ce, dans des conditions où les parlars en présence sont rarement canoniques.

L'exemple que je propose d'examiner se situe au sud de l'Arabie, et concerne des langues – l'arabe et le sudarabique moderne – de la famille sémitique, dont on peut dire qu'elles sont assez proches au sein de cette famille, mais parfaitement distinctes contrairement à ce que pourrait laisser croire la dénomination *sudarabique* ; les sémitisants savent bien qu'il n'y a pas la moindre intercompréhension entre arabe et sudarabique.

Le sudarabique moderne² regroupe plusieurs langues entre lesquelles il n'y a pas non plus d'intercompréhension. Ces langues elles-mêmes manifestent une dialectalisation aussi développée que celle de l'arabe (de l'arabe en Arabie du sud : donc cette dialectalisation est *très* développée). Cette réalité dialectale doit être prise en compte pour se prononcer valablement sur les emprunts : sans connaissance précise des variétés de la langue donneuse³ et de celles de la langue receveuse, on perd de vue la réalité sociale, historique, géographique ... J'ai essayé ici, sans rentrer dans trop de détails, de rendre compte des principales variétés.

Je n'examinerai que les emprunts dans un sens ; l'influence du sudarabique moderne sur l'arabe reste à évaluer : les travaux sur l'arabe ancien ont une tendance excessive à considérer les relations avec le sudarabique épigraphique, et méconnaissent le sudarabique moderne (désormais : 'SAM'). Or, dans sa pénétration vers le sud, c'est avec cette langue que l'arabe est entré en rivalité, et donc en interférence.

Les langues sudarabiques modernes constituent un substrat pour l'arabe au sud de la Péninsule, comme elles l'ont fait pour le sudarabique épigraphique. Le rapport avec l'arabe n'est pas le même selon les époques et, en particulier, le contact épisodique avec des tribus arabes rivales ne produit pas les mêmes effets linguistiques que l'inexorable croissance du poids culturel de la langue de l'Islam.

La famille sudarabique moderne a été façonnée par deux faits historiques: la séparation ouest-est –qui répond à l'apparition des royaumes sudarabiques antiques– suivie, quelques siècles plus tard, par la migration vers l'île de Socotra d'un groupe de la branche orientale. Le SAM occidental, *i. e.* le **mehri** [mɸhri],⁴ est marqué par le contact ininterrompu avec les langues des communautés au pouvoir depuis plus de 3000 ans, langues sudarabiques anciennes, puis langue arabe. Le SAM oriental, continental (**jibbali**) [jibbāli], et insulaire (**socotri**) [sɔqɔtri] s'est développé à l'écart de ces influences, isolé par les conditions géographiques. Ce tableau se complique il y a quelques siècles (six siècles ?), lorsque des tribus mehri venues de l'ouest s'emparent de la zone orientale de l'aire sudarabophone continentale, adoptant la langue des populations assujetties (le jibbali) sauf dans le cas de deux minorités qui ont adopté la langue des vainqueurs : celle qui parle le **harsusi** [ħarsūsi] et celle qui parle le **bathari** [bāṭhari], les deux langues étant des dialectes du mehri. De ce rapide aperçu, on tire les conclusions schématiques suivantes : au Yémen, le mehri le plus occidental est affecté assez profondément par le contact avec l'arabe, le mehri du Dhofar l'est nettement moins. Le jibbali des tribus (mehri) dominantes (les ʕḥklóʔ ou Qarā) est peu marqué par l'arabe, mais un peu plus que celui des groupes non-tribaux asservis (les Šḥeróʔ ou Gəblóʔ). Les variétés de socotri les plus isolées, dans les hautes montagnes, ont presque complètement échappé à

2. Cf. Johnstone 1977, 1981, 1987, Leslau 1938, Lonnet 1994. Les exemples sont cités d'après les trois dictionnaires de Johnstone et d'après le corpus que j'ai constitué au cours de mes enquêtes de terrain.

3. Pour donner un exemple, bien des travaux sur les emprunts à l'arabe dans les langues d'Afrique, d'Europe, d'Asie, omettent la dialectologie arabe : la validité des résultats en est souvent compromise.

