Strata on loanwords from Arabic and other Semitic languages in Northern Somali

Giorgio Banti*

1. Introduction

The term Somali (af-Soomaaliga ‘the Somali language’) is frequently used for referring to two different things that should not be confused: (a.) the written Somali language and the corresponding spoken koiné and (b.) the Northern Somali (NS) dialects and their most closely related varieties. After Heine (1978), Ehret and Mohamed Nuuh Ali (1984), and Lamberti (1986, 1988) there has been a general consensus that the following subgroups should be included in the Somali cluster:¹

(i.a.) the NS dialects spoken in Djibouti, present-day Somaliland and Puntland, eastern and southern Ethiopia, north-eastern Kenya and the adjoining regions of southern Somalia west of the Juba river (ca. 14.000.000 speakers together with the Benaadir subgroup);

(i.b.) the Benaadir dialects spoken in Somalia from the southern Mudug region to the left bank of the Shabelle river and the area of Merka (ca. 14.000.000 speakers together with the NS subgroup);

* The author is grateful to Alessandra Avanzini, Peter Behnstedt, Cabdalla Cumar Mansuur, Riccardo Contini, and Mena Lafkioui who, at different times and in different ways, helped him in preparing this chapter. He also thanks Bertrand Hirsch and François-Xavier Fauvelle-Aymar for allowing him to reproduce as Figure 1 one of their maps. Obviously, any mistake in this contribution is only due to the present author. His special compliments go to the team of Corpus of South Arabian Inscriptions of the University of Pisa for having prepared such a rich and useful database. Somali is spelt here with its official orthography; its main peculiarities are: ’ for glottal [ʔ], c for [͡ʃ] (i.e., Arabic ‘ayn), dh for retroflex [d] and its allophone [tʃ], and x for [h]. Notice Ar.= Arabic, ASA= Ancient South Arabian, Har.= Harari; MSA= Modern South Arabian.

¹ The distribution of the subgroups indicated below does not take into account the worldwide Somali diaspora in several Arab countries, India, eastern and southern Africa, Australia, Europe, and North America.
(i.c.) the so-called Ashraaf dialects spoken mainly for in-group communication in several families of Mogadishu, Merka and some smaller settlements along the intervening coast;

(ii.) the Maay dialects spoken by almost 2,000,000 pastoralists and farmers in the interriverine area of Somalia from Wanle Weyn to the eastern bank of the Juba River;

(iii.a.) a cluster of rather diverse dialects spoken by ca. 130,000 pastoralists and farmers partly in the same area as Maay, from Baidoa to the mouth of the Juba River. It includes Tunni, Dabarre (and Irroole), Garre and Jiiddu. The Garre who live near Moyale, on the border between Kenya and Ethiopia, speak a southern Oromo dialect;

(iii.b.) the Boni dialects spoken by ca. 8000 hunter-gatherers from the southern tip of Somalia to the eastern bank of the Tana River in southeastern Kenya, and

(iv.) Rendille, spoken by ca. 35,000 pastoralists between Marsabit and Lake Turkana in northern Kenya.

Boni has been shown by Tosco (1994) to be closely linked to Garre, and thus to be actually a southern extension of subgroup (iii). Abdirachid Mohamed Ismaïl (2011) has suggested on the basis of mainly phonological and lexical isoglosses that subgroups (ii.) and (iii.) should be regarded as a single subgroup of dialects, that he called entre-deux-fleuves, i.e., inter-riverine. Notice that this is basically a subgroup formed by language contact and convergence, rather than a genetic taxon. The ensuing major articulation of the Somali cluster is thus:

1) Somali strictu sensu: Northern Somali + Benaadir + Ashraaf;

2) Inter-riverine: Maay + Tunni + Dabarre + Garre–Boni + Jiiddu;

3) Rendille.

Small groups of speakers such as the Bayso and the Girirra use languages related to the Somali cluster in southern Ethiopia.

Together with Arbore, its close relative Elmolo, and Dhaasanač, the Somali cluster is regarded by linguists (cf. Sasse 1979: 3 f.) as forming the Omo-Tana subgroup of East Cushitic, one of the major branches of the Cushitic language family. NS, the Benaadir and the Ashraaf dialects are characterized by the following historical features:
they preserve *c and *h as separate phonemes, e.g., NS mágac ‘name’ vs. Maay magó, cAfar migáč, Oromo maqáa; NS madáx ‘head’ vs. Maay madó, Oromo matáa, etc.;

they develop *z > d in all positions, e.g., NS wádne ‘heart’ vs. Tunni wésna, Rendille wéyna (regularly with *z > y), Northern Saho and Kambata wazana, etc.;

they develop *m > n in syllable codas, e.g., NS nín sg. ‘man’ but pl. nimán (still namán in the southern Mudug region) vs. Jiiddu lám, Oromo namá, cAfar núm, etc.

NS, the Banaadir dialects and, in most but not all position, the Ashraaf ones also develop *k > g and *t > d after vowels as in:

NS ilk-ó pl. ‘teeth’ vs. ilíg sg. (< *ilk-) ‘tooth’, cf. Oromo ilk-áan ‘teeth’, etc.;

NS madáx ‘head’ vs. Oromo matáa, etc.

It should also be remarked that dh, i.e., [d], whatever its origin, evolves into r after a vowel in the Ashraaf and Benaadir dialects and in present-day eastern NS, but is either preserved or weakened to [ɾ] in the rest of NS, e.g., Benaadir and eastern NS yiri ‘he said (it)’ vs. other varieties of NS yidhi [jídi] ~ [jíri] ‘id.’ from the prefix-conjugated verb *-eqh-‘say’, cf. Rendille yidæh ‘he said’, Saho-cAfar yedhe ‘id.’, southern Oromo yed’e ‘id.’

