John Hajek and Valentina De Iacovo*

Lexical contact in Africa: Italian loanwords in the Mà'dí language of South Sudan and Uganda

Abstract: We examine the presence of Italian loanwords in Mà'dí (ISO 630 code [mhi]), a Moru-Madi language spoken in South Sudan and Uganda. The existence of Italian loanwords elsewhere in Africa is the result primarily of colonial contact, and/ or the presence of communities of Italian-speakers originally from Italy. We show that the Italian loanwords in Mà'dí have two sources of transmission: (a) direct contact with Italian Catholic missionaries; (b) indirect mediation via Arabic (mostly via a creole form of it, known as Juba Arabic). We argue that what are in some cases loanwords borrowed from Arabic are in fact ultimately indirect loans from Italian, having first entered into Arabic. Other loanwords, presumed to be Anglicisms, may be or are also Italian in origin or show some evidence of contact with Italian. The situation is rendered more complicated by the presence of borrowings that may instead be Latin or even Portuguese in origin – the latter potentially having entered Mà'dí indirectly through Swahili.

1 Introduction

1.1 What is Mà'dí?

A member of the Nilo-Saharan language phylum, Mà'dí (the apostrophe indicates that the following consonant is implosive in articulation) belongs to the Moru-Madi group which falls within the Eastern group of the Central Sudanic languages, spoken mainly across an area of Central Africa that touches upon Uganda (Mà'dí, Lugbara),

---

*Corresponding author: Valentina De Iacovo, Department of Foreign Languages, Literatures and Modern Cultures, University of Turin, 10124, Italy, E-mail: valentina.deiacovo@unito.it. https://orcid.org/0000-0002-3913-2727
John Hajek, School of Languages and Linguistics, The University of Melbourne, Melbourne, 3010, Australia, E-mail: j.hajek@unimelb.edu.au. https://orcid.org/0000-0003-1335-4273

South Sudan (Mà'dí, Moru, Lugbara, Kaliko, Avokaya, Lolubo, Logo) and the Democratic Republic of Congo (Logo, Lugbara, Kaliko, Avukaya). According to Blackings and Fabb (2003), the specifically Mà'dí-speaking population straddles Uganda (est. 250,000 people) and South Sudan (est. 50,000). The Ugandan Mà'dí live along both sides of the Nile while the South Sudanese Mà'dí reside along the eastern bank of the Nile only – possibly as a result of a relocation ordered at one point by the British (Molloy 1957). The Mà'dí-speaking area is also known to be ethnolinguistically very complex: in and around these territories one finds the Bari and the Lolubo (North), the South Sudanese Acholi and the Latuka (East), the Ugandan Acholi (South) and the Alur, Lugbara, the Kuku and the Kakwa (West). The predominant economic activity of the Mà'dí is agriculture and, in addition to food crops, cotton and tobacco cultivation is common in Uganda and South Sudan respectively – alongside traditional occupations of fishing and hunting (Blackings and Fabb 2003). While the term Mà'dí is used in English with reference to both the language and its speakers, Blackings and Fabb (2003) report that the speakers refer to their language as *màdí̵ tī*, literally "Mà'dí mouth" while Tucker and Bryan (1966: 36) suggest that *ma-* is a 'tribal prefix' used in some languages in the region to name either the "tribe" itself or another "tribe".

Linguistic interactions involving the Mà'dí have been strongly influenced by the complex recent history of Sudan of which South Sudan was a part until it achieved full independence on July 9th 2011 (see below for more detail). As Blackings and Fabb (2003) note, for example, in part due to the first civil war in Sudan (1955–72) when many South Sudanese were relocated among the Acholi population in Uganda, most Mà'dí speak Acholi but very few Acholi speak Mà'dí. At the same time, for reasons not entirely clear, most Kukus in Uganda speak fluent Mà'dí, but barely any Mà'dí speak Kuku. It is also still possible even today, as Blackings and Fabb (2003: 4) report, "to find among the [South] Sudanese Mà'dí people who can trace their ancestry to the neighbouring tribes – Bari, Kuku, Pajulu, Acholi etc… Hardly any of them can now speak their 'ancestral' languages; they speak Mà'dí only and have become fully absorbed into the Mà'dí community".

The territories where these languages are spoken in South Sudan and Uganda are characterised by a strong multilingual component (e.g. Amone 2021). However, the most widely used lingua franca in South Sudan is Juba Arabic, a creole variety of Arabic that first developed in the South Sudanese region of Equatoria, and in particular in Juba, its capital (for a discussion on the sociolinguistic dynamics of this variety, see, e.g. Tosco and Manfredi 2018, and for a grammatical overview see Manfredi and Petrollino 2013; Manfredi 2017). Equatoria is also the area where the Moru-Madi languages are concentrated in South Sudan. However, along and across the neighbouring borders of Uganda, Swahili and English are more widely spoken than Juba Arabic. It is important to note that knowledge of Swahili is common in Uganda, especially in the North, where it is spoken by speakers of a range of non-

Bantu languages, and is the preferred local lingua franca (Blackings 2011: iii). Further to the South in Uganda, several Bantu languages are also used as languages of wider communication (see Boone and Richards 1996).

The multilingual setting also includes local variants of Mà'dí (in the towns of Wau and Malakal for example), while the Juba Arabic spoken by the Mà'dí is also influenced by another creole variety of Arabic, Nubi/Kinubi, spoken in Uganda among Moslems who are reported to be mainly descendants of British General Gordon's troops who settled there after his Sudanese campaign (Blackings and Fabb 2003). It is also important to emphasise the potential for linguistic change caused by the various population displacements in both directions between South Sudan and Uganda during the years of civil war in Sudan (see below) and periods of civil unrest in northern Uganda. As a result, some Ugandan Mà'dí have learned Juba Arabic, while the South Sudanese Mà'dí have in turn also been exposed to Swahili (Blackings and Fabb 2003).

1.2 Materials in and on Mà'dí

Being predominantly an oral language, there is only limited written production in the Mà'dí language. As reported by Blackings and Fabb (2003), published works in Mà'dí concern missionary publications such as the translation of the New Testament, and prayer and song booklets by the Catholic missionaries. In their grammar, Blackings and Fabb (2003) also include texts from the Mà'dí Ethnic and Heritage Welfare Association in Britain which published a quarterly bilingual (English and Mà'dí) paper called *Mà'dí Lelego*. A more positive situation, however, concerns radio where Mà'dí is used in broadcasting in Uganda and in South Sudan (Abdelhay et al. 2015).

Early descriptive studies on Mà'dí language are by Molinaro (1925) and Tucker (1940) and later Bryan and Tucker and Bryan (1966) who include a chapter on Mà'dí. There are also some works on language varieties related to Mà'dí such as Logbara (Crazzolara 1960) and Ngiti (Kutsch Lojenga 1994).

With respect to dictionaries on Mà'dí, these include Bilbao and Moizi (1984), Bilbao and Ferrazin (n.d.) and Ferrazin (1996) – none of which we have been able to locate, with the last two identified by Blackings and Fabb (2003). The only lexicon we have access to and on which we focus our attention is the *Mà'dí-English Dictionary* compiled by Mairi Blackings. Published first in 2000, a second expanded edition appeared in 2011. It contains some 5,000 items in the main Mà'dí-English section, in addition to a shorter English-Mà'dí finder list (~2000 items). Given it is our primary source, we provide additional details about the dictionary further below and also discuss our methodological approach with respect to data extraction and evaluation.

1.3 A brief sociopolitical and linguistic history of (South) Sudan in the 19th and 20th century

Before becoming independent in 2011, South Sudan had for some time been a part of Sudan by which name the entire entity including the South was usually referred to. In the early 19th century Egypt had begun to expand territorial control over Sudan and in 1821 the Egyptian pasha Ismāʿīl established the colony of Equatoria in the southernmost part of what is now South Sudan. With the declining fortunes of Egypt and the rise of political conflict, in 1882 Egypt sought assistance from Britain, which legitimised its military presence in Sudan by acquiring significant political power. The situation was subsequently destabilised by Muhammad Ahmad the self-styled *Mahdi*, who led an Islamic revolt throughout Sudan, leading ultimately to the capture of Sudan's main city, Khartoum, in 1885. The revolt was eventually put down by the British and their allies, and in 1899 Sudan became an Anglo-Egyptian Condominium – notionally under shared jurisdiction but under actual *de facto* British control until Sudanese independence in 1956. During much of this period the British made efforts to limit Arab influence in South Sudan, simultaneously favouring the spread of English and Christianity. This situation was further encouraged by the so-called Southern Policy (1930) formulated by the British authorities, which preferred to administer the northern (Arabic- and Islam-dominant) and southern (non-Arabic and non-Islamic) parts of Sudan separately, until their administrative reunification was decided at the Juba Conference in 1947 (Okeny 1991). While Sudan achieved independence in 1956, the disparities between the two territories (with the North heavily Arabised and Islamicised as well as more economically developed) and efforts by the North to impose its linguistic and religious practices on the South led to two protracted civil wars (1955–1972 and 1983–2005) with over 2 million dead, until the successful referendum for South Sudanese independence in 2011. This complex political history has also had consequences at a linguistic level and as a result of their long-term presence in one way or another in Sudan and more broadly in the region, English and Arabic have left substantial traces in the form of loanwords: in Sudanese Mà'dí they are mostly of English and/or Arabic (or more specifically of Juba Arabic origin), while in Ugandan Mà'dí they are mainly of English and/or Swahili origin (Blackings 2011). We return to discussion of English and Arabic loans below.