4. Les noms vernaculaires entre crochets sont en arabe et non en SAM.

l'influence de l'arabe,⁵ ce qui n'a pas été possible pour les dialectes côtiers ; de plus, il est probable que des fractions tribales arabes venues des montagnes du Yémen aient migré vers Socotra, où elles ont acquis le socotri, mais sans perdre l'arabe.

Les groupes sudarabophones peuvent avoir individuellement des relations étroites avec des groupes arabophones : c'est le cas des Harāsīs, dont la langue ḥarsūsi contient de très nombreux emprunts à l'arabe et manifeste des symptômes d'extinction. C'est aussi le cas des habitants de 'Abd-el-Kūri, île voisine de Socotra, qui entretiennent des relations privilégiées avec certains villages de Socotra, mais surtout avec deux villages (arabophones) du Hadramout, au point que leur parler a connu une évolution très particulière sous l'influence de l'arabe. Il reste à dire un mot du bathari, qui s'éteint autant sous la pression du jibbali que sous celle de l'arabe, et surtout du **hobyot** [həwbyūt],⁶ petite langue très bien préservée, à la frontière du Dhofar et du Mahra, profondément marquée par le mehri et le jibbali, et guère par l'arabe.

À l'époque contemporaine, une nouvelle vague d'emprunts arabes est provoquée par la "modernité" : d'une part, les États (Yémen, Oman) affirment leur autorité, cherchent à fixer les nomades, soutiennent un développement économique centralisé –donc arabisé–, autant d'obstacles à la survie des identités culturelles locales ; d'autre part les moyens modernes de diffusion des langues (ondes hertziennes, bandes magnétiques, ...) sont presque entièrement au service de l'arabe.

Le plurilinguisme est la règle chez les sudarabophones, qui pratiquent presque toujours l'arabe et souvent (sauf à Socotra) une ou deux autres langue SAM.

Les emprunts arabes en SAM proviennent de l'arabe "classique" lorsqu'il s'agit de faits culturels islamiques, mais aussi, et beaucoup plus abondamment, de divers dialectes du Yémen et de l'Oman. On peut attribuer les emprunts les plus récents à un arabe moderne international. Les plus anciens emprunts sont à situer à une époque qui précède l'arrivée des Sabéens, lorsque des dialectes nord-arabiques et sud-arabiques se partageaient la péninsule.

Depuis la fin des royaumes sudarabiques antiques, et en faisant abstraction de l'épisode colonial britannique –qui ne les a guère directement touchées–, les langues sudarabiques ne sont plus en contact qu'avec l'arabe sur la terre ferme. Cependant, la circulation maritime des hommes et de leurs biens matériels et culturels –au premier rang desquels les langues, les lexiques– est très intense dans le nord-est de l'océan Indien, et l'arabe et le sudarabique moderne y ont joué leur rôle, donnant et recevant à travers les siècles. Il s'ensuit qu'il n'est pas toujours facile de déterminer les voies d'emprunt d'un mot donné.

Une importante caractéristique du sudarabique moderne est sa parenté avec l'arabe, qui se manifeste par un fonds lexical commun considérable, une morphologie du même type et une phonétique partiellement partagée, ce qui facilite hautement l'intégration des emprunts. Nous verrons ci-après que les différences morphologiques et phonétiques permettent d'identifier les emprunts les plus récents. On peut classer les langues par *ordre de ressemblance décroissante* avec l'arabe, selon des critères intuitifs que confirment d'une part l'analyse des difficultés d'acquisition pour les apprenants, d'autre part l'analyse structurale :

HARSUSI, MEHRI, BATHARI, HOBYOT, JIBBALI, SOCOTRI.