2. Avenues of contact of Northern Somali with Semitic

Now the main NS-speaking area is in contact with the following languages and language (sub-)groups:

a) cAfar (East Cushitic) in the north-west;
b) Oromo (East Cushitic) in the west and south-west;
c) Harari (Southern Ethiosemitic) in the west;
d) inter-riverine Somali dialects in the centre and south;
e) minority Bantu languages (Mushunguli, and the two northernmost Swahili dialects, i.e., Mwiini and Bajuni) in the south;
f) Amharic, widely spoken as a second language in the Somali-speaking areas of Ethiopia;
g) Arabic (Ar.), known by religious people, educated Somalis, Somali traders and business people, and by Somalis who have been living and working in the Arab countries (see also Simeone-Senelle 2006).

This picture, however, has been changing through time, and cannot be naively projected into the past. For instance, it is reasonable to assume that the protracted wars between Christian and Muslim polities in central Ethiopia during the late Middle Ages caused considerable changes in the distribution of languages after the 16th century in the central and eastern Horn (see later in this section).

On the other hand, present-day NS includes a considerable amount of loans from the European colonial languages that came in contact with it since the end of the 19th century: Italian in former Italian Somalia, English in Somaliland and north-eastern Kenya, and French in Djibouti. Kenyan NS also borrowed more heavily than elsewhere from Swahili, whose role as a lingua franca in Kenya has grown considerably from colonial times till now. In addition to this, Cardona (1988) and Cabdalla Cumar Mansuur (2008: 26) identified a number of Persian and Indian loanwords, many of which entered NS through the lingua francas used across the Indian Ocean “among seamen, traders, soldiers” (Cardona 1988: 24). Cardona writes on the same page that such lingua francas have been in use “since the Middle Ages”, but it is likely that some forms of them were already used on the northern Somali coast in classical antiquity. Indeed the Periplus of the Erythraean Sea, a description of the sea lanes and trading opportunities from Roman Egypt to India probably written in the middle of the 1st century, explicitly states that the τοῦ πέραν ἐμπόρια “far-side ports of trade”, i.e., the ports that lay beyond the Bab-el-Mandeb strait on the coast of northern Somalia and immediately to the south of Cape Guardafui traded at that time both with Egypt and with the Indian port of Barygaza, which is usually identified with present-day Bharuch in Gujarat, at the mouth of the Narmada river. These “far-side ports of trade” were, e.g., Avalites (ὁ Ἀυαλίτης), Malao (ἡ Μαλαώ), Mundu (ἡ Μούνδου), Tabai (τὸ ... Τάβαι) and Opone (‘Οπωνή).
προηγουµένως εἰς ταύτα τὰ ἐμπόρια πλέουσιν, οἱ δὲ [κατὰ] τὸν παράπλουν ἀντιφορτίζονται τὰ ἐμπεσόντα. Οὐ βασιλεύεται δὲ ὁ τόπος, ἀλλὰ τυράννοις ἰδίοις καθ’ ἐκαστὸν ἐμπόριον διοικεῖται.

‘Departure from Egypt to all these far-side ports of trade is made around the month of July, that is Epiphi. To these far-side ports of trade it is also common to ship in from the inner regions of Ariaca and Barygaza goods from those places: wheat, rice, clarified butter, sesame oil, cotton cloth, both the monache and the sagmatogene kinds, and girdles, and honey from the reed called sacchari. Some make the voyage especially to these ports of trade, and others exchange their cargoes while sailing along the coast. This area is not ruled by a king, but each port of trade is ruled by its separate chief.’

(Periplus of the Erythrean Sea, 14.1 f.; translation adapted from Casson 1989)

It is reasonable to assume that the lingua franca used along this trade route was at that time heavily lexified with Indian, South Arabian, Coptic and Greek terms. Old Ar., i.e., the older form of what is now called Ar., probably played just a minor role at that time.

A little later, while speaking of the ports that lie farther south along the African coast of the Indian Ocean, such as Rhapta (τὰ Ῥαπτά), the author of the Periplus wrote:

Μέγιστοι δὲ (ἐν) σώμασι περὶ ταύτην τὴν χώραν ἀνθρωποί ὑπατοὶ κατοικοῦσιν καὶ κατὰ τόπον ἐκαστὸς ὁμώνυμοι τιθέμενοι τυράννοις. Νέμεται δὲ αὐτήν, κατὰ τὸ δίκαιον ἀρχαῖον ὑποπίπτουσαν τῇ βασιλείᾳ τῆς πρώτης γιονμένης Ἀραβίας, ὁ Μοφαρείτης τύραννος. Παρὰ δὲ τοῦ βασιλέως ὑποθορον αὐτήν ἔχουσιν οἱ ἀπὸ Μοῦζα, καὶ πέμπουσιν εἰς αὐτήν ἐφόλκια τὰ πλέονα κυβερνήταις καὶ χρειακοίς Ἀραβικές κράτον εἰς τοὺς καὶ εὐρετούς ἔχουσιν ἐμπειροὺς τοὺς καὶ τῆς φονῆς αὐτῶν.

‘Very big-bodied men inhabit the region, under separate chiefs for each place. The region is under the rule of the governor of Mopharitis, since by some ancient right it is subject to the kingdom that has become first in Arabia. The people of Muza receive tributes from it through a grant from the king, and send out to it many large ships, using Arab captains and crews, who are familiar with the natives and intermarry with them, and who know the places and their language.’

(Periplus of the Erythrean Sea, 16.5 f.; translation adapted from Casson 1989)
At that time, thus, part of the East African coast to the south of Cape Guardafui had political and commercial links with the port of Muza, in present day south-western Yemen not far north from the Bab-el-Mandeb, and with the country of Mopharitis (aka Mapharitis, possibly to be identified with present-day Ma‘āfar), that lay inland of it. It is difficult to assume that the ships from Muza sailed directly to the region of Rhapta without stopping at the “far-side ports of trade” mentioned earlier, in *Periplus* 14.1 f. As a consequence, even though it is possible that Malao, Mundu or Opone were not paying tributes to Muza like the more southern region of Rhapta, they had to have contacts with its sailors and merchants.