Abu-Manga (1991) notes with respect to the contact between Arabic and the languages of South Sudan, that it was not so much the religious sphere that has had an impact as the commercial one. During the Turkish-Egyptian administration (1821–85) Arab merchants involved in trade in the South favoured the use of a pidgin Arabic and for the first time too Arabic served as the language of official correspondence (Elrayah 1999). English came into the area instead through the arrival

first of British functionaries and then of British missionaries some time after General Gordon's time in the employ of Egyptian authorities. The missionaries were invited by the Church Missionary Society in 1871 to establish their activities there (Elrayah 1999). As a result, English was firstly limited to government work and mission schools, with colloquial Arabic present in more informal contexts during the Anglo-Egyptian Condominium phase (1899–1956). English was eventually adopted in 1929 as the official language of government and of school instruction in the South. During this period of British rule of Sudan (including South Sudan), the use of Arabic (together with propagation of the Islamic religion) was seen as a dangerous unifying factor so efforts were made, as noted above, to limits its use in the South (with the exception of the Juba area) where it was only reintroduced in significant fashion with Sudan's independence in 1956. From then on, Arabic increased its role alongside English, which in 1972 was designated to be the language of administration and education in the South while Arabic was declared in the constitution as the official language of the Sudanese Republic (Elrayah 1999). Subsequently, the role of Arabic was expanded at the expense of English in the South as part of significant processes of Arabization and Islamization, triggering strong local resistance over many years (Abdelhay et al. 2015, 2016; Musgrave and Hajek 2014).

The school system in South Sudan was at different times before independence primarily either English- or Arabic-medium (with Arabic strongly favoured at the expense of English by the Arabic-speaking Sudanese government based in Sudan's national capital Khartoum). Since independence in 2010, these processes have been completely reversed: English is now the only official language, as well as the primary language of instruction (UNICEF 2016). In primary school, teaching in the local languages also often occurs (including Mà'dí in the areas where it is spoken), while in Juba itself local Juba Arabic may also be used. In Uganda, on the other hand, English is predominant, while in some contexts Mà'dí is also used in school (see Boone and Richards 1996; Heugh and Mathias Mulumba 2014).

1.4 The presence of Italian(s) in Sudan: the role of the Comboni Missionaries

A further piece must be added to the complex linguistic situation described above, which relates to the presence in Sudan of the Italian Catholic order of Comboni Missionaries of the Heart of Jesus. Its founder, Daniele Comboni, who left for his first mission to Africa in 1857, is considered by many an important symbol of Catholic evangelisation in the Sudanese territories and eventually across large parts of

Central and Southern African continent.[1] Through his "Plan for the Regeneration of Africa", he devoted himself to the establishment of schools and the training of teachers and doctors in Africa. In 1867 he founded in Verona the Institute for the Mission of Africa and a few years later he petitioned for the evangelisation of Central Africa, which enabled him to carry out further missions in Africa and to found schools in Sudan (Chemello 2021). After his death (1881), the fate of the *Verona fathers* (as the Comboni missionaries are often called in English) depended on the political events that took place in those territories, while managing to maintain an important role of reference for Catholicism (which has since become one of the major religions practised along with Protestantism and to a lesser degree Islam) in South Sudan as well as in Uganda. Available information (e.g. Agostoni 1996; Chemello 2021) indicates that over time most Comboni priests were and continue to be from northern Italy, and in particular from the Veneto region in which the city of Verona is located. We also know the Comboni priests lived and worked very closely with different linguistic communities in South Sudan over extended periods, intensifying linguistic contact as a result. A number of them were also actively involved in preparing detailed grammatical descriptions and dictionaries of South Sudanese languages over a period of decades, e.g. Father Pasquale Crazzolara wrote a grammar of Nuer in 1933, followed by a series of other linguistic works, including a description of the Pokot language in 1978.

As we will demonstrate below, it is no surprise therefore if in this particular missionary context, the influence of Italian in the languages spoken in these territories (including Mà'dí) most typically concerns the religious sphere, as Blackings (2011: iii) also notes. That the presence of Italian missionaries, specifically of the Verona fathers, has led to the borrowing of Italian words in South Sudanese languages has also been noted by Storch (2014), albeit in only brief terms, for Luwo, a Nilotic language spoken some way to the North of Mà'dí.

2 Understanding borrowing in Mà'dí

2.1 The challenge of parallel pathways, convergence, mediation and multiple source languages

There are significant challenges for the correct identification of the source language for borrowings in Mà'dí. In the first instance, there may be mediation of a loanword through another language, masking the original source language, e.g. a loan may be

---

1 The first Catholic school in Khartoum was in fact opened by an Italian priest called Luigi Montuori in 1843 (Alcaide 2019).

borrowed directly from Arabic but the word is ultimately Italian in origin through indirect transmission (see below).

In other cases, more than one source language can potentially be identified. In the case of words listed by Blackings as Swahili in origin, he notes that they may as well have been marked as Arabic given that they mostly have their origins in that language (p.iii). However, Swahili has also borrowed extensively from English and from Portuguese (the latter the result of the strong Portuguese presence along the eastern coast of Africa from the 1500s to the 1700s. See, e.g. Baldi 2018, 2024). It is also a well-known donor language to other languages in Africa (Baldi 2011). Given the shared genetic origins as Romance languages and therefore the often similar nature of Italian and Portuguese lexical items, it can be difficult to separate the two when considering the ultimate origins of loans in Mà'dí. The large Romance lexical component (mostly from Latin and French) in English further complicates matters, since English words may also appear very similar to their cognates in Italian and/or Portuguese. As a result, a specific word could come from any one of those languages or even if borrowed from one specific language, it can also show evidence of overlap and/or mixing in phonetic form in Mà'dí as a result of complex language contact patterns in that language.

Finally, we have to add one more language, Latin, to the complex mix. It remains the official language of the Roman Catholic Church and, until the 1960s, it was the sole liturgical language of the Catholic mass, hence during Catholic missionization in South Sudan, as elsewhere in the world, through the 19th century and until well into the second half of the 20th century, local populations would have had some contact with Latin as well.

2.2 Italian loanwords in the Horn of Africa and North Africa

As a European nation, Italy's colonial presence in Africa was relatively late and somewhat limited. By the end of the 19th century it had established itself in Eritrea and part of Somalia (Italian Somaliland). As Tosco (2023) points out in the case of Somalia and Eritrea, given Italy's limited provision of education during its period of colonial control in the Horn of Africa, contact with Italian in these parts was generally more informal in nature. However, this did not prevent some terms belonging to mostly specialised lexicons (cuisine, tools, furniture, household, vehicles) from spreading in languages of the Horn, e.g. in Saho spoken in Eritrea, Ethiopia and Djibouti we find *biyatti* (It. piatto) 'dish', *kashshabiito* (It. cacciavite) 'screwdriver'; Amharic: *kanatera* or *kanitera* (It. canottiera) 'undershirt', and *färäfango* (It. parafango) 'mudguard/fender' (Banti and Vergari 2008). The introduction of Italian loanwords in North Africa, particularly in Libya, Tunisia, and Egypt was

largely the result of colonial contact in the case of Libya, and the presence of a large resident Italian population and a strong commercial presence in all three countries from the mid-19th century on.

It is within this context, that we can identify a small number of studies on Italian loanwords in the languages spoken in Africa such as Somali (Mioni 1988), and Eritrean (Tosco 2008), as well as Arabic as spoken in the Dahlak islands of Eritrea (Simeone-Senelle 2018), in Tunisia (Cifoletti 1994), Libya (Abdu 1988; D'Anna 2018), and Egypt (Moustafa 2019-20). These sources are useful when considering the possible Italian origins of loanwords in Mà'dí. Moreover, as noted in some of these studies, the spread of Italian loans was sometimes mediated indirectly by the other languages spoken in those areas, including not just Arabic, for instance, but also dominant local languages such as Amharic in Ethiopia (cf. Tosco 2009 on borrowing in the Ethiopian language, Gawwada).

3 Italian loanwords

3.1 Methodology

As noted above, for this study we rely on Blackings' (2011) *Madi-English / English-Madi Dictionary* as our primary data source. It consists of a short introductory section (I-VII) explaining the organisation of the entries, the language's phonological system, suggested orthographic norms orthographical norms, and a brief reference to the grammar found in the dictionary. This section is then followed by a full listing of entries in Madi-English (pp.1-191) as well as a shorter and less detailed English-Madi wordfinder (pp.192-232).

An example of an entry in the Madi-English section is as follows:

<blockquote>

aduresi [àdùrēsì] *n.* address ENG.

ALSO **àdùrésì**, **àdērēsi**

</blockquote>

The headword in bold is in what Blacking (2011) refers to as conventional Mà'dí orthography while the form in square brackets is in the new Mà'dí spelling system that he proposes. This revised spelling system, which is practically phonetic, allows for both vowel quality (the distinction between the advanced and retracted tongue root articulations) and tone (high, low, mid and contour) to be indicated. Alternative forms which follow the word ALSO are given in combined form, i.e. bold with tone and vowel quality markings. Each entry also includes the word class and the English gloss. In some cases, derived or inflected words, idioms and example sentences are

also provided. Information about alternative spellings (and pronunciations) and whether the word is a loan from another language (e.g. English) is usually also given (although not all loans are necessarily identified as such). As Blackings (2011: 233) notes, the source languages for loanwords are most commonly Arabic (in South Sudan, and which he indicates elsewhere to be specifically Juba Arabic), English (Uganda and South Sudan), and Swahili (mostly Uganda) and, more rarely, Italian (mostly religious register) – all pointing to the complex socio-cultural, linguistic and political contacts of the Mà'dí people.