L'invasion du lexique harsusi par l'arabe se voit très bien à la lecture du *Ḥarsūsi Lexicon* (Johnstone, 1977). Il y a une abondance de lexèmes qui se présentent en harsusi avec une racine arabe tandis que toutes les autres langues SAM ont un équivalent dont la racine est proprement SAM. Ex. : "aisselle", H

5. Les dernières nouvelles de Socotra laissent craindre que la fin de ces dialectes préservés s'approche.

6. Ce nom vernaculaire est celui par lequel la langue se désigne elle-même.

ʔəbāṭ (< ar. clas. *ʔibṭ*, ar. d'Oman *ʔibāṭ*) ; les autres langues SAM ont, avec le même sens, des noms de racine SAM *šxw* : J *šxot*, S *šḥoh*, M *xōt*.⁷

Notre présentation n'est qu'une esquisse, car elle ne fait qu'effleurer la très riche diversité dialectale sudarabique. Elle ne prend pas non plus en compte les faits socio-linguistiques cruciaux qu'on regroupe sous l'étiquette de "code-switching" : les emprunts examinés ici se présentent dans des discours aussi monolingues que possible.

Avant toute chose, rappelons que toute partie du discours peut être empruntée. Par exemple, dans le système pronominal : le SAM se caractérise par un pronom suffixe de 1^{ère} personne sg. de forme *-i* en tous contextes (même après un verbe), mais il y a, surtout en harsusi, des occurrences de la forme *-ni* après un verbe. L'influence de l'arabe est évidente, mais la propagation de ce *-ni* à la plupart des prépositions et particules révèle une situation structurale particulière qui n'est pas celle de l'arabe (ni local, ni standard) ni celle du SAM.

Le cas le plus simple est celui de l'emprunt lexical. Le nom de la 'vache' illustre la variété des situations : la racine arabe *bqr* et la racine SAM (chamito-sémitique) *Iʔy / Ihy* sont en rivalité :

1. L'emprunt *bəḳərēt*, pluriel *bəḳār*, complètement intégré (il a même un diminutif : *bəḳərēnōt*), se trouve en harsusi et en mehri yéménite occidental.
2. La racine SAM apparaît
 - A. *seulement* pour le collectif *Iháyṭən*, dans *certain*s dialectes mehri,
 - B. mais pour *tous* les nombres :
 - a. en mehri de Jādib (Yémen) : *Ieh* pluriel et collectif *Iháyṭən*
 - b. et dans toutes les autres langues SAM, lesquelles ignorent la racine arabe.

La racine *Iʔy* se retrouvant au nord du Yémen, loin de l'aire SAM, la possibilité d'un emprunt très ancien (dans un sens ou dans l'autre) doit être envisagée.

Il n'est pas rare que deux **mots** d'origine différente soient en concurrence pour exprimer le même sens. Dans les paires d'exemples suivantes, en hobyot, le premier terme est SAM et le second emprunté à l'arabe :

"cœur"	<i>ʔəlbēb</i>	&	<i>ḳalb</i> ,
"ciel"	<i>hétəm</i>	&	<i>sēmēʔ</i> ,
"cracher"	<i>fosg</i>	&	<i>tfōl</i>
"robe"	<i>xəllōḳ</i>	&	<i>ksuwét</i> .

Le mot emprunté peut prendre un sens s'écartant plus ou moins du sens d'origine : ainsi, en mehri de Qishn, dans *ḥarmēt*, pl. *ʔāgzōn* "femme", on reconnaît deux mots d'origine arabe fonctionnant exactement comme l'arabe *marʔa* et son pl. *nīsāʔ*, plutôt que comme les mots arabes *ḥurma* "femme mariée" ou *ʔajūz* "vieille femme" (qui est sans doute le sg. dont *ʔāgzōn* est le pl. (à suffixe **-ān*).

L'emprunt arabe peut comporter une racine présente en SAM avec un sens différent : ainsi J *žerbét* "coup" est un emprunt arabe ; la racine SAM *žrb* signifie "éprouver une douleur" et non "frapper" (qui s'exprime par les racines *krš*, *rgt*, *sbṭ*, *sḳḥ* etc.)

7. Le harsusi présente en outre un curieux doublet *məšxáwt*, dans lequel *š* est préservé (il passe normalement à *h* ou zéro comme en mehri).

Souvent, l'emprunt arabe se manifeste sur le plan phonétique.