However, one should beware of assuming that these “Arab captains and crews” spoke Old Ar. The *Arabes* of classical antiquity were generically the inhabitants of the Arabian Peninsula, whatever language they spoke or ethnical group they belonged to. In particular, Muza and the country of Mopharitis were on the western margin of the area of the Himyarites, already known to the author of the *Periplus as Homērītai* (‘Οµηρίται), who generally used south and late Sabaic in their inscriptions. There are different opinions about the language they actually spoke: Robin (2007) insists upon the differences between pre-Islamic Himyaritic and Sabaic, and seems to regard it as a Semitic language altogether different from ASA, whereas Stein (2008, 2012) stresses their similarities and regards Himyaritic as a development of Late Sabaic. Himyaritic was still known to medieval Arab scholars as a separate language, and later converged into Ar. leaving even morphological relics in some contemporary Yemeni dialects, such as the article (*V)n- ~ (*V)m- (< Himyaritic <hn->) and the 1st and 2nd persons with -k- in the perfect.

In any case, the language(s) spoken by the seamen and traders from Muza can be regarded as a source of Semitic loanwords in the language(s) of the communities that lived on the southern coast of the Gulf of Aden. As a working hypothesis, Stein’s suggestion will be followed here, and possible cognates of NS terms will be looked for in what is known about Middle and Late Sabaic.

There is no direct evidence about the language(s) of the northern Somali coast at the time of the *Periplus*, beyond the obscure place names it mentions such as Malao (f.), Mundu (f.), Tabai (neuter), etc. But there is no compelling reason for thinking that it could not be an older form of NS. Indeed, even if Mohamed Nuuh Ali (1985: 82 f.) suggests that “a North

---

2. Possibly corresponding to Mawza’, which now lies several miles inland because the coastline has receded.
Lowland Eastern Cushitic”, i.e., Saho-‘Afar “language could have been spoken both in the trading ports and in the hinterland”, he concedes that the evidence for this is very scarce. The group of items he lists as Saho-‘Afar loanwords in NS (1985: 144 ff.) has to be taken with great care because (i.) they may have been borrowed much later, when NS was indisputably in contact with ‘Afar as it is today, and (ii.) for many of them there is no evidence that the direction was from Saho-‘Afar into NS. For instance, he regards NS magáalo ‘town’ as borrowed from ‘Afar magaalá ‘town, village’, because of its “distribution” (1985: 148). However, the word also occurs in Harari magâla ‘market’ and eastern Oromo magaláa ~ magaalá ‘market, town’. It is thus attested in the four main languages spoken today in the eastern Horn. But since it lacks an obvious etymology in any of these four languages, it may also have been borrowed from a fifth, unidentified language. The only clear fact is that there is no compelling evidence that it has been borrowed from ‘Afar into NS!

A second wave of Semitic loanwords reached the region where NS is spoken today through the spread of Islam and the ensuing growing number of people who learnt some Ar. because of their religious studies and practice, as well as the growing role of some form of Ar., both in many areas of the southern Arabian Peninsula where South Arabian and Himyaritic were spoken previously, and as a trade language along the sea routes that connected the eastern Horn with Egypt, the southern Arabian Peninsula, and Mesopotamia.

The oldest known Ar. inscription from Somalia lay on the tomb of Abū ʿAbdallāh bin Rāyā an-Naysābūrī, found by Cerulli to the north-east of the old district of Shingaani in Mogadishu (Cerulli 1957: 2 f.). This Persian from Nishapur in the northeastern province of Khorasan died in June-July 1217:

\begin{quote}
	توفّى العبد الضعيف الراجى إلى رحمة الله ابّو عبد الله بن رايا بن محمد بن أحمد النيسابوري الخراسانى توفيّ يوم الثلث في شهر ربيع الأول مضى ستة مائة واربعة عشر سنة ١١٤
\end{quote}

‘The weak slave, hoping in God’s mercy, Abū ʿAbdallāh bin Rāyā bin Muḥammad bin Aḥmad an-Naysābūrī al-Ḥurāsānī died. He died on the day of Tuesday in the month of Rabī’ al-ʿAwal; six hundred and fourteen years had passed (since the Hijra).’

Since then, Ar. has been the only language written by the Somalis until the end if the 19.th century, when the first examples of Somali written in Ajami, in new local alphabets such as the Osmaniya, and in Latin script start to appear (see, e.g., Banti 2011b).
One important avenue through which the knowledge and use of Ar. spread to areas here NS is now spoken were the above-mentioned sea routes that connected the northern and eastern coast of the Horn with the Arabian Peninsula, Egypt and beyond. But commercial, religious and political relationships with the Islamic polities in the interior of the Horn were an important and often disregarded avenue for the spread of Ar. loanwords into the NS-speaking communities. In recent years, Fauvelle-Aymar and Hirsch (e.g., 2008 and 2011) have published the results of several archaeological campaigns and historical studies aimed at shedding more light on such polities during the Middle Ages.

One of their maps, reproduced here as (Figure 1), shows a major cluster of early Islamic sites pointing south from Berbera, and south-east from this town and Zeyla towards the Chercher Mountains and the lakes of the Rift Valley. Many of these sites are towns with stone buildings, that were already abandoned and in ruins when the first European travelers saw them in the 19th century. Harar and the Argobba villages to the south-west of it, like Koromi, are the only ones that are still thriving today.

In the first half of the 16th century, speakers of NS were already present in this region. The Futūḥ al-Ḥabaša, an eye-witness account of the wars between Christian Abyssinia and the Islamic polities led by Aḥmed the Left-handed (1506–1543 ca.), mention the Somali as one of the major elements in his army, and name among them several now NS-speaking tribes such as the Harti, the Marreexaan, and the Habar Magadle. Harar has been a major political and religious centre also for many NS areas at least from that time. Nowadays, several Somali clans still use for their traditional chiefs Har. titles such as garāad and malâaq, and the Har. word āw ‘father’ as a religious title.

More recently, direct political, educational, and cultural relations with the Ar.-speaking countries have increased during the 20th century, especially after independence was reached. Since then, Ar. has been one of the three languages of the country. Many educated Somali’s have studied in Arab universities, and Somalia joined the Arab league during the seventies.