We first note that loans are more typically adapted from the oral rather than the written form: this is particularly clear in words from English such as *namba* [námbà] (number), *sikuta* [sɪ̀kútà] (scooter). which show regular syllable-final r-loss typical of spoken British English which is used in Sudan, and South Sudan as well as throughout East Africa. It is also important to observe the regular relationship between tone in loanwords in Mà'dí and stress placement in the source language: the non-low tone (typically the high tone, but also the mid tone when high is not present) generally coincidences with primary stress placement in the source language, as seen in the case of *sikuta* [sɪ̀kútà] (scooter), with the high tone on the [u] which is stressed in English. Similarly, the mid tone in [àdùrēsì], and the high tone in the alternative [àdùrésì] are consistent with stress on the final syllable in English *address.*[2]

In our analysis of loanword entries, we discuss them in the following order: (1) the entries that Blackings (2011) specifically identifies as loans from Italian; (2) the entries that the author attributes to English (2a) or Arabic (2b) but for which we need to consider an Italian origin; (3) the entries that have two outcomes, one of which could be from Italian but which the author does not indicate; and (4) the entries whose language of origin is not given but which may be Italian in origin. Finally, for completeness, we also consider briefly a small set of items that are borrowed from Latin, rather than from Italian or Portuguese.

After listing and then discussing the various entries, we then discuss overall patterns, before considering some of the phonetic and phonological processes as well as semantic categories that can be identified in the borrowed forms. To facilitate our analysis, we mark stress on all Italian orthographic forms, although this is not usually done in Italian spelling. Since Italian spelling is highly phonetically transparent, we do not provide phonetic transcription, except in a few specific cases. We sometimes also mark stress placement in the spelling of words from other languages where this is particularly helpful to our analysis.

---

2 Ironically, stress placement in English 'address' may be on the initial or final syllable in the noun (the first author accepts both in this case, while it is on the final in verb). It is clear that the stress-final variant was borrowed in Mà'dí (where it is now on the non-final penult due to final vowel epenthesis), since the high(er) tone is never on the initial syllable.

3.2 Corpus analysis

A small number of items correctly identified by Blackings (2011) as direct Italian loans are listed in (1):

(1)   **asunda** [àsṵ́ndà] / **asunta** [àsṵ́ntà] < *assúnta* (Assumption)
        duluvio [dṵ̀lṵ́vḭ̄ọ̀] < *dilúvio* (deluge, flood)
        faratelo [fàràtẹ́lọ̀] or [fārātẹ́lọ̀] < *fratéllo* (a Brother, in the Catholic Church)
        nyatali [nyàtàlì] / **natali** [nàtàlì] < *Natále* (Christmas)
        papa [pápà] < *pápa* (pope**)**
        suere [sṵ́ẹ̄rẹ̀] / **suwere** [sṵ́wẹ̄rẹ̀] < *suóra* (nun, Sister)

As is immediately obvious, all six items have a religious connection. The doublet **asunda/ asunta** is from *la Festa di Maria Assunta* (the feast of the Assumption of Mary), a major Catholic feast day. While an Italian origin is sufficient for **papa**, it is also possible it is from Latin *papa* or is a Portuguese loan that has entered via Swahili which both also have *papa*. The term *papa* is commonly used throughout central (e.g. Kinyarwanda) and southern Africa (e.g. Zulu *upapa*). On the other hand, such an explanation more likely reflects coincidence and reinforcement through converging loan pathways – given the otherwise consistent use in Mà'dí of Italian forms such as *fratello* and *suora,* to describe other basic religious roles in the Catholic church in the items above.

In the case of **duluvio** [dṵ̀lṵ́vḭ̄ọ̀], the traditional and new spellings do not quite match. It is not clear if this discrepancy is due to an error in the former, or if the latter reflects a more recent change in pronunciation. The new spelling appears in both parts of the dictionary. There is a third form [dèlūjḭ̀ọ̀] which appears to be a cross between English *déluge* and Italian *dilúvio*, a phenomenon which highlights the possibility of multiple language contact effects in Mà'dí. In addition, it is worth noting the absence of a high tone in this word, with mid tone on what would be the stressed vowel in Italian.

In the case of **suere** [sṵ́ẹ̄rẹ̀] / **suwere** [sṵ́wẹ̄rẹ̀], this appears to be from the plural form in Italian, i.e. *suóre* (nuns, sisters), since the change from -a to -e would not otherwise be expected. The use of the Italian plural form as the borrowed source also occurs in a small number of other cases, as discussed below.

An additional item is also identified in the dictionary as Italian in origin:

        peregelemu [pẹ̄rẹ̄gẹ̀lẹ́mù] (riddle) < *probléma* (problem)

However, the Italian origins of this word are less clear to us, albeit not impossible, given the significant phonetic and phonological distance between the Italian and Mà'dí forms. On the one hand, the high tone on the penult is consistent with primary stress placement in Italian *probléma* (but not English *próblem*). But the appearance, amongst other things, of final -*u* rather than (Italian) final -*a* is unexpected, and is not

consistent with other Italianisms where the original final vowel is normally maintained. The Mà'dí form is in our view more likely to be from English *problem* with vowel harmony and the predictable addition of final -u after a bilabial consonant, cf. *bomb* > **bomu**. What is unexpected is the high tone on the new penult – which is usually indicative of the stressed syllable in the source language. In English the stress is on the initial syllable. This type of stress shift in Mà'dí occurs only very rarely, and may in this case be due to contamination with Italian *probléma*.

A number of items identified as Anglicisms by Blackings (2011) listed in (2a) appear instead in our view to most likely be direct loans from Italian. In the list we have immediately suggested the Italian source, alongside the meaning of each item in English, before providing more detailed explanation further below.

(2a) **alitari** [àlìtárì] also [àlìtárè] < *altáre* (altar)
　　　anjeli [ánjḛ̄lḭ̄] < *ángelo* (angel)
　　　azima [ázìmà] < *ásma* (asthma)
　　　batisimu [bàtḭ́sḭ̄mù] < ?*battésimo* (baptism)[3]
　　　girasia [gìrásīà] < *grázia* (grace)
　　　kandela [kàndélà] < *candéla* (candle)
　　　kina [kḭ́nà] < *china* (quinine)
　　　kolera [kòlérà] < *coléra* (cholera)
　　　kuruce [kùrúcè] < *króce* (cross, crucifix)
　　　lamba [lámbà] / [lāmbáà] < *lámpa* (lamp)
　　　matirimonio [matḭ̀rḭ̀mónḭ̄ọ̄] < *matrimónio* (matrimony)
　　　paradizo [pàràdḭ́zò] < *paradíso* (paradise)
　　　penitesia [pēnìtésīà] < *peniténza* (sacrament of penance)
　　　purofeta [pùròfétà] < *proféta* (prophet)
　　　ruzario [rùzárḭ̄ọ̀] < *rosário* (rosary)
　　　seminario [sḛ̀mḭ̀nárḭ̄ọ̀] < *seminário* (seminary)
　　　semiterio [sḛ̄mḭ̄térḭ̄ọ̀] < ?*cimitéro/cimitério* (cemetery)

We see in this set mostly terms relevant to religious practices and life:

> **alitari** [àlìtárì] and also [àlìtárè]. It is highly unlikely this is from (British) English *alter* for which **alita** or even **olita** would be expected, given both the pre-lateral back and rounded vowel quality and regular loss of the final rhotic in that variety of English (see also **bia** 'beer' below). The alternate form [àlìtárè] with final -e is also entirely consistent with the original Italian form. However, it is difficult to discern the specific origin of this word since there may be up to three sources: Italian and Latin *altare* and Swahili *altare* (and also *alitari*, M. Tosco, pers. comm). It is not clear if the loan into Swahili is from Latin *altare* or Portuguese *altar*. However, Baldi (2024: 100) reports the borrowing of Portuguese *altar* in various languages of Sub-Saharan

---

3 Where an Italian origin is potentially less likely, this is marked with ? before the word.

Africa, including in Macua (spoken in northern Mozambique) which present multiple forms, two of which are identical to those in Mà'dí, i.e. *alitari, alitare, altari* and *altare*. That said, given the important religious nature of the term, Latin as a source is also plausible in the cases identified by Baldi (2024) as well and cannot *a priori* be excluded.

anjeli [ánjēl̤ī]. It is difficult to be certain if this is a loan from English *angel* or from Italian *angeli* (lit. angels, plural) or even a cross-over form from both. If English were the source, we would expect **enjeli** with vowel raising as more likely, cf. **beteri** < *battery*. If it is from Italian, it appears to be based on the plural form in that language (cf. *angelo* in the singular). Although this is unusual, we suggest there is a parallel with **suwere** 'sister' above and other forms such as **kiristiani** 'Christian' discussed below.

azima [ázìmà] appears to be a direct loan from Italian *asma* ['azma] (asthma). Retention of the voiced [z] points to clear Italian origin, whereas the English *asthma* has only voiceless [s]: **esima** in light of **beteri** may also be expected if it were to be borrowed from English.

batisimu [bàt̤ís̤ī̤mù] may be from Italian *battésimo* (baptism) where the stressed vowel matches the high tone in Mà'dí. Portuguese *batísmo* (also *baptísmo*) is also a possible source. However, in both cases, the voiceless [s] is unexpected, where [z] would be more likely, given sibilant voicing before /m/ as well as between vowels in both languages. Portuguese origin is problematic all the same given that the form in the expected mediating language, Swahili, is *ubatizo* based on the Portuguese verb *batizar* rather than on the original nominal form. The form **batisimu** or similar is also found elsewhere in Central Africa, including Lingala *batisimo* and Kinyarwanda *batisimu* alongside *umubatizo*.[4] It is not likely from English *baptism* in which primary stress is on the initial syllable and for which we would expect the [p] to be retained in some way, e.g. ****bapitisimu**. Similarly, Latin *baptismus* can also be excluded here.