Toutes les langues SAM distinguent nettement *z* et *ʃ* (cette dernière réalisée souvent *ʃ̣*, surtout en H) : ainsi la racine de M *zəfərēt*, H *zəfərōt*, J *zəfrēt*, S *záfreh*, "tresse de cheveux", n'est pas celle de M *ʃayfēr*, etc., "ongle" (en socotri avec *t* correspondant à *ʃ*).

Ce n'est pas le cas des dialectes arabes voisins, qui ne distinguent pas *dād* ض et *ḏā'* ظ, réalisés tous deux *ḏ* ; on peut donc en tirer des indications sur les emprunts : H, M, J, H *ʃarb*, S *ʃarb* "morceau de bois" est SAM (racine *ʃrb*), tandis que H *ʃarb* "sorte de"; H M *ʃərbēt*, J *ʃarbēt* "coup de vent", J aussi "détonation d'une arme", sont des emprunts à l'ARABE (racine *ʃrb*). De même, M *ʃərūr* "mal" (nom), est un emprunt (ar. *ḏrr*), mais H *zərr*, J *zerr*, M *zər* "faire du mal" sont SAM.

Un cas particulier est celui de la racine **ʃhr* : en relèvent H *ʃəhōr*, *ʃəhār*, *zəhōr*, *ləhōr*, M *zəhēr*, *zəhār*, J *zəhēr* "apparaître" et de nombreuses formes dérivées en *zhr* ainsi que M H *ʃahr*, J *ʃəhur* "midi", H *ʃahr*, *zahr*, M *zāhər*, *zāhər*, J *zəhər* "dos". Il y a sans doute là le résultat d'une série d'emprunts arabes, plus ou moins intégrés, faits à diverses dates. Les formes J *ʃəhēr* "être fini", S *ʃāhar* "partir" sont, elles, proprement SAM.

Une correspondance anormale des sifflantes sémitiques est un indice : ainsi M *sənēt* "année" n'est pas conforme : on attendrait **hənēt*. Certains dialectes ont *sənéh* dont la marque de féminin arabe est encore visible. De même, dans : H *ʔətōš*, M *ʔawtəh*, *ʔətōs*, J *ʔətōš*, *ʔətōš*, S *ʔətāš*, "éternuer", sémitique **ʔataš*, il y a deux anomalies :

a) Le harsusi devrait avoir une forme en *h*, telle que *ʔawtəh*. Il peut s'agir d'un emprunt au jibbali, mais aussi d'un cas de conservation de *š* comme cela se produit dans quelques mots (voir n. 7 ci-dessus).

b) La deuxième variante du mehri, *ʔətōs* est un emprunt de l'arabe *ʔatas*. Les deux variantes du mehri sont l'une et l'autre conformes/intégrées du point de vue du schème : la profonde différence *CáwCəC* :: *CaCōC* s'explique régulièrement par la laryngale en troisième position.

Le harsusi présente un dédoublement *g/y* qui révèle l'emprunt d'un mot arabe (d'un dialecte d'Oman) en *yīm* (pour *jīm*). Les mots non-empruntés n'ont pas la variante *y*. Ex. *fayr*, *fagr* "aube", *neyáh*, *nəgáh* "réussir", *wāyeb* "nécessaire" ; comparer *yann* "folie", *génni* "djinn", tous deux empruntés, mais le second à un arabe plus standard (le mot SAM pour "djinn" est *kēʕi*).

Le mehri du Dhofar présente, beaucoup moins souvent, une variation correspondante *g/j* : comparer *jamʕ* "poing" et *gamʔ* "prière communautaire", emprunts respectifs à un dialecte local et à l'arabe standard, le second mieux intégré.

Il en est de même pour *k/g(j)* : le harsusi emprunte des mots arabes en *gāf* et même *jāf* (pour *qāf*).⁸ Comparer H *gəbāyli* "membre d'une tribu" (emprunt ar.) et *kəbīyəl* "tribus", *tagrīb* "presque" (emprunt) et *karb* "proximité" (*garb* est la "gale" !), H *gháweh* (emprunt) et M *kəhwēt*⁹ "café", H *gəšəbəh* "roseau" (ar. d'Oman *gšibah*) et *kəšəbət* "barre à mine", ces deux derniers mots se confondant dans le M *kəšəbēt*. Les mots non empruntés n'ont pas la variante *g(j)* mais des emprunts à un arabe plus "littéraire" peuvent être en *k* : *məklūʔ* "détestable" (racine *klʕ*).¹⁰