Not only Ar. has been used in Harar. Its inhabitants speak a Southern Ethiosemitic language, Harari, which has been written in Ajami for several centuries. The oldest dated manuscript is from the very beginning of the 18th century, but it has to be assumed that Ajami Harari was already written at least during the second half of the 16th century (cf. Cerulli 1936: 439; Banti 2005a: 77 etc., 2010: 153 f.). The Somalis who fought in the
armies of Ahmed the Left-handed were certainly in contact with speakers of Harari in the first half of the 16th century.

It is not known when Southern Ethiosemitic spread into central Ethiopia, and we can only suggest hypotheses about the languages that were spoken in the medieval Muslim towns along the northern and southern sides of the Rift valley. What appears to be certain is that:

i.) the town of Harar has been a major political, religious and economic centre at least from the 16th century;

ii.) there are no obviously Oromo names in medieval Christian and Muslim Ethiopian sources before the XVI century, nor in the Futūḥ al-Ḥabaša (cf. Banti 2011a);

iii.) the Muslim Argobba to the south-east of Harar now speak Oromo and some Harari, but they still preserved in the late 19th century songs in their separate Ethiosemitic language;

iv.) other Muslim Argobba groups, on the western escarpment of the Rift Valley, speak several varieties of Argobba, a Southern Ethiosemitic language closely related to Amharic;

v.) Harari is closely related to Eastern Gurage languages such as Silt’e, now separated from it by a wide stretch of Oromo-speaking territory;

vi.) medieval Ethiopian sources, the Futūḥ al-Habaša, and local oral traditions mention the Ḥarla as an important ethnic group in Muslim eastern Ethiopia, but it is not known what language or language(s) they spoke at that time (cf. Banti 2005b; Chekroun et al. 2011). Now there are tiny groups calling themselves Xarla, i.e., [Ḥarla], among the north-western Somalis and Haralla among the ḌAfar, and speaking the language of the people among whom they live.

Fact (ii.), together with the Abyssinian historical traditions, is generally construed as an indication (a.) that Oromo ethnic identities and the Oromo language reached their present distribution in Ethiopia only after the first half of the 16th century and, (b.) that the protracted wars between Christian and Muslim polities in central Ethiopia during the late Middle Ages caused considerable changes in the distribution of languages and ethnic identities after that century. For instance, fact (vi.) shows that even major groups vanished after that time, and that their remnants merged with the new dominant groups and languages. On the other hand, facts (iii.) to (v.) above are seen as evidence of a much wider distribution of Southern Ethiosemitic in the
eastern Horn before the spread of Oromo, as Wagner (2009: 112) repeated recently. It is thus quite likely that there were substantial numbers of speakers of some variety of Southern Ethiosemitic also in the Muslim towns whose ruins lie in presently NS-speaking territory, along the old caravan routes between Harar and the coastal towns of Zeila and Berbera. Just like Harar is still now a place where NS is in contact with Southern Ethiosemitic, also these other towns probably must have played a similar role as places where speakers of NS met speakers of Southern Ethiosemitic languages.

The roles of Harar and, possibly, also of speakers of other Southern Ethiosemitic language varieties in an Islamic setting in the eastern Horn of Africa, make it likely that Southern Ethiosemitic not only transmitted Ethiosemitic terms to NS, but also Ar. words it had already borrowed.

3. South Arabian (SA) loan-words

In this section and in the two that follow it, NS items have been drawn mainly from the two major monolingual dictionaries that have been published till now: Yaasiin Cismaan Keenadiid (1976), and Saalax Xaashi Carab (2004). Although the Somali spelling does not mark pitch accent, tone marks have been added here for ease of pronunciation.

An Ancient South Arabian (ASA) loanword in NS was identified by Cerulli (1959: 119), i.e., NS addōon m. ‘slave’. But a systematic study of the contacts between ASA, Himyaritic, and NS has not been done yet. Nevertheless, a few possible ASA (or Himyaritic) loanwords can be identified in NS.

NS addōon m. and addōon f. ‘slave’ (pl. addoomō m.) has -n < -m in syllable coda, a regular development in NS (cf. § 1.) Cerulli (1959: 119) connected it with Sabaic 'dm ‘servants, vassals, religious servants’. This particular semantic development of the well-known Semitic root 'DM still survives in Modern Yemeni Ar. 'awādim ‘domestics’ (Ricks 1989: 5). Notice that -oo- is here probably from *-a- or *-ā-, a development that took place in the donor language, not in NS. Modern South Arabian (MSA) languages have several instances of accented *-ā- > -ō- (> -ū-).

3. A third monolingual Somali dictionary, edited by Annarita Pugielli and Cabdalla Cumar Mansuur, was recently published but not yet available to the author while preparing this chapter.
NS gêel m. ‘camel’ has -ee- < -aa-, that is still preserved, e.g., in NS compounds like haruub-gâal m. ‘vessel for milking camels’ or daba-ggâalle m. ‘ground squirrel’ (lit. ‘having a camel tail’), etc. East Cushitic *gaala ‘camel’ is from Semitic *gamal- ‘id.’ and the most likely intermediary has been a language spoken in Southern Arabia when camels were introduced to the eastern Horn ca. 2000 years ago, as argued in Banti (2000: 49 ff.).

Gml ‘camel’ is well attested in Sabaic. However, the loss of -m- is not regular in ASA, nor in NS or any known East Cushitic language, and points to a different language: loss of -m- between vowels occurs several times in Southern Ethiosemitic and in MSA.

NS jâdar m. ‘incense from Boswellia frereana’ (cf. Axmed Cartan Xaange 1984: 23 ff.; Saalax Xaashi Carab 2004: 295a “pure incense”). The root occurs in Middle Sabaic mqârt ‘incense altar’ and in Ar. miqâtar ‘censer’, but Ar. q never gives NS j. However, if ASA (and Himyaritic) had an ejective realization of q like MSA and Ethiosemitic, i.e., [k’], *q- > j in NS jâdar parallels *k’ ~ j- in NS jëb-i ‘break (tr.)’ < *k’eb-, a word with a good East Cushitic etymology, as shown by Dhaasanech g’ëb ‘break (tr.)’, Oromo ç’ab-s- ‘id.’, Konso qep-š- ‘id.’, etc. (cf. Sasse 1979: 49; Banti and Contini 1997: 173 f.)