girasia [gìrásìà] is from Italian *grázia* ['grattsja] (grace) and/or Latin *gratia* which in Italian-style pronunciation is pronounced similarly. We note however that the reduction of affricate /ts/ to voiceless [s] can be explained in two ways: (a) it is a common process in northern Italo-Romance languages and is very typical of the Veneto language of the Verona fathers, e.g. ['grasja]; and (b) in light of the absence of native affricates /ts, dz/ in Mà'dí. In Mà'dí the glide [j] becomes a fully syllabic vowel. Portuguese *graça* is not a source since ****girasa** would be expected, and *graça* it is not in any case borrowed into Swahili which instead has *neema*.

kandela matches Italian *candéla* (candle), with no corresponding competitor in Latin (*lucerna*), Portuguese (*vela*) or Swahili (*mshumaa*). Candles play an important role in Catholic religious practices.

kina [k̤ínà] matches perfectly Italian *chína* [kina] "quinine".[5] The Italian form is itself a borrowing from Spanish *quina* (ultimately from Quechua *kina* 'tree-bark'). The word has entered Mà'dí most likely through Arabic. Baldi (2021) lists *kina* (alongside *kinin*) as Arabic loans

---

4 Recall that the Verona Fathers operate across Central and Southern Africa, such that borrowing from Italian in the religious sphere in particular is likely to be evident well beyond South Sudan and Uganda.

5 Today the more usual forms in Italian are *chinina* and *chinino*.

in other South Sudanese languages, such as Acholi, Bari and Lotuxo, but he fails to identify it as a loan into Arabic. Elsewhere, it is clearly identified as an Italianism in Palestinian and Jordanian Arabic (Butros 1973).

kolera [kòlérà] < Italian *coléra* (cholera). Here the stressed syllable in Italian matches the high tone in Mà'dí. English and Portuguese have stress on the word-initial syllable, while the form in Swahili (*kipindupindu*) is unrelated. The form *kole:ra* is recognised explicitly to be a loan from Italian in Palestinian and Jordanian Arabic (Butros 1973), while *kolera* is also found in Juba Arabic.

kuruce [kùrúcè] < Italian *cróce* /krotʃe/ – with matching stressed syllable and orthographic and phonetic correspondence between Mà'dí *c* /tʃ/ and Italian *c* /tʃ/ in this case.

lamba [lámbà]/[lāmbáà] is not from English but is instead borrowed from Arabic *lamba* (Wehr 1994), cf. also *laamba* in Libyan Arabic (Abdu 1988) and *lampa* in Chadian Arabic (Jullien de Pommerol 1999). It is ultimately from Italian *lampa* which is today little used in that language where it has now been replaced by *lampada.* The alternative form [lāmbáà] with high tone on the first post-cluster vowel is unexpected. The Italian loan **lamba** is also widely attested across the Horn of Africa (M. Tosco, pers. comm.).

matirimonio [maṭìrìmónĭó̧] is a clear loan from Italian *matrimónio* (mátrimony, marriage), supported by the close phonological matching including correspondence between primary stress in Italian and high tone in Mà'dí. While Portuguese also has *matrimónio,* it can be excluded as Swahili has *ndoa*. Matrimony is of course a major religious sacrament in Catholicism.

paradizo [pàràdízò] is a perfect match with Italian *paradíso* [para'dizo] (paradise) – with an identical number of syllables, and vowels as well as retention of [z] – the result of regular intervocalic voicing of /s/ in northern Italy from where the Combonian order originates. Portuguese *paraíso* is easily excluded given the deleted obstruent.

penitesia [pēnìtésīà] is likely from Italian *peniténza* [peni'tentsa] (sacrament of penance) but shows the influence of Latin *poenitentia* which in traditional Italian pronunciation of that language is [peni'tentsja] with a glide. Reduction of [tsja] to [sja], e.g. [peni'tensja] is also typical of many northern Italian languages, including in Veneto. Portuguese *penitência* also has a post-sibilant glide, but Swahili does not appear to have borrowed this word. It has *kitubio* instead.

purofeta [pùròfétà] is potentially from Italian *proféta* (prophet), or Latin *prophéta* which are in turn mutually reinforcing. The form is also identical in Portuguese *proféta* but there is no evidence this has been borrowed in Swahili which has *nabii*.

ruzario [rùzárĭò̧] appears to be a direct loan from Italian *rosário* [ro'zario]/[ro'zarjo]. Portuguese also has *rosário* but this does not appear to be the source for Swahili *rozari* which is instead from English *rosary*.

seminario [sèmìṇárḭ́ò̰] also appears to be a direct loan from Italian *seminário* (seminary). Portuguese likewise has *seminário* but again it does not appear to be the source for Swahili *seminari* which is from English *seminary*.

semiterio [sḛ̄mḭ̌térḭ́ò̰] is not directly from Italian *cimitéro* (and rarer *cimitério*) 'cemetery' given the phonetic mismatch particularly between initial [s] in Mà'dí and [t͡ʃ] in Italian. One possibility may be phonetic loan convergence between Italian *cimiter(i)o* and English *cemetery* which would then explain the initial [s] in Mà'dí. To complicate matters, we note it matches perfectly with Portuguese *cemitério* (where initial orthographic *c* before *e* is [s]). However, it does not appear to be recorded in Swahili where *makaburini* is the usual form. It may be a previously unrecorded Portuguese loan in Swahili that has since spread elsewhere. Of all the items in this particular category it appears to be the least directly Italian in origin. Given the origins of many Verona fathers, interference from the Veneto forms *simitéro* and *simitério* with initial /s/ is also possible and also cannot be excluded.

There is a small number of items identified as loans from Arabic by Blackings (2011) which are certainly in most cases Italian in origin but transmitted into Mà'dí via Arabic:

(2b) **botoloni** [bòtòlóṇḭ̀], **motoloni** [mò̰tò̰lóṇḭ̀] < Juba Arabic *bantaloon*[6]
   ultimately from Italian *pantalóni* (trousers)
   busita [búsutà] (also **búsutà**, **pósutà**, also **póstà)** < Juba Arabic *busta*
   ultimately from Italian *pósta* (post, post office)
   gonila [gò̰ṇílà] < Juba Arabic *guniila* ultimately from Italian *gonnélla* (skirt)
   salasione [sàlàsḭó̰nḛ̀] (rubber solution for mending punctures) < Arabic
   salsion but ultimately, or more likely, directly from Italian *soluzione*
   (lit. solution)
   susita [súsḭtà] (zip for fastening clothing, alongside **zipu** from English *zip*)
   < Arabic *susta* but ultimately from Italian *sústa* (spring).

These items are used across the Arabic-speaking world and can also be found in languages in contact with Arabic in South Sudan as well as elsewhere in Africa and the Middle East (see, e.g. Baldi 2021). While Blackings (2011) emphasizes Arabic loans in Mà'dí are typically from the Juba creole, Abu-Manga (1991) suggests that such loans enter into South Sudanese languages via a greater range of Arabic varieties, e.g. central Sudanese colloquial Arabic, and Turco-Egyptian Arabic, with the latter in particular associated with administrative terminology.

botoloni [bòtòlóṇḭ̀] **motoloni** [mò̰tò̰lóṇḭ̀] (from It. *pantalóni*) in Baldi (2021): *bantaloon* in Juba Arabic, this appears in different forms across Arabic, e.g. *banṭalūn* in Wehr (1994), *pantalon* / *panâtilîn* in Jullien de Pommerol (1999), *banṭloon* in Abdu (1988). There is also a similar word-initial alternation involving the bilabials, /b m/, in Saho in Eritrea where it appears as

---

6 Juba Arabic forms are taken from Smith and Morris (2005).

banthallooni or *manthalloon* (Banti and Vergari 2008). The loan is also found in other South Sudanese languages, such as *bantalun* in Dinka (Blench 2005).

busita [búsutà] (also **búsutà**, **pósutà**, **póstà)** < Arabic *busta* according to Baldi (2021) who appears to be unaware of its Italian origins – unlike Abdu (1988: 192) who also notes its widespread use throughout the Arabic-speaking world. The postal system in Egypt, as in other parts of Africa, was established by Italians (see, e.g. Moustafa 2019-2020). The transformation of voiceless /p/ to /b/ is typical in Arabic which lacks the former as a phoneme in its sound system. The form **póstà** with retained initial voiceless stop may be from Swahili *posta* which Rajki (2005) identifies as Italian in origin. The word also appears in different South Sudanese languages, including *bosta* in Bari and *posta* in Dinka (Baldi 2021: 31). Forms with the initial /p/ and the vowel /o/, are in our view likely direct loans from Italian (or mediated via Swahili) and have not entered via Arabic.

gonila [gònɪ̃́là] < Arabic *gonella* and *guniila* (from It. *gonnella* 'skirt') according to Baldi (2021), *gonella* in Wehr (1994). There is no doubt that this loan is ultimately Italian in origin.

salasione [sàlàsɪ̀ɔ́nὲ] (rubber solution for mending punctures). While Blackings (2011) thinks this is directly from Arabic *salsion*, we think instead this is much more likely directly from Italian *soluzióne* (lit. 'solution') for a whole series of reasons. In the first instance the Juba Arabic form is given as *selsiyon* by Smith and Ama (2005: 52, 172) who translate it as 'glue' and 'rubber solution'. Stress is on the initial syllable, i.e. *sélsiyon*, according to a predictable rule in Arabic that a single final consonant does not count in the calculation of stress placement resulting in stress shift to the antepenult in CVCVCV(C) forms (see also above). In [sàlàsìɔnὲ] we note that the high tone in Mà'dí matches penultimate stress in the original Italian, alongside the retention of the original final low mid -*e*. While some contaminating influence from Juba Arabic cannot be entirely excluded, if the form was directly from the Arabic form which has a final nasal, the final epenthetic vowel in Mè'dí would predictably be -*i*.

susita [sʊ́sɪ̄tà] (zip). This word has undoubtedly entered Mà'dí via Arabic *susta* 'zip' whch is originally from Italian *sústa* 'spring'. It was borrowed first into Egyptian Arabic[7] before appearing in turn in Juba Arabic. We see in the Mà'dí form elimination of the medial cluster through vowel-insertion. The high vowel on the first vowel matches the stressed syllable in Italian.