Certains parlars SAM montrent une tendance à réaliser les consonnes emphatiques comme des pharyngalisées-uvularisées de type arabe au lieu des éjectives d'origine. Les phonèmes atteints par le phénomène sont d'abord les fricatives :

8. Ex. *bəriḡ*, *əbrīḡ* "pichet", formes identiques en ar. d'Oman.

9. Le sens moderne, "café", est donné à une vieille forme sémitique **kəhw-at* que le SAM et l'arabe possédaient.

10. Ici, l'emprunt est révélé par le schème de participe passif arabe en *ū* (voir ci-dessous).

<i>t'</i>	> <i>ṭ</i>	IPA :	[θ'] > [ðʕ],
<i>s'</i>	> <i>ṣ</i>	IPA :	[ʃ'] > [ṣʕ],
<i>š'</i>	> <i>ṣ̣</i>	IPA :	[ʃ'] > [ʕʕ],
<i>s'</i>	> <i>z, ṣ</i>	IPA :	[s'] > [zʕ], [sʕ]

[noter le voisement : c'est dans une première phase une *creaky voice* = laryngalisation].

Puis les occlusives sont affectées.

<i>t'</i>	> <i>t</i>	IPA :	[t'] > [tʕ],
<i>k'</i>	> <i>k, ǧ</i>	IPA :	[k'] > [kʕ], [ǧʕ]

On peut penser que cette évolution est à peu près celle que l'arabe le plus ancien a connue. En outre, la situation n'est pas si simple, car on peut observer des dialectes arabes à emphatiques (*t, ṣ, ṭ, q*) réalisées éjectives (API [t', s', θ', k']) dans certaines zones marginales bilingues.

Il peut se produire que pour un mot SAM il y ait interférence avec un mot arabe presque identique, et que l'accentuation arabe s'impose sur le mot SAM, ex. "elle a commencé" S *bádʕʕh*, ar. *bādáʕat*, S avec accentuation arabe : *bádʕʕh*.

On lit parfois que les mots S comportant un *x* ou un *ǧ*, ou les mots M comportant un *ʕ*, sont normalement des emprunts à l'arabe car ces phonèmes sont dans ces langues en coalescence avec d'autres (resp. *h, ʕ, ʔ*). Il y a là une vieille erreur due au fait qu'on a pris les dialectes étudiés les premiers pour la langue elle-même. Quiconque se rend à Socotra peut entendre *ǧayg* "homme", *ḥa* "bouche" dès qu'il étend un peu sa visite au-delà des parlers qui prononcent *ʕayg* et *ḥa*. De même en pays mehri peut-on entendre, pour "il grandit", bien des variantes dialectales : *yáwǧǧar, yəʕǧǧar, iʕǧǧar, iʕǧǧar ...*

L'emprunt est, dans de nombreux cas, révélé par un indice morphologique, tel que la présence d'un schème étranger au SAM. Par exemple, le schème de participe passif du verbe simple est *mVCCūC* en arabe mais *mVCCīC* en SAM. Dans le socotri *máxlǧǧ* "être humain", on reconnaît l'arabe *maxlūq* "être humain" < "créature" = "créé", dont l'équivalent proprement SAM est H *mǧǧlǧǧ* "être humain", J *mǧǧléǧǧ*, M *mǧǧlǧǧ* "pauvre malheureux", de même pour M *mǧǧbrūǧǧ*, H *ǧmǧǧbrūǧǧ*, vs. J *mǧǧbréǧǧ* "béni". Notons que dans S *máxlǧǧ* le schème a connu une adaptation accentuelle, typique de cette langue.

On voit parfois un doublet :

M	"maçon"	<i>bǧnnāy</i> (< ar.)	&	<i>mǧnnōy</i> ,
H	"finir"	<i>xǧllǧǧ</i> (< ar.)	&	<i>xǧllǧǧ</i> .