In several other instances, words that have been mentioned in the literature as loans from Ar. are already attested in ASA, and may thus have entered NS at a much earlier date than usually assumed:

1. NS luubâan f. (also lubâan and lubaanâd) ‘frankincense’ either directly from ASA lbny (attested in Sabaic and Qatabanian), or through Ar. lubân.

2. NS berkêd f. ‘cistern, artificial pond’ from Sabaic brkt ‘id.’ (Beeston et al. 1982: 31) either directly or through Ar. birka ‘cistern’, ‘pool’. According to Behnstedt and Woidich (2011, Map 272) this word is used only in the mountains of southern Yemen and the Tihama, with some further isolated attestations in Jordan, northern Syria, the Gulf and southern Oman. Notice that a direct borrowing from Ar. would have been expected to be *birkad or *birko, as in Har. birka.

3. NS gidâar m. ‘wall’ either directly from ASA gdr ‘id.’ (e.g., Qatabanian), or from an Ar. dialect that had a velar realization of ġim in Ar. ġidâr ‘id.’, cf. Har. ġidâr ‘thickness of wall’ with the palatal affricate.

4. NS qaalin m. and qaalîn f. (pl. qaalm-ô) ‘young camel or calf when it approaches sexual maturity’ (nêef gêel âh ama lô’ oo dá’ yár Yaasiin Cismaan Keenadiid 1976: 333b), with ġ > q and -m > -n in
syllable coda, that are regular developments. From the Semitic root \( \text{GLM} \), attested in Sabaic \( \text{ glm} \) ‘boy, child’, or directly from Ar. \( \text{ gālim} \) (participle of \( \text{ galima} \)) or \( \text{ galim} \) both meaning ‘excited by lust’, from the same root Ar. has \( \text{ gillīm} \) (he-camel) excited by lust’ and \( \text{ gulām} \) ‘a young man, youth’. The present author already pointed out (Banti 2000) that it also occurs in Rendille \( \text{ khāalim} \) m. ‘male camel calf’ and \( \text{ khaalím} \) f. ‘female camel calf’ and is thus unlikely to be a recent loanword. The ASA hypothesis assigns \( \text{ qaalin} \) to the oldest core of camel-related terms that entered the eastern Horn, together with *\( \text{ gaala} \) ‘camel’ > NS \( \text{ gēel} \); in this case, the extension to bovine calves is a secondary development.

The fragmentary nature of what is known about ASA and, even worse, about Himyaritic makes it difficult to identify many other terms as old loanwords from these languages. But in the four above cases, there is no compelling evidence for excluding a pre-Islamic origin. A more careful examination of the NS lexicon related to camels, building, and incense may yield some other findings.

### 4. Arabic loanwords

Present-day NS and Benaadiri varieties are full of Ar. loanwords, even though their occurrence may vary in individual texts or dictionary sources. For instance:

- the two above-mentioned monolingual dictionaries for native speakers of Somali don’t include many of the more obvious Ar. loanwords, because their authors regarded them as being “just plain Arabic”, as one of them explained almost thirty years ago;
- instead, Agostini et al. (1985), a Somali-Italian dictionary prepared by a team of Italian and Somali linguists, includes a higher amount of Ar. loanwords, because many of them are not obvious for European users who are not familiar with this language;
- sermons and written texts about Islam are teeming with Ar. words;
- newspaper articles about contemporary politics display a considerable number of terms of Ar. origin;
- traditional narratives display lower percentages of Ar. loanwords.
The issue of Ar. loanwords in NS and their phonological and grammatical adaptation have been discussed by several authors, with different methodologies, e. g., Reinisch (1903: 1–23), Zaborski (1967, 2008), Callegari (1988), Cardona (1988), Soravia (1994), Cabdalla Cumar Mansuur (2008: 25 ff.), Vasaturo (2012), etc. Callegari (1988) indentified 1297 Ar. loanwords in her mainly NS corpus, and classified them in the following meaning groups:

1) religione (‘religion’);
2) utensili, vestiti, attrezzi (‘household implements, garments, tools’);
3) cultura, insegnamento, grammatical (‘culture, education, grammar’);
4) amministrazione, politica (‘public administration, politics’);
5) relazioni sociali (‘social relationships’);
6) commercio, economia (‘trade, economy’);
7) corpo umano, medicina (‘human body, medicine’);
8) sentimenti, comportamenti umani (‘emotions, human behaviors’);
9) legge (‘law’);
10) calendario, tempo, giorni della settimana (‘calendar, time, days of the week’);
11) piante, frutta, profumi (‘plants, fruits, perfumes’);
12) cibi, bevande (‘foods, beverages’);
13) edifici, abitazioni (‘buildings, dwellings’);
14) navigazione (‘seafaring’);
15) professioni, mondo del lavoro (‘crafts, work’);
16) geografia, natura (‘geography, nature’);
17) parentela, rapporti familiari (‘kinship, family relations’);
18) gioco, prostituzione, alcolismo (‘games, prostitution, alcoholism’);
19) spostamenti, viaggi (‘trips, travels’);
20) animali (‘animals’);
21) minerali, gioielli (‘minerals, jewels’);
22) misure (‘measures’);
23) insulti (‘insults’);
24) nazionalità (‘nationalities’);
25) varie (‘miscellaneous’, i.e., words that had not been listed above);
26) avverbi, preposizioni, particelle, esclamazioni (‘adverbs, prepositions, particles, exclamations’).

Soravia (1994: 201) identified 1436 Arabic loanwords using Agostini et al. (1985) as his corpus. As mentioned above, this dictionary tends to record more Arabic loanwords than the two monolingual ones. In addition to this,
it should also be pointed out that it includes some Benaadiri items beside the NS lexicon, because it aimed at recording “Somali” in a broader sense. Soravia (1994: 202) remarks that he assembled a 100-item Swadesh list of basic lexicon on the basis of his corpus, and that it included just one item borrowed from Ar., i.e., ákhdar ‘green’. Yet even this loanword co-occurs with the non-Ar. word cagâar ‘green’ and ‘verdure, lush vegetation’. In other words, he claims that the core vocabulary of Somali and of NS as well, is almost devoid of Ar. loanwords, even though they are extremely frequent in more culturally-bound areas of the lexicon.