To this group we add one more word for which an Italian origin appears less certain:

baburu [bàbʊ́rʊ̀] < Juba Arabic *babuur* (ship, lit. 'steam' from 'steamboat') < ?Italian *vapore* 'steam'

baburu [bàbʊ́rʊ̀] in Mà'dí has two meanings: (a) ship; (b) a small pressure stove which uses paraffin. According to Baldi (2021) it is from Arabic *bābūr* 'locomotive'. He does not identify it as a loan into Arabic. It is also *mabu(w)r* 'ship' in Dinka and Bari (Abu-Manga 1991). The word is widespread across the Arabic-speaking world, including *baabuur* in Libyan Arabic (Abdu 1988), while in the Horn of Africa we find *baabur* in Saho (Banti and Vergari 2008), and *baabuur* in Somali (Tosco 2023);

---

7 https://en.wiktionary.org/wiki/%D8%B3%D9%88%D8%B3%D8%AA%D8%A9.

While Abdu (1988), for instance, considers the source for Libyan Arabic *baabuur* (i.e. *bābūr*) to be Italian *vapóre*, it may be more likely from French *vapeur* given the word-final rhotic that appears in the partly nativized form. If it were from Italian, we might have expected *bābūri* with a retained and raised final vowel.

We also find in Mà'dí a small set of doublet forms – which may point to different source languages – in each case, English and Italian. Four such pairs have been identified:

(3)   Doublets with a potential Italian source:
      baa [báà] (Eng. *bar*)   **bari** [bárì] (?It. *bar* < Eng *bar*)
      bia [bíà], **biya** [bíyà] (Eng. *beer*), **bira** [bírà] (It. *bírra*)
      redio [rẹ́dị̀ọ̀] (Eng. *radio*)   **radio** [rádị̀ọ̀] (?It. *rádio*)
      ziro [zị́rọ̀] (Eng. *zero*)   **zero** [zẹ́rọ̀] / (?It. *zéro*)

With respect to **baa** and **bari**, syllable- and word-final orthographic 'r' in English loans in general does not surface in Mà'dí. In British English, syllable-final 'r' in word-medial position is always silent in that variety and this is always the case in Mà'dí, e.g. **kona** 'corner'. Word-final /r/ in British English only surfaces as a rhotic if the following word is vowel-initial (so-called r-linking) but it does not surface in isolation, before a consonant or phrase finally. As previously noted, this is generally the case in Mà'dí of which ones find many such examples (including previously sited **kona**). The loss of the rhotic also points to oral rather than written transmission of English loans as more likely in Mà'dí. Blackings identifies **baa** as English in origin, but says nothing about **bari**. It is tempting to consider the latter as a loan from Italian *bar* /bar/ (itself borrowed from English *bar*) with word-final vowel epenthesis. However, on occasion one does find other rare examples of apparent final r-retention, often in doublet form, e.g. **leta** [létà] and **letari** [lètàrì] both from English *letter* – pointing to what appears to be occasional borrowing of the written form (rather than only from spoken language). Given the final vowel, the latter is not from Italian *léttera* for which **letera** or **letara** would have been expected. We note English 'bar' also occurs in Juba Arabic as **bar** /bar/ and this is most likely the mediating source.

The possibility that *bar* has been borrowed independently twice (directly or indirectly) from English into Mà'dí is further supported by the almost identical doublet **ka** and **kari** 'car', with and without rhotic retention. These can only have entered directly from English without mediation from Juba Arabic which has an entirely different lexical form for this concept.

In the case of **bia** and **bira**, the former is clearly English in original with regular final r-deletion and open final /a/, while the latter is ultimately from Italian *birra* and refers across large parts of African and the Middle East to bottled beer. It is commonly used across the Arabic-speaking world (Abdu 1988), including in Juba Arabic, and is also found in Amharic and Tigrinya *bira*, in Saho *birra* while in Somali

biir is from English (Tosco 2023). The borrowing of *birra* is widespread across the former Ottoman Empire – hence *bîre* in Kirmancî and *bira* in Zazakî spoken away from Africa in Kurdistan in the Middle East are also identified as being Italian loanwords by Manzelli (2008).

With respect to **redio** and **radio,** given the vowel raising of *a* to *e*, the former is clearly English in origin. With respect to the latter form, this may be a reading pronunciation of English, but it is also consistent with Italian. That said, Baldi (2021) also identifies it as an Arabic loan *rādiyō* into Dinka. He reports it as having been borrowed into Arabic from French, but Abdu (1988) says it is from Italian in Libyan Arabic. A similar vowel alternation is also found in Ethiopian languages (M. Tosco, pers. comm.).

In the case of **ziro** and **zero**, once more given the vowel raising of *e* to *i*, the former is clearly a loan from English. Again, the second form may also be a reading pronunciation of the English written form. However, it is also consistent with the previously noted (a) deaffrication of the initial /dz/ in Italian *zero* [dzero] to [zero], which is also a common phenomenon in the Veneto region where Verona is located; and (b) absence of affricates /ts dz/ in Mà'dí.

There are also a number of words not identified as loans but which may be or are definitely Italian in origin, whether in part or in full, as listed in (4) with suggested Italian source. Each item is discussed in turn, and as will become evident, some are mediated via Arabic.

(4)   **erenia** [ɛ́rɛ́nɨ̏à] (hernia) < *érnia*
      faratelo asi [fàràtɛ̀lɔ̀ ásɨ́] (custard apple) < *fratéllo* 'brother'
      jibiterio [jɨ̀bɨ̀tɛ́rɨ̏ɔ̀] (cemetery) < *cimitéro* and *cimitério*
      katekisimu [kàtàkɨ́sɨ̄mù] / [kàtɛ̀kɨ́sɨ̄mù] (catechism) < ?*catechísmo*
      katoli [kàtɔ́lɨ̏] (Catholic) < *cattólico*
      kirisitiani [kìrìsìtìánì] (Christian) < *cristiáno*
      makanika [màkánɨ̏kà] (mechanic) < ?*meccánico*
      makina [mákìnà] (engine, sewing machine) < *mácchina*
      padere [pádērè] (father) < *pádre*
      pagani [pàgánì] (kaffir) < *pagáni* (pagans, heathen)
      salata [sálātà] (salad) < *insalata* (salad)
      tegula [tégū̀là] (roof tile) < ?*tégola*
      turumba [tùrúmbà] (pump, bicycle pump, petrol station) < *trómba*
      vino [vɨ́nɔ̀] (wine) < *víno*

   erenia [ɛ́rɛ́nɨ̏à] (hernia) is from Italian *érnia* given the absence of the initial [h] found in the English *hernia* and the retention of the syllable-final rhotic. Mà'dí has /h/ in its phonemic system, so its loss would not be expected here. Portuguese also has *hérnia* [ernia] but it is not borrowed into Swahili which has *ngiri* instead.

faratelo asi [fàràtélò̰ ásḭ́] lit. 'brother's heart' (custard apple), given its large size and general heart shape. The first element is clearly from Italian *fratéllo* in the sense of '(religious) Brother'. The English equivalent 'Brother's heart' is also used in Uganda.[8]

jibiterio [jḭ̀bḭ̀térḭ̀ò̰] (cemetery) < Italian *cimitéro/cimitério*. Given the voiced affricate [d͡ʒ] in word initial position, this matches most closely Italian voiceless affricate [t͡ʃ], and is an alternate form to previously discussed **semiterio** [sḛmḭ̀térḭ̀ò̰] which shows evidence of contact with English (or Portuguese and even potentially Veneto). In Italian *cimitero* is normally preferred, with *cimiterio* now generally considered somewhat archaic. The ending is however also reinforced by the Latin form *cimeterium* – which in Italian-style reading also has a voiceless affricate [t͡ʃ].

katekisimu [kàtàkḭ́sḭ́mù] / [kàtḛkḭ́sḭ́mù] (catechism) may be from Italian *catechísmo* given the shared location of the stressed vowel in Mà'dí and in Italian. However, it is not certain if this is ultimately the source. English may also be the source with word-final post-labial addition of -u, but we note stress placement in that language is usually on the first syllable i.e. *cátechism*. On the other hand, Swahili also has *katekisimu* which is identical in form to the Mà'dí form. Baldi (2024: 144) considers Portuguese *catequismo* [kateˈkizmu] (alternative to the more common *catecismo* [kateˈsizmu]) to be the loan source in Central Africa. It is less likely that the Swahili form is from Italian. The Mà'dí and Swahili forms would, however, also be reinforced by Latin *catechismus*, Italian *catechismo* and English *catechism*.