Il est probable que les termes qui ne se présentent pas comme empruntés l'aient été en réalité à date ancienne et se soient intégrés.

Les noms quadrilitères *CvCCǧC donnent en SAM *CvCCē/īC. Un tel quadrilitère, typiquement, sera caractérisé par sa voyelle longue (M *ǧǧǧmīm* "beurre frais"), ou, selon les langues, par la voyelle accentuée qui en résulte (J *ǧǧǧmīm*), ou encore par le *h parasite* qui en résulte (S *ǧǧǧmǧhīm*).¹¹ Donc sont manifestement empruntés des noms tels que H *mǧǧbrǧǧ* "lime" (vs. M *mǧǧbrēǧǧ*), M J *mǧǧlǧǧ* "sincère" (< ar. classique *mǧǧlǧǧ*).

L'emprunt peut révéler une forme dérivée verbale surnuméraire : H *ǧǧǧma, ǧǧǧma, ǧǧǧmǧǧ* "se réunir", présente, respectivement, les formes dérivées SAM T1 et T2, et un emprunt arabe, ou encore H *ǧǧǧǧǧar* "devenir insuffisant", *ǧǧǧǧǧar, ǧǧǧǧǧar* "passer (temps)" sont, respectivement, la forme dérivée

11. Voir Lonnet (1994).

SAM II-III et la II^e forme arabe (avec deux variantes en *k/g*). Le participe H *məṭḥáffi* "nu-pieds" (ar. d'Oman *miṭḥáffi*, V^e forme dérivée) ne peut être qu'un emprunt : la dérivation verbale SAM n'a pas de forme à *t* préfixé (il est toujours infixé) ni de gémination.

Les emprunts arabes se signalent aussi souvent par une désinence de féminin arabe en $-v(h)$ au lieu du $v(t)$ SAM :¹² J *tórəh* "révolution", *sáʿáh* "heure ; montre". Mais il arrive que le mot emprunté s'adapte à la morphologie d'accueil : J *rišt* "gâchette, détente" < ar. d'Oman *rīšah*.

Avec une telle proximité morphologique et phonétique, un tel partage du fonds lexical (sud-) arabe et de la culture bédouine ancestrale (sans parler du fait que les Mahra ont été parmi les premiers groupes à suivre le Prophète), et une telle extension du bilinguisme, il est très difficile de trouver un champ lexical qui se révélerait particulièrement riche en emprunts arabes, hormis celui de la modernité (par exemple le lexique du mobilier, celui de l'administration, etc.) ...

Les évaluations et les échanges commerciaux traditionnels se font encore avec les nombres et les unités de mesure traditionnels, mais les activités orientées vers l'extérieur ont introduit assez d'arabe pour éliminer, par exemple, la plupart des noms de nombres SAM au-delà de 10, et produire des systèmes composites assez perturbés :¹³

J <i>stīn</i> "60" (< ar. * <i>sittīn</i>),	mais	<i>šet</i> , f. <i>štət</i> "6",
M <i>sōdəs</i> (< ar. * <i>sādis</i>),	mais	f. <i>šədtīt</i> "6 ^{ème} "; <i>hət</i> , f. <i>yəttīt</i> "6",
H <i>təlatṭáʿəs</i> "13" (ar.),	mais	<i>šəlayś</i> , f. <i>šāfayt</i> (< <i>šaʿṭayt</i>) "3"; <i>šələt</i> ou <i>šələs</i> "3 ^{ème} ".

Cependant, certains dialectes socotri ont encore des noms de dizaines tels que *ħaʿt ʿəsárhən* "60" [six dizaines],¹⁴ ce qui est une remarquable résistance à l'emprunt.