There also are several syntactic constructions in NS that can be regarded as grammatical and syntactic borrowings from Arabic, e.g.:

- *iláa Hargeysá* ‘until Hargeisa’, ‘from Hargeisa till here’ replacing the inherited type táníy o H., lit. ‘this and H.’, that has parallels in Saho etc. (an intermediate type is *iláa iyo H.* lit. ‘to and H.’); notice that the only two prepositions of Somali, *iláa* and *mín*, are borrowed from Ar.: otherwise Somali uses local nouns and oblique particles before predicates where most other Semitic languages use prepositions;
- *in* as a generic complementizer for several types of sentential complements, from dialectal Arabic *ín ~ inn* (cf. Banti 2011b: 39 f.);
- the type *iságà oo óoyayá* ‘while he cries, crying’ (lit. ‘he and he cries’) from the Ar. wâw al-ḥāl construction wa-huwa yakbî, dialectally *u-hū bi-yibki* (lit. ‘and he cries’.)

Many of the above-mentioned studies list how Ar. sounds are rendered in the receiving language. Frequently they focus only on the written forms, without taking in consideration that educated speakers may tend to follow an Ar. pronunciation on several occasions. For instance, NS lacks *z*, and Ar. wâzîr ‘minister’ is spelt *wasiïr* in NS, yet educated speakers frequently realize it as *wa[z]ïr* with a voiced fricative. There is thus a considerable degree of variation according to the speaker and to the register of his speech. Nevertheless, some Ar. sounds or clusters are rendered in NS in several cases in different ways that require a more detailed discussion. One of these is Ar. *ɗ* (*dâd*, ض):

- NS *d* as in:
  - *mucáarad* m. ‘opponent’ (< *mušûrid* ‘opponent’ and *mušûraḏa* ‘opposition’);
  - *fadéexo* f. and *fadeexád* f. ‘great shame, scandal’ (< *faḏiha* ‘id.’);
  - *cudúd* f. ‘the upper arm between the shoulder and the elbow; strength’ (< *ṣudud* ~ *ṣudud* ‘id.’); etc.
NS l ~ ll as in:
qaalli m. ‘cadi, Islamic judge’ (< qādin ~ al-qādī ‘id.’); árli m. ~ árlo ‘earth’ (< ard ‘id.’); lid m. and liddi m. ‘person who is against something’ (< ḍidd ‘against’); etc.

- NS s in:
wéeso ‘ritual ablution’ (< wuḏū ‘id.’).
- Possibly NS dh [Ǧ ~ ŏ] in:
wádhaf m. ‘sling’, cognate of Yemeni Ar. waḏafa ‘id.’ according to Behnstedt and Woydich (2011: map 270); however, they record also waẓaf ~ waẓif ‘id.’, awoldtow ‘id.’, and developments of these forms in Maltese and Ar. dialects from Tunisia, northern Algeria and Morocco, and medieval Andalusia, while Piamenta (1990–91: 526a–b) has both waṯuf ‘id.’ and mizāfa ‘id.’; bādh m. ‘half’ if it really is from Ar. baʾd ‘part’, ‘some’.

The most frequent developments are d and l ~ ll, the second one in more common and everyday words (but with exceptions like the above cudūd f. ‘the upper arm between the shoulder and the elbow; strength’), the fist one seemingly in more specialized terms, and in contemporary political language. D > l ~ ll reflects a lateral articulation of this phoneme, that is now lost in most contemporary varieties of Arabic. It is thus an old pronunciation that was established in NS quite early. Interestingly, Ar. d never occurs as l in loanwords in Har. (Leslau 1956) and ḋAfar (Leslau 1995), nor in Swahili (Baldi 2008; Vasaturo 2012): its most frequent rendering is d, but dh and sometimes z in Swahili. Amharic has either of d, z or ṱ (Leslau 1957a, 1988), while Argobba has a few occurrences of d according to Leslau (1957b).

The isolated rendering as s in wéeso, a technical term of Islamic ritual, is also isolated in Har. wussa, where Ar. d generally gives d as in bidā’a ‘merchandise’ < bidā’a ‘id.’ ḋAfar has wadó with d. A possible explanation for this irregularity in NS and Har. is that it is a word that entered the Horn during the early diffusion of Islam, and then spread to Har. and NS through a third language where Ar. d > s was a regular development. Leslau (1963: 162) regards wussa as a loanword from Somali in Harari, without taking into account the fact that it is also irregular in Somali.

On the other hand, wádhaf is an old word in the Horn, as attested by the Geez verb waḏafa ~ waḏdafa ~ waṣafa ‘hurl with a sling’. Akkadian already had waṣpu > aspu ‘sling’, but the occurrence of both d, z and j in its Yemeni Ar. cognates is highly irregular. Har. has woč’āfa ‘id.’, Amharic.
wanč’əf ‘id.’ with a non-etymological -n- that also occurs in ḨAfar wandef ‘id.’ The irregular correspondences in dialectal Arabic and the isolated [d] in Somali and ḨAfar indicate that this word, even though it is likely to be of Semitic origin, had a complex history, and is not necessarily a loanword from regional Arabic into Ethiosemitic and East Cushitic. More likely, it re-entered Ar. from the Horn, and the occurrence of [d] in NS and ḨAfar has to be explained in the context of the bilateral history of East Cushitic and Ethiosemitic. Also NS bánθ ‘half’ is suspect as a loan from Ar. ba’d ‘part’, ‘some’, because ˈayn is usually retained in NS loanwords from Ar. It thus has probably a different etymon.

In some cases, one Ar. word has produced two different NS words. For instance, there are from Ar. zakā ‘zakat alms, tithe, the giving of 2.5% of one’s surplus wealth once a year to charity, generally to the poor and needy’:

- NS dágo f. 1. ‘the first month of the Muslim year’; 2. ‘domestic animals that are given away as alms once every year’.
- NS sáko ~ séko f. 1. ‘zakat alms’; 2. ‘the first month of the Muslim year’.