katoli [kàtɔ́lḭ̀] (Catholic) may be from Italian *cattólico*, supported by Latin *cathólicus*. Portuguese also has *católico* but can most likely be excluded as this form does not appear to be the source for Swahili *katoliki* and *mkatoliki* which are from English *Catholic*. It is not at all clear why the form in Mà'dí is reduced but it is also found with the same phonetic form in Acholi, spoken in the same region.

kirisitiani [kḭ̀rḭ̀sḭ̀tḭ̀ánḭ̀] may be from English *Chrístian*, or is some form of Italian *cristiáno*. It may also represent a convergence of both. The final -i in the Mà'dí form could in fact reflect the plural in Italian, i.e. *cristiani* or point to vowel epenthesis after the final alveolar nasal consonant in English. We also note that English has stress on the initial high vowel, while the high tone on the penult in Mà'dí aligns with stress placement on the same penultimate low vowel in Italian. On balance, and in light of shared high tone/stress placement, an Italian origin for the word in Mà'dí seems more likely.

makanika [màkánḭ́kà] (mechanic) is similar to Italian *meccanico*. However, the identical form is also in Swahili where it is considered to be a loan from English *mechanic* (Rajki 2005). Given the shared form, it is more likely to have entered Mà'dí from Swahili, and an Italian origin can be excluded.

makina [mákínà] (engine) is from Arabic *makina, mākīna* according to Baldi (2021). It was previously noted by Roth-Laly (1969) for Arabic as *mákina*, and by Wehr (1994) as *makina* / *mākīna* 'machine', *makana* / *makanât* 'machine, engine' (Jullien de Pommerol 1999). This is a

---

8 https://www.facebook.com/LocalFlavoursJinja/posts/from-the-the-tropical-fruit-farm-seedlings-sold-at-local-flavours-jinja-prices-r/4970848679619107/ (accessed 6 June 2025).

clear Italianism in Arabic, taken from *mácchina* 'machine, car'. It also appears in Ottoman Turkish in identical form, as well as in the Horn of Africa where there has been more direct contact with Italian in the 19th and 20th centuries, e.g., the word for 'car' in Amharic *mäkina* and Tigrinya. *makina* (Beyene 2011: 130), Saho *makiina* and Oromo *makinaa* (alongside native *konkolaataa*). It is also found in other languages spoken in and around South Sudan, such as Acholi, Bari, Dinka, Lotuxo and Ndogo (Baldi 2021: 279).

padere [pádērè] (father) is from Italian *padre* 'father' but also reinforced by Swahili *padre, padri, padiri* from Portuguese *padre.* The final *-e* in the Mà'dí is consistent with an Italian origin, and given the specific nature of the term is also consistent with the borrowing from Italian of other role-related religious terms such as *suora,* and *fratello.*

pagani [pàgánì] (kaffir, i.e. pagan, heathen) appears to be another example of an Italian loan in which the plural form serves as the source, i.e. *pagáni* (pagans, heathen, pl, rather than the singular *pagáno*) with stress on the same syllable as the high tone in the Mà'dí form.

salata [sálātà] (salad) is a common Italianism which is found in different varieties of Arabic (e.g. Cairene Arabic. See Moustafa 2019-2020), as well as in languages of the Horn of Africa (including Amharic, Tigrinya and Saho) as well as in Central Europe (e.g. Slovene, Croatian). It is a reduced form of original *insalata* in which the initial syllable appears to have been reanalysed as the preposition *in.* In the case of Mà'dí the word has clearly been mediated through Arabic. This is evident in the regular shift in stress placement from the original penult to the antepenult in CVCVCV forms where is CV is a light syllable, i.e. from Italian *(in)saláta* to Arabic *sálata.* The result is a predictable high tone in the initial syllable in the Mà'dí form.

tegula [tég̱ūlà] (roof tile). It is tempting to link this to Italian *tégola* as it finds no match in other potential source languages such as Swahili (*vigae*) or Portuguese (*telha*). However, it does exist in some varieties of English specifically as a Roman-style roof tile and it also found in Luganda *akategula* and *e'ttegula.* The word can also be used in Uganda specifically with reference to church roofing.[9] Given the identical form, this word may potentially be a loan from Latin *tegula*, although Italian *tegola* still seems more likely.

turumba [tùrúmbà] (pump) is borrowed from Arabic as indicated by Baldi (2021) who notes it is also found in Dinka, a major South Sudanese language. However, it is ultimately Italian in origin, from *tromba* 'trumpet, horn' i.e. by extension 'a (water-)blowing instrument' (cf. *tromp* in English). The original Italian word is also widely borrowed across northern Africa and the Middle East. Abdu (1988: 238), for instance, lists it as an Italian loan in Libyan Arabic with various related meanings (e.g. horn, suction pump, trumpet, pump, plunger).

vino [vɪ́n�ò] (wine) is clearly from Italian *vino* which it matches perfectly. It is reinforced by Latin *vinus* (also with initial [v] in Italian-style pronunciation) and is not from Portuguese *vinho* (with intervocalic palatal nasal [ɲ]) which is retained in Swahili *mvinyo*. The item plays an important part in Catholic liturgy.

---

9 https://www.anglicannews.org/news/2014/12/namirembe-diocese-celebrates-the-completion-of-project-ttegula.aspx.

Finally, we note a small number of religious terms that have been borrowed not from Italian but come instead from Latin. These items reflect the previously noted importance of that language in the Catholic Church, particularly before the so-called Vatican II reforms of the 1960s which allowed for the liturgy to be celebrated in languages other than Latin. They are:

(5) **ekelesia** [ɛ̀kɛ̀lɛ́sɪ̄à] (church) < Latin *ecclésia* (ultimately Greek in origin)
episikopi [ɛ̀pɪ̀sɪ̀kópɪ̀] (bishop) < Latin *episcopus* (ultimately Greek in origin)
pasika [pásɪ̄kà] (Easter) < Latin *páscha* (from Greek but ultimately Hebrew in origin)

Neither of the first two words exists in sufficiently similar form in Swahili (*kanisa, askofu* respectively), Portuguese (*gereja, bispo*) or Italian (*chiesa, vescovo*).

> **ekelesia** or similar is found (alongside *ekkanisa* from Arabic *kanīsa*) in Luganda (Uganda) for instance, but it also appears elsewhere in Africa as far south as languages in Zambia, e.g. Bemba (Bantu) *ekelesha*. It also found elsewhere in the world, including in the Pacific, e.g. Samoan and Wallisian *ekelesia*.

> **episikopi** is also found in other languages in Africa, including in Ateso (Eastern Nilotic) spoken in Eastern Uganda, Nyanja (also known as Chichewa, Bantu) spoken in Malawi and Zambia, as well as Kinyarwanda and Kirundi (Bantu) spoken in Rwanda and Burundi respectively. Unusually in the case of Mà'dí the form shows unexpected stress shift from the antepenult in Latin (*episcopus*) to the penult given its unexpected high tone, i.e. [ɛ̀pɪ̀sɪ̀kópɪ̀].

> **pasika** (and related *pasaka*) is also widespread across Africa, e.g. in Lingala (spoken in large parts of Central Africa), as well as Kirundi and many other languages. The word in Mà'dí cannot be from Swahili given it has only *pasaka* which Rajki (2005) considers to be from Portuguese *páscoa* (and ultimately from Hebrew *pesah*). But a Portuguese source also appears unnecessary given the closer phonetic relationship of *pasika* (and *pasaka*) with Latin *pascha* ['paska] than with Portuguese *páscoa* ['paskua ~ 'paskwa] for which we might expect *pasikoa ~ pasakoa* or similar in Mà'dí, Swahili and elsewhere. A similar explanation excludes Italian *pásqua* ['paskwa] as a source for the Mà'dí word.

Finally, we have one additional religious term: **misa** [mɪ́sà] (mass). This is not from Italian *messa*, but has two potential converging origins. It is more likely to be a direct loan from Latin *missa*, and reinforced by Portuguese *missa* via Swahili *misa*. On the other hand, it is also possible that it has entered into Mà'dí via Swahili (with ultimately Portuguese origins) and in turn reinforced by Latin. This second pathway seems less likely to us, given the need to use the term from earliest missionization in South Sudan in the 19th century – before contact with Swahili through movement across the border into Uganda in the second half of the 20th century.

4 Italian loanwords in Mà'dí

4.1 Phonetic and phonological system of Mà'dí language

Before proceeding to a summary overview of the various phonetic and phonological processes evident in the nativization of potential Italian loanwords in Mà'dí, we present the phonological system of the Lokai dialect of Mà'dí based on Blackings and Fabb (2003).

The vowel system includes four high vowels (non-back, /ɨ/ and /i/ and back /ʊ/ and /u/), four mid vowels (non-back /ɛ/ and /e/ and back /ɔ/ and /o/) and two low vowels (non-back /a/ and /ʌ/), also distinguished according to the Advanced (+ATR) and Retracted Tongue Root (-ATR).[10] The language also has four tones (high, mid, low, high-low).[11]

The phonological consonant system of the Mà'dí language is quite rich and shows particular overlap with the Italian phonological one. There are seven plain stops (/p/, /b/, /t/, /d/, /k/, /g/, /ʔ/) and three implosive stops (/ɓ/, /ɗ/, /ʄ/). There are four plain nasals (/m/, /n/, /ɲ/, /ŋ/), five fricatives (/f/, /v/, /s/, /z/, /h/) as well as two affricates (/t͡ʃ/, /d͡ʒ/), one lateral /l/, one trill /r/ and two glides (/j/, /w/). There are also five prenasalized phonemes (/mb/, /nd/, /ŋg/, /ɱv/, /nd͡ʒ/). Only a small number of Italian phonemes are missing from Mà'dí: the alveolar affricates /ts, dz/, post-alveolar fricative /ʃ/ and the palatal lateral approximant /ʎ/. The phonological system also includes double and secondary articulated sounds with some additional labial articulation, such as the following stops (/kp/, /gb/, /tʷ/, /dʷ/, /kʷ/, /gʷ/, /ʔʷ/), prenasalized (/mgb/, /ndʷ/, /ŋgʷ/) and implosive (/ʄɓ/, /ɓʷ/), as well as two nasals (/ŋm/, /mʷ/), a fricative (/hʷ/), an affricate (/t͡ʃʷ/), a lateral (/lʷ/) and a trill (/rʷ/).