Enfin, on pourra noter l'abondance de particules grammaticales ou particules de discours empruntées à l'arabe :

- M H *yəkūn* "peut-être", mais SAM J (*ħa-*)*yékən* "peut-être ; à peu près" ;
- M *yómkən* "peut-être" ;
- H *yā* "ou bien ; je veux dire" ;
- H *yā-rēt*, M *yā-rayt*, J *ya-rét* "ah ! si seulement ... !" ;
- HM *yā* "ô ..." (vocatif), mais la particule SAM est représentée par M *ʿā* ;
- J *bdan*, M *ʿábdan* "jamais, pas du tout", mais SAM MH *bəhāwʿ*?, J *bəhóʿ*?, particule empruntée par l'ar. d'Oman : *bhawʿ* ;
- S *mseb* "parce que" < ar. *min sabab* ;
- MJ *ʿəkīd* "certainement !" ;
- MJ *ʿaywa(h)*, *ēwa* "oui", mais SAM M *yēya*, *ʿahā*, J *yéyφ*, *ʿahān* ;
- M *hēsən* "quoi ?" < ar. *ʿāšən* < *ʿayy šayʿin* "quelle chose ?".

12. Le socotri n'est pas ici concerné car il présente *-h* et non *-t*.

13. « [M]any centuries of symbiosis with a culturally dominant language have brought with them large-scale borrowings from Arabic, to produce in these languages to a greater or lesser degree mixed systems. This is most noticeable in the numbers above ten ... » T. M. Johnstone : "The system of enumeration in the South Arabian Languages", *Arabian and Islamic Studies. Articles presented to R. B. Serjeant on the occasion of his retirement from the Sir Thomas Adam's Chair of Arabic at the University of Cambridge*, R. L. Bidwell and G. R. Smith, eds., London & New York, Longman, 1983, p. 225-228.

14. On remarque dans *ħaʿt* un *ʿ* non-étymologique.

Une forme d'emprunt particulièrement « avancée » est celle où une particule empruntée présente une certaine intégration (dans l'exemple mehri suivant, la perte de gémination), mais fonctionne exclusivement avec les pronoms suffixes arabes :

M ən-, lən- H yənn- + suffixe arabe, "parce que" < arabe *(li-)?anna-.

Les emprunts attendus qui *ne se produisent pas* sont significatifs au plan culturel : le nom d'Allah est d'usage courant en socotri, mais en mehri on rencontre, à sa place, en phraséologie islamique, et même en référence non technique à Dieu, les mots :

(a)bēli < "mon seigneur", racine bʕl,

rahmōn (J erḥəmún) "le miséricordieux",¹⁵

ʔāl < sémitique *ʔal "dieu" (ou, moins probablement, emprunté au jibbali : *a-baʕl > *āʕl > J ʕʕʕ > ʕʕ et autres variantes : ʔʕʕʕ, ʔʕʕʕʕ).¹⁶

Une formule de remerciement archaïque est ʔāl lūk byāz "que Dieu soit généreux sur toi", dans laquelle le mot byāz est homologue de l'arabe ʔabyaḍ "blanc". La formule moderne est bēli šūk "le Seigneur (soit) avec toi."¹⁷

Quelques exclamations comportant l'invocation d'Allah ont été empruntées, sans que leur motivation soit perceptible (elle l'est à peine en arabe), et l'on entend :

J ʔəllə! "eh bien !", bəllé? "vraiment ?", wəllə! "parfaitement !"

Références

JOHNSTONE, Thomas M. 1987. *Mehri lexicon and English-Mehri word-list*. London: School of Oriental and African Studies. [Bound with *Index of the English definitions in the Jibbāli Lexicon, compiled by G. Rex Smith*]

JOHNSTONE, Thomas M. 1981. *Jibbāli Lexicon*. London: Oxford University Press.

JOHNSTONE, Thomas M. 1977. *Ḥarsūsi Lexicon and English-Ḥarsūsi word-list*. London: Oxford University Press.

LESLAU, Wolf. 1938. *Lexique soqotri (sudarabique moderne) avec comparaisons et explications étymologiques*. Paris: Librairie C. Klincksieck.

LONNET, Antoine. 1994. "Quelques résultats en linguistique sudarabique moderne". *Quaderni di Studi Arabi* 11, p. 37–82.

15. Plus probablement la persistance du nom d'un dieu pré-islamique, associé par le Prophète à celui d'Allah, et (re-)devenu en arabe un simple qualificatif.

16. Notons qu'en mehri le nom ʕAbdallah se dit aussi Abdāl.

17. Au masc. ; fém. šīk, mpl. šīkəm, fpl. šīkən.