The first term develops Ar. z- and -k- in the same way as inherited East Cushitic *z and postvocalic *k, as seen above in § 1. It is not used as a technical term for the Islamic zakat. The second term devolves Ar. z into s and seems to preserve Ar. -k- after the first vowel. However, one should not forget that inherited East Cushitic long *-tt- and *-kk- occur as -t- and, respectively, -k- in NS. Since the Har. term for zakat is zakka, the occurrence of -k- in NS sáko is best explained as a direct loanword from Harari, that is, from the language of the town that has been the major centre of Islam in the eastern Horn for the last five centuries.

Ar. z merges with d, i.e., is treated like inherited East Cushitic *z, only in a few other words that belong to common everyday life, and that probably were borrowed quite early. Otherwise it is just devoiced.

Ar. z:

- NS d in:
  miúndul m. (> miuddul ~ muddülle) ‘round dwelling with a thatched conical roof’ (< manzil ~ manzal ‘house, dwelling’); etc.
- NS s as in:
  wasaarād f. ‘ministry’ (< wizāra ‘id.’);
  sâman m. (~ sêmen ~ sêben) ‘time’ (< zaman ‘id.’); etc.
That múndul is an old loanword is also indicated by the complete change of its vocalic pattern, from -a-a- into -u-u-, probably through labialization triggered by the initial m- and subsequent assimilation of the second vowel to the first one, in order to have full vowel harmony. A similar development can be observed in NS musqūl f. ‘washroom, toilet’ < Ar maḡsil ‘washroom, lavatory’, with the regular development Ar. g > NS q, metathesis -qs- > -sq-, and -a-a- > -u-u-. Ar. vowels are otherwise generally well preserved, even though there can be some changes in vowel length, usually with V > VV, while Ar. a-a sometimes becomes e-e.

Finally, Ar. words with a final consonant cluster are rendered in NS in two different ways:

- NS -CCi in:
  - digri m. ‘dhikr’ (< Ar. ḍikr ‘id.’);
  - shārci m. ‘law’ (< Ar. šarī ‘id.’);
  - dérbi m. ‘wall’ (< Yemeni Ar. darb ‘wall’, cf. Piamenta 1990–91: 147a); etc.

- NS -CVC with echo vowel insertion as in:
  - kidib m. ‘lie’ (< Ar. kiḏb);
  - méher m. ‘dowry’ (< Ar. mahr);
  - dúhur m. ‘noon’, ‘noon prayer’ (< Ar. zuhr); etc.

It is interesting to notice that echo vowel insertion is an active process in NS nominal and verbal phonology, e.g., īlig ‘tooth’ vs. ilk-ó ‘teeth’, ārag ‘see (sg.) it!’ and wāy arag-tay ‘she saw it’ vs. ārk-a ‘see (pl.) it!’ from the roots /ilk-/ and /ark-/. Instead adding a final -i is an active process in Har. phonology for avoiding final consonant clusters, as in the simple imperfect Har. 3sg.m. yilabsi /yilabs/ ‘he dresses’, Har. karsi ‘belly’ vs. Geez karš ‘id.’, Argobba and Gurage kārs ‘id.’, etc.

Accordingly, Ar. words in -CC are generally adapted with the -CCi pattern in Harari, e.g., zikri ‘dhikr’, Old Har. šarī > Modern Har. šarī ‘law’, maḥrī ‘dowry’, zuhri salāt ‘noon prayer’, etc.; occasionally they lose their second consonant as in kiz ‘lie’. A likely explanation of the occurrence of the two patterns in NS is thus that the kidib type is an internal development of NS, while the digri type spread into NS through Harari, possibly through words like zikri > digri and šarī > sharci whose pattern was extended analogically to other loanwords from Arabic.4

---

4. It is interesting that also Ar. loanwords in ʼAfar systematically display -i for avoiding final consonant clusters, e. g., sārī ‘doctrine, theology’, māhrī ‘alimony given to a divorced wife’, etc. Yet, differently from Har., ʼAfar phonology either degeminates final long consonants as in enged ‘deny (it)!’ (from
5. Ethiosemitic loanwords

Not much attention has been devoted to the bilateral contacts between Ethiosemitic and NS, even though several authors have addressed the issue of the impact of Cushitic, and thus also of Somali, upon Ethiosemitic, cf. recently Appleyard (2012), who also mentions the most recent literature on this topic. But borrowing from Ethiosemitic into NS has been discussed, at least to the present author’s knowledge, only by Mohamed Nuuh Ali (1985: 150 ff.) and Banti (1988: 57).

The former identifies some Har. loanwords in NS, but commits some astonishing oversights such as considering Northern and Benaadir Somali sitimāan ‘week’ as a borrowing from Har. sātti ‘seven, week’ with “added suffixation” (ibid.: 153), whereas its Italian origin (< settimana ‘week’) is obvious. Banti, on the other hand, points out that ugāas ‘tribal chief’ and agāas in m. ‘good management, good government, good management’ have Southern Ethiosemitic etymologies:

NS agāas-in m. ‘good management, orderly arrangement’, and ugāas m. ‘tribal chief’ are from the Ethiosemitic root GZ’: Geez gaz’a ‘dominate, master’ and əgzi’-ənnā ‘dominion, sovereignty’, Har. gaza’a ‘govern, own’, Gurage languages gaza ‘id.’ (Chaha and Endegebn gasa ‘id.’ with their regular devoicing of etymological z), Amh. gaza ‘id.’. The patterns uCaaC- and aCaaC- are typical of NS derivation from prefix-conjugated verbs, as shown in Banti (1988).

A few other Ethiosemitic loanwords can be pointed out here, in a list that is in no way exhaustive:

- NS maalin f. (pl. maalmó m.) ‘day’ < *maal-im cf. máan-ta ‘today’ with *-lt- -> -nt- partial assimilation (cf. Ar. al-yawm ‘today’, lit. ‘the day’), from Ethiosemitic: Geez. moʾālt ~ maʾalt ‘day’, Har. maʾaltu ~ māltu ‘day’ < W Ł ‘spend the day’. The suffix -t and -tu was lost also in several Southern Ethiosemitic palatalized forms like Har. mōy ‘day’, Zway māy-a ‘id.’, etc.; yet NS *maal-ta ‘the day, today’ can also be regarded as a reanalysis of a form like Har. māltu, because the f. article in NS is -ta, with Nominative case -tu.