4.2 Phonological adaptation of Italian (and other) loanwords in Mà'dí

When considering the nativization of borrowed forms we have explored above, amongst the most salient processes we note the following:

---

10 This distinction is also orthographically marked (e.g. [ạ] vs [a] where the diacritic indicates the +ATR).

11 Note that in the phonetic transcription used in the dictionary by Blackings (2011), the symbols [c], [j], [ny], [ng] and [y] correspond respectively to the IPA sounds [t͡ʃ], [d͡ʒ], [ɲ], [ŋ] and [j]. Moreover, the tongue root advancement of a vowel is indicated with the diacritic under the symbol [ạ], unlike IPA which uses [a�].

sonorisation or voicing: *assunta* [as'sunta] → **asunda** (alongside retention in **asunta**), *cimiterio* [t͡ʃimiˈtɛːrjo] → **jibiterio** [jɪ̀bɪ̀tɛ́rɪ̯̀ò]);

degemination: *assunta* → **asunda**, **asunta**, *fratello* → **faratelo**, *grazia* → **girasia**, *macchina* ['makkina] → **makina**, *birra* → **bira**, *cattolico* → **katoli**. There are no long consonants in Mà'dí and degemination is regular;

vowel epenthesis: *fratello* → **faratelo**, *altare* → **alitari**, *grazia* → **girasia**, *croce* → **kuruce** [kùrúcè̩], profeta → **purofeta**, → **matirimonio**, *padre* → **padere**, *catechismo*→ **katekisimu** [kàtèki̩si̩mù], *tromba* → **turumba**. The result is a general avoidance of obstruent + liquid (e.g. /fr, dr/, and other clusters, e.g. /lt, sm/. A rare counter-example to this rule is the presence of the variant **posta** alongside **busita;**

final vowel epenthesis: *bar* → **bari**. No word in Mà'dí ends in a consonant so the epenthesis of a vowel after the final consonant in a loan is normal. As previously noted, the identity of the epenthetic vowel depends on the place of articulation of the preceding consonant: -u if it is bilabial, otherwise -i, e.g. **bomu** < *bomb*, **filimu** < *film,* and **begi** < *bag;*

resyllabification of diphthongs: *suora* ['swɔre] → **suere** and **suwere**; both forms are unexpected as /Cw/ clusters are permitted in Mà'dí, e.g. **swe** 'to become fat';

vowel assimilation: *diluvio* → **dulúvio**; *problem* → **peregelému**. In these cases, we find anticipatory vowel assimilation or harmony before the vowel bearing a high tone;

vowel raising: o > u: *rosario* → **ruzario**, *catechismo* → **katekisimu**, *croce* → **kuruce**, *tromba* → **turumba** [tùrúmbà], *battesimo* → **batisimu**; and e > i: *altare* → **alitari**, *natale* → **natali**;

palatalization: (*Natale* [na'taːle] → **nyatali** [nyàtàlì]); this is unexpected given /n/ is a phoneme in Mà'dí;

deaffrication: *zero* ['d͡zɛːro] → **zero** [zɛ́rò], *penitenza* [peni'tɛnt͡sa] → **penitesia**; Mà'dí has neither /t͡s/ nor /d͡z/ in its system. Simplification here is further supported by the fact that in many northern Italian languages including Veneto, /d͡z/ and /t͡s/ are typically also reduced to /z/ and /s/ respectively;

denasalisation: *cimiterio* → **jibiterio** [jɪ̀bɪ̀tɛ́rɪ̯̀ò];

stress position and tone assignment: we have already pointed out the close correlation between tone assignment in Mà'dí and stress placement in the source languages. This has proven itself to be very helpful in determining the loan source. Given the stressed syllable of the original word usually receives a high tone in Mà'dí (or more rarely a mid tone), it makes it possible to identify with greater certainty whether the word derives from Italian ([kàndélà] < *candéla*, [kɪ̀-ìsìtiánì] < *cristiáni*) or English ([kàtɪ́dūrālì] < *cathédral*, [kɒ̀mónɪ̯̀ò] < *commúnion*). We have noted above a number of cases where loanwords appear to have Italian-style (and also Portuguese-style) stress placement, marked by Mà'dí high tone on the same syllable of the Italian, but not the English word, e.g. [bàtɪ́si̩mù], [kàtɒ́lì], [mákìnà], [matɪ̀rɪ̀mónɪ̯̀ó], [pàràdɪ́zò], [pēnìtésīà], [pùròféta], [rùzárɪ̯̀ò], [sɛ̀mìnárɪ̯̀ò]. In very rare cases, there is no high or mid tone assigned to the nativized loan, e.g. **nyatali** [nyàtàlì] which has only low tones. It is not clear why this might be the case in such words.

In a number of other cases there is no or little change in segmental form, e.g. *papa* > **papa** [pápà], *paradiso* [para'di:zo] > **paradizo** [pàràd́ɩ̀zɔ̀], *seminario* [semi'na:rjo] > **seminario** [sɛ̀m̀ɩ̀nárɩ̀̀ɔ̀], *colera* [ko'lɛ:ra] > **kolera** [kòlérà], *radio* ['ra:djo] > **radio** [rádɩ̀̀ò], *vino* ['vi:no] > **vino** [v́ɩ̀nɔ̀].

4.3 Some semantic considerations about Italian loanwords in Mà'dí

The semantic analysis of loanwords in a language allows a better understanding of the relationships established between the people in contact. In the case of Mà'dí, the most marked lexical imprint is certainly that of the Comboni missionaries with the consequent presence of loanwords (nouns only) linked to the religious sphere, regardless of whether they are Italian, Latin or even Portuguese in origin. In fact, words are related to Catholic feasts and rites functions (e.g. **asunda**, **batisimu**, **katekisimu**, **matrimonio**, **misa**, **natali** ~ **nyatali**, **pasika**, **penitesia**), items used in religious activities (**alitari**, **ruzario**, **vino**), religious titles and names (**faratelo**, **papa**, **padere**, **purofeta**, **suere**/**suwere**), and places of religious activity (**jibiterio**/**semiterio**, **seminario**), Biblical or other religious terms and concepts (**duluvio**, **girasia**, **paradizo**).

In other cases, items identified as Italianisms are occasionally related to health conditions, i.e. **azima** and **erenia**, and slightly more often to everyday services and items, such as **botoloni** (trousers), **busita** (post) and **gonila** (skirt), as well as technology, e.g. **turumba** and **makina**. With respect to the last category these are borrowings from Italian that are also common in the wider Arabic-speaking world and have entered Mà'dí in those forms via Arabic.

5 Discussion and conclusions

While we are not the first to note the presence of Italian loanwords in Mà'dí, ours is the first to explore them in detail in that language (or indeed in any South Sudanese language). It is clear that items we have identified as potential Italianisms fall into different categories of certainty – something that is not unexpected given the complex linguistic situation Mà'dí speakers find themselves in – in particular with respect to important superstrate languages such as English, Arabic (including in particular Juba Arabic), Swahili and Italian, amongst others.

Having now considered in detail the potential sources for a range of lexical items in Mà'dí, an Italian origin is clear for a number of them, such as **faradelo** 'brother', **suwere** 'sister' and **natali** 'Christmas'. At the same time, the situation is a complex

one – with a number of different pathways, competing source languages, as well as patterns of potential reinforcement and convergence of forms from different source languages, e.g. English, Italian, Latin and Portuguese (via Swahili). It is not surprising then if Blackings (2011) often ascribes a different origin to some of the Italian loanwords that we have been able to identify with certainty, e.g. **kandela**, **paradizo**, **seminario**, which he treats as Anglicisms. In some cases, Italian origin is in our view less certain or possibly unlikely, e.g. **baburu**. In other cases again, it is hard to choose one language over another as the source.

The religious domain is not surprisingly the one in which most Italianisms are found. The presence of the Comboni missionaries (or Verona Fathers) has been pivotal in Catholic evangelization throughout South Sudan, Uganda and elsewhere in Africa. A striking feature of some of these loans is what appears to be the borrowing of the Italian plural form, e.g. **suwere** 'nun' < *suore* (f. pl), and **pagani** 'kaffir' < *pagani* (m. pl). This is likely a reflection of a more frequent use of those particular forms rather than in the singular in everyday discourse. On the other hand, it is perhaps a little surprising that two specific medical conditions **erenia** and **azima** have been borrowed: they appear by all indications to be solid Italianisms. But Italian missionaries, amongst their many tasks over many decades, will also have provided basic medical assistance.

Of course, future access to additional Mà'dí lexical material is likely to add to the list of Italianisms in that language. Given the specific context here involving linguistic contact through missionization, we are mindful that there is generally very little inclusion of religious vocabulary in dictionaries of lesser described languages. This lack of information means of course that the true extent of borrowing from Italian through contact with Italian missionaries, such as the Verona Fathers, cannot be properly determined.

That said, future work on Italianisms in South Sudan and surrounding areas should nevertheless also be extended beyond the Mà'dí language. The wide presence of the Comboni mission across Central and Southern Africa means that Italianisms are likely to be found across many different languages of these regions, something Storch (2014) has, for instance, already identified for Luwo.

References

Abdelhay, Ashraf, Al-Amin Abu-Manga & Catherine Miller. 2015. Language policy and planning in Sudan: From local vernaculars to national languages. In Barbara Casciarri, Munzoul A. M. Assal & François Ireton (eds.), *Multidimensional change in Sudan (1989–2011): Reshaping livelihoods, conflicts and identities*, 263–280. New York: Bergbahn Books.

Abdelhay, Ashraf, Busi Makoni & Sinfree Makoni. 2016. The colonial linguistics of governance in Sudan: The Rejaf Language Conference, 1928. *Journal of African Cultural Studies* 28(3). 343–358.

Abdu, Hussein Ramadan. 1988. *Italian loanwords in colloquial Libyan Arabic as spoken in the Tripoli region*. PhD thesis. Tucson: University of Arizona.

Abu-Manga, Al-Amin. 1991. Arabic loanwords in the Nilotic languages of the Southern Sudan. In Lionel Bender (ed.), *Proceedings of 4th Fourth Nilo-Saharan Linguistics Colloquium*, 133–145. Hamburg: Helmut Buske Verlag.

Agostoni, Arcisio. 1996. *History of the Institute of the Comboni Missionaries*. Rome: Comboni Missionaries of the Heart of Jesus.

Alcaide, Jorge Carlos Naranjo. 2019. The development of Catholic schools in the Republic of Sudan. *Social and Education History* 8(1). 83–111.

Amone, Charles. 2021. Ethnic pluralism and the challenge of thematic curriculum implementation in Uganda. *Journal of Multilingual and Multicultural Development* 42(1). 52–63.

Baldi, Sergio. 2011. Swahili: A donor language. *Lingua Posnaniensis* 53(1). 7–24.

Baldi, Sergio. 2018. Portuguese loans in some African languages. *Annali, Sezione orientale* 78. 19–44.

Baldi, Sergio. 2021. *Dictionary of Arabic loanwords in the languages of Central and East Africa*. Leiden: Brill.

Baldi, Sergio. 2024. *Dictionary of Portuguese loanwords in the languages of Sub-Saharan Africa*. Leiden: Brill.

Banti, Giorgio & Moreno Vergari. 2008. Italianismi lessicali in saho. *Ethnorêma* 4. 67–93.

Beyene, Yaqob. 2011. I prestiti italiani in amarico e tigrino. *Rassegna di Studi Etiopici*. *Nuova Serie* 3(46). 97–140.

Bilbao, Osmundo & Giovanni Ferrazin. n.d. *Mà'dí English dictionary*. Unpublished manuscript.

Bilbao, Osmundo & Luigi Moizi. 1984. *Mà'dí-English dictionary*. Limone: Casa Comboni.

Blackings, Mairi. 2011. *Mà'dí-English dictionary*, 2nd edn. Munich: Lincom.

Blackings, Mairi & Nigel Fabb. 2003. *A grammar of Mà'dí*. Berlin: Walter de Gruyter Mouton.

Blench, Roger. 2005. *Dinka-English and English-Dinka dictionary*. Unpublished pdf.

Boone, Douglas & L. Watson Richard. 1996. *Moru-Mà'dí survey report*. Khartoum: SIL-Summer Institute of Linguistics, Sudan Branch.

Butros, Albert. 1973. Turkish, Italian and French loanwords in the Colloquial Arabic of Palestine and Jordan. *Studies in Linguistics* 23. 87–104.

Chemello, Francesco. 2021. *Una grande storia d'amore. La missione comboniana in Sud Sudan dal 1857 al 2017*. Rome: Biblioteca Colombiana.

Cifoletti, Guido. 1994. Italianismi nel dialetto di Tunisi. In Dominique Caubet & Martine Vanhove (eds.), *Actes des premières journées internationales de dialectologie arabe de Paris*, 451–458. Paris: Langues' O.

Crazzolara, Joseph Pasquale. 1960. *A study of the Logbara (Mà'dí) language: Grammar and vocabulary*. Oxford: Oxford University Press.

D'Anna, Luca. 2018. Phonetical and morphological remarks on the adaptation of Italian loanwords in Libyan Arabic. In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 171–187. Amsterdam & Philadelphia: John Benjamins.

Elrayah, Tagelsir. 1999. Arabic and English in Sudan: 1821–1985. *Islamic Studies* 38(4). 603–620.

Ferrazin, John. 1996. *Ma 'di English-English Ma 'di Dictionary*. Unpublished manuscript.

Heugh, Kathleen & Bwanika Mathias Mulumba. 2014. *Implementing local languages medium education in the early primary curriculum of Ugandan schools: A Literacy and Adult Basic Education (LABE) intervention in six districts in north and north west Uganda*. Kampala: Literacy and Adult Basic Education (LABE).

Jullien de Pommerol, Patrice. 1999. *Dictionnaire arabe tchadien-français suivi d'un index français-arabe et d'un index des racines arabes*. Paris: Karthala.

Kutsch Lojenga, Constance. 1994. *Ngiti: A Central-Sudanic language of Zaire*. Cologne: Rüdiger Köppe.

Manfredi, Stefano. 2017. *Árabi Júba: Un pidgin-créole du Soudan du Sud*. Leuven & Paris: Peeters.

Manfredi, Stefano & Sara Petrollino. 2013. Juba Arabic. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages. vol. 3: Contact languages based on languages from Africa, Asia, Australia, and the Americas*, 54–65. Oxford: Oxford University Press.

Manzelli, Gianguido. 2008. Gli italianismi di mediazione turca nel curdo della Turchia e dell'Iraq. In Lazzeroni Romano, Banfi Emanuele, Bernini Giuliano, Chini Marina & Marotta Giovanna (eds.), *Diachronica et Synchronica*, 301–326. Pisa: Edizioni ETS.

Mioni, Alberto. 1988. Italian and English loanwords in Somali. In Annarita Puglielli (ed.), *Proceedings of the Third International Congress of Somali Studies*, 36–42. Rome: Il Pensiero Scientifico.

Molinaro, Luigi. 1925. *Appunti di grammatica della lingua madi*. Verona: Istituto missioni africane.

Molloy, Peter. 1957. *The cry of the Fish Eagle; The personal experiences of a game warden and his wife in the Southern Sudan*. London: Joseph.

Moustafa, Ahmed. 2019-20. *Gli italianismi nel dialetto del Cairo: Una lunga storia di contatti*. Master's thesis, University of Padua.

Musgrave, Simon & John Hajek. 2014. Linguistic diversity and early language maintenance efforts in a recent migrant community in Australia: Sudanese languages, their speakers and the challenge of engagement. In John Hajek & Yvette Slaughter (eds.), *Challenging the monolingual mindset*, 113–130. Bristol: Multilingual Matters.

Okeny, Kenneth. 1991. The 1947 Juba Conference. *Northeast African Studies* 13(1). 39–58.

Rajki, Andras. 2005. *Swahili Dictionary (with etymologies)*. Unpublished ms. Available at: https://www.academia.edu/12788108/Swahili_Etymological_Dictionary (accessed 05 February 2024).

Roth-Laly, Arlette. 1969–1971–1972. *Lexique des parlers arabes tchado-soudanais*. Paris: CNRS. [Fascicule i (1969), 106 pp.; ii (1969), pp. 107–264; iii (1971), pp. 265–402; iv (1972), pp. 403–545].

Simeone-Senelle, Marie-Claude. 2018. Arabic on the Dahlak Islands (Eritrea). In Stefano Manfredi & Mauro Tosco (eds.), *Arabic in contact*, 111–134. Amsterdam & Philadelphia: John Benjamins.

Smith, Ian & Timothy Ama Morris. 2005. *Juba Arabic–English dictionary. Kamuus ta Arabi Juba wa Ingiliizi*. Kampala: Fountain Publishers.

Storch, Anne. 2014. *A grammar of Luwo*. Amsterdam: John Benjamins.

Tosco, Mauro. 2008. A case of weak Romancisation: Italian in East Africa. In Thomas Stolz, Dik Bakker & Rosa Salas Palomo (eds.), *Aspects of language contact. New theoretical, methodological and empirical findings with special focus on Romancisation processes*, 377–398. Berlin: de Gruyter Mouton.

Tosco, Mauro. 2009. Loanwords in Gawwada, a Cushitic language of Ethiopia. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages. A comparative handbook*, 124–141. Berlin: Walter de Gruyter Mouton.

Tosco, Mauro. 2023. Eritrea, Ethiopia, and Somalia. In Ursula Reutner (ed.), *Manual of Romance languages in Africa*, 787–807. Berlin: Walter de Gruyter Mouton.

Tosco, Mauro & Stefano Manfredi. 2018. Juba Arabic (Árabi Júba): A "less indigenous" language of South Sudan. *Sociolinguistic Studies* 12(2). 209–230.

Tucker, Archibald Norman. 1940. *The Eastern Sudanic languages*. London: Oxford University Press.

Tucker, Archibald Norman & Margaret Arminel Bryan. 1966. *Linguistic analyses, the non-Bantu languages of north-eastern Africa*, 2nd edn. London: Oxford University Press.

UNICEF. 2016. *South Sudan. The impact of language policy and practice on children's learning: Evidence from Eastern and Southern Africa* (country review), 4pp.

Wehr, Hans. 1994. In J. Milton Cowan (ed.), *A dictionary of modern written Arabic: (Arabic-English)*. Ithaca, N.Y: Spoken Language Services.