/engedd/, cf. yengedde ‘he denied it’), or inserts an echo-vowel as in subaáhat ‘with butter’ (from /sabaáh + t/), cf. Parker and Hayward 1985: 216). Is there also here a Har. model, even though ʿAfar is not in direct contact with Har. today?
– NS badán ‘be much, be many’, bádi f. ‘being much, being many’, badí v2 ‘be much, be abundant, be strong in’, from Ethiosemitic: Geez bazḥa ~ bazha ‘be abundant, be much, be many’, Har. bazha ‘be numerous, be abundant’ and bizha ‘abundance’. Bazza ‘be abundant’ with loss of ḥ occurs in most Gurage languages and in Amharic.

Astonishingly, these are words from the basic lexicon, i.e., from the very core of the lexicon. For both of them, the donor language was most likely a Southern Ethiosemitic language. A few other Ethiosemitic loanwords in NS are listed below, five of them most likely from Har., the last one from Amharic during the last decades:

– NS géyi m. (~ gáyi ~ gíyi) ‘land, country’, from Southern Ethiosemitic: Har, gê ‘Harar’ (lit. ‘the Country’), several Gurage languages ge ‘country village’ (with possible parallels in other Semitic languages).

– NS gówrac v. and n.m. ‘slaughtering an animal by slitting its throat’ from Southern Ethiosem.: Har. gōra’a v. ‘id.’, several Gurage languages gorā v. ‘id.’, Geez gʷarəya ‘id.’ denominative verb from Geez gʷər’e ‘throat, neck’ with several other Semitic cognates. The development of ô after the first root consonant only occurs in Southern Ethiosemitic.

– NS arabīkhī f. ‘maize, corn’ from Har. arab ihi ‘sorghum of the Arabs’, with Har. loss of ʾayn in arab < ʿarab (cf. NS Caráb-ta ‘the Arabs’) and the typically Har. inverted genitive construction.

– NS áw m. religious title, similar to shêekh, from Har. aw ‘father, elder, religious title’ (already mentioned above in § 2).

– NS malāakh m. ‘tribal chief’ (already mentioned above in § 2) from Har. malāq ‘official who is responsible for the welfare of each locality connected with one of the five gates of the city of Harar’, from Har. mala ‘ways and means’ + āqa ‘know’ (cf. Leslau 1963: 107b).

– NS kilil m. ‘region’ in Kilil Shán ‘Region 5’, i.e, the Somali region of Ethiopia, from Amharic kollol ‘zone’.

Notice that NS garāad ‘tribal chief’, which was mentioned above in § 2, does occur also in Har. with the same meaning, but doesn’t have a clear etymology. As noted by Gori (2005) it already occurs as a title of Muslim rulers in medieval Abyssinian sources, and has been used subsequently also by some Christian rulers. It has thus to be regarded as a title that has been
used since the Middle Ages in Muslim polities of the Horn, whatever its linguistic origin, but there is not enough evidence for regarding it as a specifically Ethiosemitic loanword. Indeed, Leslau (1963: 75a) regards it as borrowed from Cushitic into Harari, but also for this there is very little evidence.
Figure 1. Penetration routes of Islam into the Horn and Ethiopia (from Fauvelle-Aymar and Hirsch 2011: 20.)
References

Abdirachid Mohamed Ismaïl

Agostini, Francesco, Annarita Puglielli, and Moxamed Siyaad Ciise (eds.)

Appleyard, David

Axmed Cartan Xaange

Baldi, Sergio

Banti, Giorgio
Banti, Giorgio, and Riccardo Contini

Beeston, A. F. L., Ghul, M. A., Müller, W. W., and J. Ryckmans

Behnstedt, Peter, and Manfred Woidich
2011 Wortatlas der arabischen Dialekte. Band II: Materielle Kultur. (Handbook of Oriental Studies, Section 1 The Near and Middle East 100/2.) Leiden/Boston: Brill.

Cabdalla Cumar Mansuur

Callegari, Laura

Cardona, Giorgio R.

Casson, Lionel

Cerulli, Enrico
1936 Studi etiopici I. La lingua e la storia di Harar. Rome: Istituto per l’Oriente.

Chekroun, Amélie, François-Xavier Fauvelle-Aymar, Bertrand Hirsch, Hailu Zeleke Deresse Ayenachew, Olivier Onezime, and Shewangizaw Addisu
Ehret, Christopher, and Mohamed N. Ali

Fauvelle-Aymar, François-Xavier, and Bertrand Hirsch
2008 Etablissements et formations politiques musulmans d’Éthiopie et de la corne de l’Afrique au Moyen Age: vers une reconstruction. Annales Islamologiques de l’Institut Français d’Archéologie Orien
tale 42: 339–75.

Gori, Alessandro

Heine, Bernd

Lamberti, Marcello
1986 Die Somali-Dialekte. (Kuschitische Sprachstudien 5.) Hamburg: Buske.

Leslau, Wolf

Mohamed Nuuh Ali

Parker, Enid M., and Richard J. Hayward

Piamenta, Moshe

Reinisch, Leo

Ricks, Stephen D.

Robin, Christian

Saalax Xaashi Carab

Sasse, Hans-Jürgen

Simeone-Senelle, Marie-Claude

Soravia, Giulio

Stein, Peter


Tosco, Mauro
Vasaturo, Laura  
2012  I prestiti arabi in somalo e swahili. Dissertation, Facoltà di Lettere e Filosofia, Università degli Studi di Napoli “L’Orientale”.

Wagner, Ewald  

Yaasiin Cismaan Keenadiid  
1976  *Qaamuuska Af-soomaaliga*. Mogadishu: Guddiga Af-soomaaliga/Akademiya Dhaqanka/Wasaaradda Hiddaha iyo Tacliinta Sare.

Zaborski, Andrzej  